



KTH Teknikvetenskap

Teknisk Fotografi

Teori och Tillämpningar

Kjell Carlsson



Hålkamera anno 1544



Digitalkamera anno 2015

Kompendium för kurs SK1140, Fotografi för Medieteknik.

©2021 Tillämpad Fysik, KTH

Senaste revision: 2021-01-25

Förord

Detta kompendium bygger på ett tidigare kompendium, ”Teknisk fotografi”, som har använts i fotoundervisningen vid KTH under många år. Det tidigare kompendiet innefattade både analog (film-) fotografi och digital fotografi. Den analoga fotografin har numera gått i limes mot noll, och dessutom har det skett en otroligt snabb utveckling på digitalfotofronten. Detta har medfört att stora delar av det gamla kompendiet blivit föråldrat och behövde revideras. Vid revisionen har delarna om analog fotografi tagits bort, och digitalfotobiten har uppdaterats och utökats. Många delar av det gamla kompendiet som var av allmängiltig karaktär (t.ex. optisk avbildning, skärpedjup mm) har behållits i det nya kompendiet. Vissa avsnitt är av överkurskaraktär, eller behandlar områden som inte ingår i KTH-kursen SK1140. Dessa avsnitt kan vara användbara om man vill tränga djupare in i vissa områden, eller om man vid senare tillfällen behöver mera information. Vilka delar av kompendiet som är av vikt vid examination på KTH-kursen SK1140 framgår av den information som ges vid kursstart.

Ursprungligen trycktes kompendiet upp inför varje kursstart och såldes till studenterna. Numera ges det ut endast i elektronisk form (som pdf), vilket har flera fördelar. Dels kan man snabbt och enkelt göra korrigeringar och förbättringar, dels kan man utnyttja färgillustrationer utan dyra tryckkostnader, och dessutom är det gratis för studenterna. Slutligen vill jag tacka Linda Lundström för korrekturläsning av kompendiet samt värdefulla synpunkter på innehållet. Om någon läsare upptäcker några felaktigheter (t.ex. tyckfel), eller har synpunkter på innehållet, går det mycket bra att ta kontakt via e-post (adress finns nedan).

Stockholm, 2019-01-31

Kjell Carlsson

e-post: kjellc@kth.se

Innehållsförteckning

Kap. 1. Historisk introduktion	7
Kap. 2. Hålkameran	10
Kap. 3. Kameratyper	15
Kap. 4. Bländaren och slutaren	19
Kap. 5. Optisk avbildning	24
Kap. 6. Objektivtyper	30
Kap. 7. Perspektiv	41
Kap. 8. Stereofotografi	46
Kap. 9. Fotometri	50
Kap. 10. Skärpedjup	65
Kap. 11. Exponeringsmätare	76
Kap. 12. Färglära	79
Kap. 13. Ljuskällor	94
Kap. 14. Matrissensorer för digital fotografi	103
Kap. 15. Optiska avbildningsfel och diffraktion	157
Kap. 16. Bildkvalitet: Upplösningsförmåga	167
Kap. 17. Bildkvalitet: MTF	175
Kap. 18. Infrarödfotografi	225
Kap. 19. Ultraviolett och fluorescensfotografi	230
Kap. 20. Röntgenfotografi	234
Kap. 21. Höghastighetsfotografi	240
Kap. 22. Sampling	247
Kap. 23. Lagring och utskrift av digitalbilder	269
Kap. 24. Automatik	277
Kap. 25. Termovision (värmekamera)	283
Referenser	286
Index	287

1. Historisk introduktion

Ordet fotografi (från grekiskan) betyder **ljusskrift**, vilket innebär att man låter ljuset självt rita upp en bild åt sig istället för att göra det manuellt. Men låt oss först gå tillbaka i tiden till den pre-fotografiska eran för att bättre förstå drivkraften bakom utvecklingen av fotografin.

Människan har sedan urminnes tider, med större eller mindre framgång, avbildat sin omgivning i form av teckningar och målningar. Detta är dock förenat med vissa problem:

1. Det tar lång tid.
2. Man måste ha konstnärlig talang.
3. Bilden blir ofta förvrängd på olika sätt.

Ett hjälpmedel som började användas redan för cirka 1000 år sedan, speciellt vid studier av sol- och månförmörkelser, var en så kallad **camera obscura** (latin ”mörk kammare”), vilken helt enkelt bestod av ett ljustätt rum med ett litet hål i den vägg som vetter mot motivet. På motsatta väggen projiceras då en upp-och-nedvänd bild av motivet. Där kan man placera t.ex. ett papper och helt enkelt fylla i motivets konturer. Ett stort problem var dock att den optiska bilden var mycket ljussvag, varför bara ljusstarka motiv kunde avbildas.

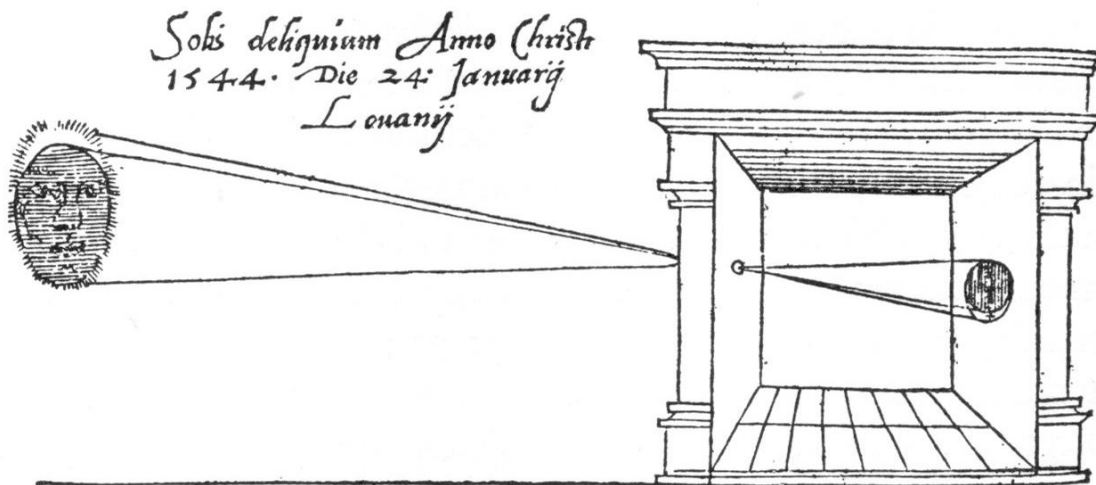


Fig. 1.1. En hålkamera används för att studera en solförmörkelse år 1544.
(från: Gemma Frisius, "De radio astronomico et geometrico liber," 1545)

Nästa steg i utvecklingen skedde i mitten av 1500-talet genom att det lilla hålet ersattes av en lins, varigenom den optiska bilden blev mycket ljusstarkare. Detta ökade användningsområdet kraftigt, och gjorde att man började använda en portabel variant av camera obscura på resor, ungefär som turisten idag är utrustad med en kamera.

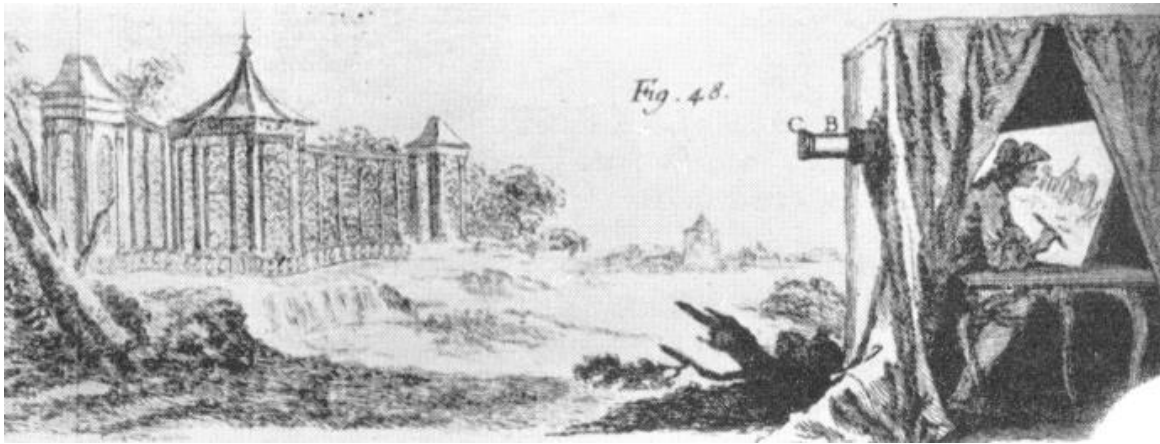


Fig. 1.2. Kameratält för ritändamål (efter Saverien 1753). (Bilden finns reproducerad i H. Bäckström, "Fotografisk Handbok," som innehåller ett fylligt kapitel om fotografins historia.)

Eftersom lättja är ett utmärkande drag hos människan, är det helt naturligt att man började ställa sig frågan om man inte kunde undvika det manuella ritarbetet, och helt enkelt låta ljuset självt "rita" upp bilden. Därmed satte jakten på ljuskänsliga fotografiska material fart. Detta blev en ganska långdragen process som utspelade sig framförallt under 1800-talet. Utrymmet medger här bara några mycket korta notiser från denna utveckling. (Den intresserade hänvisas till annan litteratur, t.ex. H. Bäckström, *Fotografisk Handbok*, eller H. Gernsheim & A. Gernsheim, *The History of Photography*.)

1700-talet: Schultze (Tyskland) och Scheele (Sverige) undersöker hur silversalter mörknar under inverkan av ljus. Inga egentliga fotografiska experiment utförs.

1820-talet: Niepce (Frankrike) lyckas framställa fotografiska bilder med hjälp av asfaltbelagda plåtar som "framkallas" med bensen. Mycket lång exponeringstid (ca. 8 timmar).

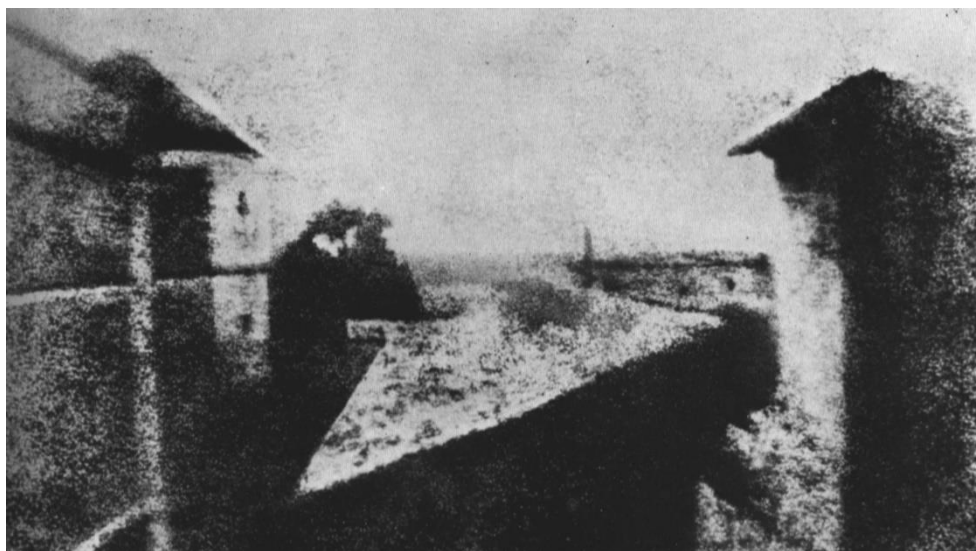


Fig. 1.3. Världens äldsta bevarade fotografi, taget av Nicéphore Niepce, 1827.

1829: Niepce inleder samarbete med Daguerre (Frankrike). Silverbaserade material i kombination med framkallning (i kvicksilverånga!) gör att exponeringstiderna kraftigt kan förkortas (några minuter räcker).

1839: "Daguerrotypien," som bygger på ovanstående framsteg offentliggörs och lanseras kommersiellt. Stor succé! Men svagheter finns, kanske framför allt att man bara får en enda bild, kopiering är inte möjlig. (Intressant historiskt material finns på www.daguerre.org)

1830-40 talen: Talbot (England) utvecklar negativ/positiv förfarandet. Kopiering i godtyckligt antal blir möjligt.

1880-talet: Rullfilmen utvecklas i USA av Goodwin och Eastman (fast inte i samarbete, utan tvärtom!). Tidigare hade fotografiska glasplåtar använts.

1888: Eastman lanserar den första massproducerade billiga lådkameran för rullfilm. Han kallar kameran "Kodak". Stor försäljningssuccé!

Man experimenterade med färgfotografi redan under 1800-talet. Första färgfotot framställdes 1861 med en metod som föreslagits 1855 av James Clerk Maxwell (ja just den store Maxwell, han med ekvationerna!)*. Färgfilm av modern typ började användas på 1930-talet (Kodachrome introduceras 1935), men kom inte i mer allmänt bruk förrän på 1950-talet. Amatörfotografen förenklades successivt genom att kamerorna automatiserades allt mer och gjordes mindre.

Mot slutet av 1900-talet började smalfilm (8 och 16 mm) allt mer ersättas av videoteknik baserad på matrissensorer i halvledarmaterial (se kap. 14). En prototyp till den första stillbildskameran utrustad med matrissensor (Sony Mavica) presenterades 1981, men bilderna lagrades fortfarande analogt enligt videostandard och antalet pixlar var bara 570 x 490. Det dröjde till 1997 innan en Mavica med digital lagring av bilderna fanns tillgänglig på marknaden. Sedan gick utvecklingen av digitalkameror mycket hastigt framåt med allt större bildsensorer och ökat antal pixlar. De från början höga priserna på digitalkameror sjönk också kraftigt, vilket gjorde att fotografering med film blev allt ovanligare för att slutligen praktiskt taget försvinna.

Slutligen kan nämnas att fotografiundervisning och forskning har en lång tradition på KTH. Det började med att John Hertzberg anställdes som lärare 1915. I början av 1930-talet utfördes under ledning av Hertzberg en berömd insats. Man lyckades då framkalla de filmer som hittats efter Andrée-expeditionen i slutet av 1800-talet. Under en period (1948-1958) var Helmer Bäckström professor i fotografi vid KTH, och det fanns då en egen institution för fotografi.

* Egentligen borde detta ha varit omöjligt, eftersom dåtidens emulsioner var helt okänsliga för rött ljus. Man tror att "rödeffekten" i verkligheten berodde på att motivets röda färg även reflekterade ultraviolett, vilket emulsionerna var känsliga för.

2. Hålkameran

Såsom omtalats i föregående kapitel var de tidigaste kamerorna inte utrustade med någon optik, utan avbildningen åstadkoms genom att ljuset passerade genom ett mycket litet hål. Genom experiment med sådana "hålobjektiv" finner man att bilden naturligt nog blir ljusstarkare ju större hålet är. Tyvärr blir den också suddigare om hålet görs stort. Detta är en naturlig konsekvens av ljusets rätlinjiga utbredning, vilket framgår av Fig. 2.1. För stora motivavstånd kan vi förvänta oss att varje punkt på det avbildade föremålet avbildas som en rund fläck med storlek $\approx D$, där D är håldiametern. (Som vi ska se lite senare kan fläckens storlek bli märkbart mindre än D .)

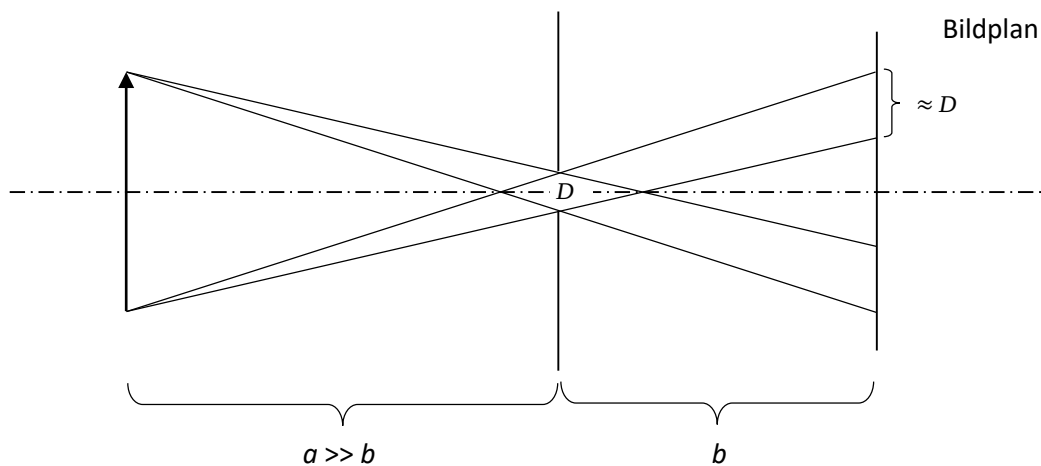


Fig. 2.1. Avbildning i hålobjektiv utgående från geometrisk optik. I verkligheten kommer diffraktion att kraftigt påverka avbildningen vid små håldiametrar.

Med utgångspunkt från Fig. 2.1 skulle man vara böjd att tro att ju mindre hålet är desto skarpare blir bilden. Detta är också sant till en viss gräns, men om man går under denna gräns blir bilden faktiskt suddigare ju mindre hålet är (plus att bilden blir extremt ljussvag). Denna försämring vid mycket små hålstorlekar beror på ljusets diffraktion när det färdas genom små öppningar. Hur ska då hålet dimensioneras för att bilden ska bli så skarp som möjligt? Det skulle föra för långt att i detta kompendium göra en stringent genomgång av detta, så låt oss istället göra en "ingenjörsmässig uppskattning"* . Uppenbarligen kan man för stora hålöppningar bortse från diffraktion och anse att ljuset utbreder sig rätlinjigt. Å andra sidan kan man för mycket små öppningar anse att diffraktionen helt tar överhanden och bestämmer bildens utseende. Det vi nu söker är gränslandet där den geometriska utbredningen vi får genom att anta rätlinjig utbredning blir ungefär lika stor som utbredningen på grund av diffraktion, se Fig. 2.2.

* Den intresserade hänvisas till <https://www.kth.se/social/files/542d2d2df276546ca71dffa/Pinhole.pdf> där avbildningskvaliteten för olika håldiametrar undersöks med datorsimuleringar, och <https://www.kth.se/social/files/5c8675e856be5bdc4e5c4033/Pinhole.camera.pdf> där praktiska mätningar redovisas.

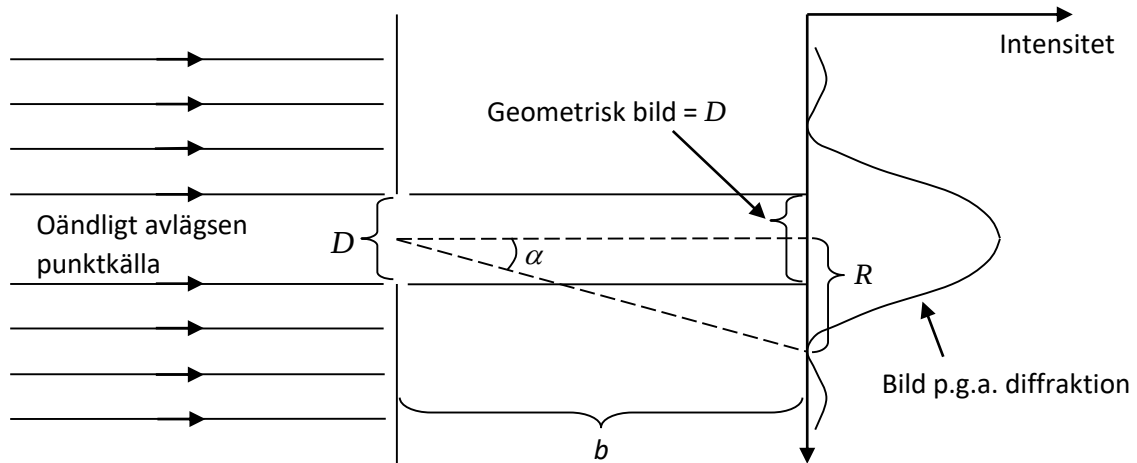


Fig. 2.2. Ljusets avböjning på grund av diffraktion.

Från fysiken vet vi att diffraktionsvinkeln

$$\alpha = \frac{1.22\lambda}{D} \approx \frac{R}{b}. \quad (2.1)$$

Lika stor oskarpefläck p.g.a. linjär utbredning som på grund av diffraktion ger $2R = D$, vilket med utnyttjande av ekvationen ovan ger

$$D = \sqrt{2.4\lambda b}. \quad (2.2)$$

Exempelvis ger $\lambda = 550 \text{ nm}$ och $b = 50 \text{ mm}$ att optimal hålstorlek $D = 0.3 \text{ mm}$. På grund av den lilla hålstorleken som krävs för bästa skärpa blir bilden mycket ljussvag. För att mera i detalj studera optimal hålstorlek, och den därvid erhållna skärpan, krävs att man gör ingående studier av diffraktionens inverkan. Detta finns gjort i referensen som anges på föregående sida. Man finner då, intressant nog, att bilden av en punktljuskälla kan bli märkbart mindre än hålstorleken D . Detta gäller om hålstorleken väljs ungefär enligt ekv. 2.2. Ett exempel på detta visas i Fig. 2.3 där belysningsfördelningen i bildplanet mätts upp för ett 0.60 mm hål placerat 160 mm från bildplanet.

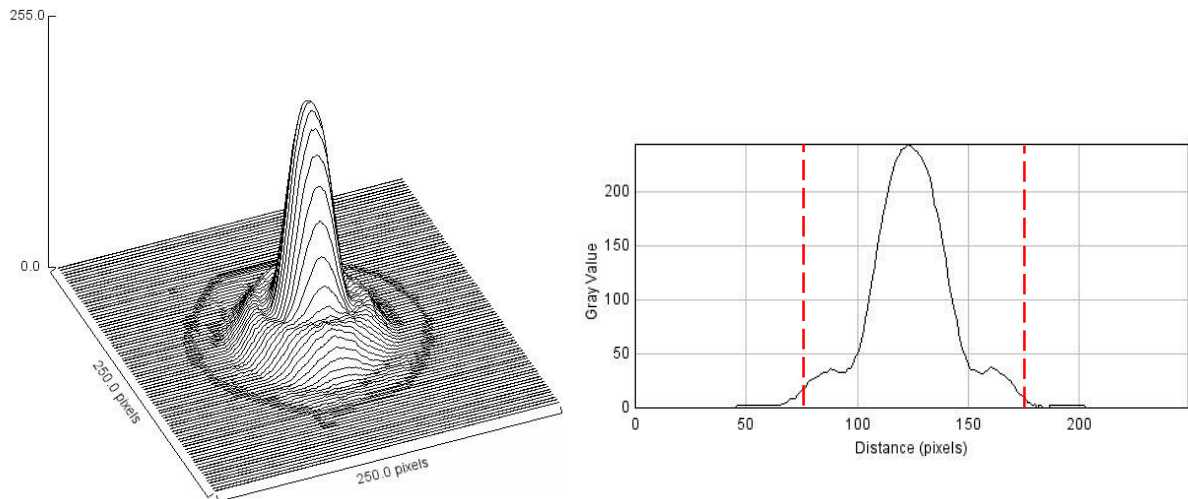


Fig. 2.3. Bilden av en punktljuskälla registrerad 160 mm bakom ett hål med diametern 0.60 mm. Vänstra figuren visar den tvådimensionella ljusfördelningen (vilken som förväntat är cirkulärsymmetrisk), medan den högra visar en profil genom centrum av ljusfördelningen. De röda vertikala linjerna anger hålets verkliga storlek. En pixel motsvarar $6.0\ \mu\text{m}$. Eftersom diffraktionen är våglängdsberoende gjordes registreringen med enbart rött ljus.

Man skulle kunna tro att hålkameran helt spelat ut sin roll eftersom man nu kan tillverka högkvalitativ och ljusstark optik. Det finns dock ett antal vetenskapliga tillämpningar för hålkameran inom våglängdsområden där man saknar lämplig avbildande optik, t.ex. för röntgen- och gammastrålning. Dessutom är det mycket roligt och tillfredsställande för vanliga fotoamatörer med barnasinnet i behåll att experimentera med hålkameror. Man kan själv snabbt och enkelt tillverka en egen kamera med vilken man kan ta hyggligt bra bilder som vi strax ska se. Nedan ges en sammanfattning av hålkamerans för- och nackdelar.

Fördelar:

- Billig
- Behöver ej fokuseras
- Fri från avbildningsfel (klarar t.ex. extremt stora bildvinklar utan distorsion)
- Kan användas för alla typer av strålning (IR, synlig, UV, röntgen, γ)

Nackdelar:

- Ljussvag bild medför lång exponeringstid
- Skärpan blir inte knivskarp (fördel om man vill ha lite ”softade” bilder)

Om man har tillgång till en systemkamera, dvs en kamera där objektivet kan avlägsnas, kan man enkelt bygga sin egen hålkamera. Ett exempel på detta visas i Fig. 2.4. Genom att göra ett hål i skyddslocket till kamerahuset (sådana brukar följa med när man köper en kamera), och där placera ett lagom stort hål, så har man fått en hålkamera. Se bildtexten för detaljer. I Fig. 2.5 visas ett exempel på bildkvaliteten man kan få från en sådan kamera.



Fig. 2.4. Ett kamerahus till en SONY $\alpha 7$ (fullformatsensor, 24 mm x 36 mm, med 24 Mpixlar) har byggts om till en hålkamera. Ett 0.2 mm hål gjordes i en Al-folie (vanlig kökstyp) genom att sticka en vass synål en liten bit in i folien. Denna folie tejpades sedan på en kartongbit som fästes med häftmassa (ljustätt!) på locket till kamerahuset (ett hål hade gjorts i locket). Det lilla hålet hamnade då ca. 24 mm från sensorn, vilket gör att bilderna uppvisar kraftig vidvinkleffekt (se kap. 6).



Fig. 2.5. Bild tagen med hålkameran i Fig. 2.4. Skärpedjupet blir extremt stort på grund av det lilla hålet. De gula blommorna ("Ryssgubbe") är ca. 8 mm i diameter, och befinner sig ca. 9 cm framför kameran. Bilderna blir också helt distorsionsfria. Distorsion är annars ett utpräglat problem för vidvinkelobjektiv.

Skärpan i hålkamerabilder kommer inte att bli lika bra som med vanliga kameraobjektiv. Bilderna får en lite ”soft”, men ofta konstnärlig, prägel. För att få skarpast möjliga bilder så ska man använda så stor sensor som möjligt, och ett kort avstånd mellan hål och sensor. Detta innebär att hålet måste vara litet (ekv. 2.2). För att få ett lagom stort hål kan man behöva göra flera försök med den metod som beskrivs i Fig. 2.4. Har man tillgång till en lupp (eller ännu hellre mikroskop) med mätskala, kan man använda denna för att kontrollera hålstorleken. Annars kan man ta några provbilder och kontrollera resultatet. Det är inte kritiskt att hålet är cirkulärt och kanterna jämna, det kan bli bra bilder ändå*. På grund av hålets ringa storlek blir exponeringstiderna betydligt längre än när man använder vanliga objektiv. Men med en fullformatsensor (24 mm x 36 mm) kan man vanligen använda ISO-tal på 1000 och ännu mer med gott resultat, och då kan man ofta ta bilder utan stativ och ändå undvika skakningsoskärpa. Bilden i Fig. 2.5 togs utan stativ med ISO 1000 och exponeringstid 1/30 sekund.

* Mera detaljer ang. att bygga sin egen hålkamera ges på
<https://www.kth.se/social/files/5c8675e856be5bdc4e5c4033/Pinhole.camera.pdf>

3. Kameratyper

Figur 3.1 visar principiella uppbyggnaden av en kamera. Även avancerade kameror fungerar efter samma grundprincip, men man har sedan lagt till finesser av olika slag. Objektivet, som i verkligheten alltid består av flera linselement, ger i sensorplanet en reell optisk bild av motivet.

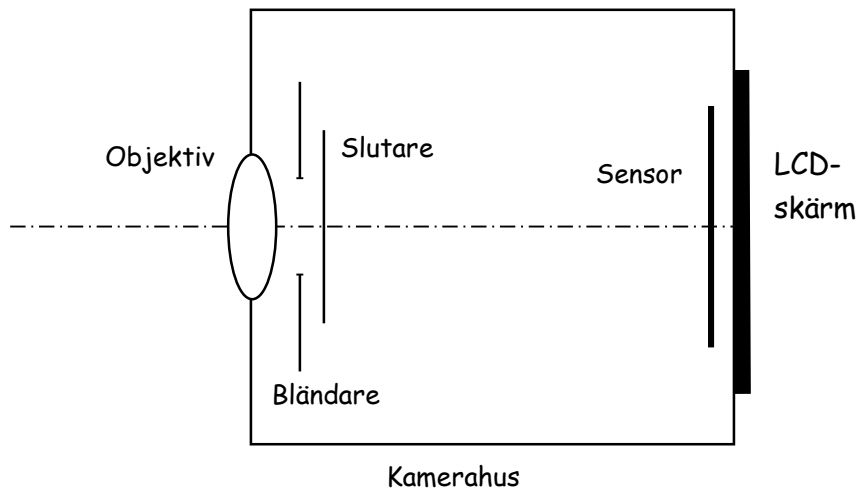


Fig. 3.1. En enkel kamera i genomskärning. Sensorn utgjordes tidigare av fotografisk film, men numera används vanligen elektroniska sensorer (CCD eller CMOS). LCD-skärmen ger en förhandstitt på hur bilden kommer att se ut. Dessutom visar den ofta kamerainställningar som exponeringstid, bländartal etc.

Kamerahuset är helt enkelt en ljustät låda (camera obscura) som endast släpper in ljus via objektivet. För att reglera exponeringen, dvs. hur många fotoner som totalt träffar sensorn, har man två reglage på kameran: bländare och slutare. Bländaren är en (mer eller mindre) cirkulär öppning, vars diameter kan varieras. På så sätt kan man variera belysningen i den optiska bilden i sensorplanet. Slutaren kan liknas vid en ljustät ridå som kan öppnas för att släppa in ljus under en viss tid, för att sedan skjutas för igen. Bländaren och slutaren kommer att gås igenom mer i detalj i senare kapitel.

Sensorn är en ljuskänslig integrerad krets som innehåller en matris av små detektorelement (pixlar), och kommer att beskrivas närmare i kapitel 14. LCD-skärmen på kamerans baksida fungerar som "sökare" (viewfinder på engelska), dvs den visar hur mycket av motivet som kommer med på bilden. Dessutom kan man överlagrat se exponeringsinformation med mera. Moderna kameror är vanligen utrustade med automatik, både vad gäller exponering och avståndsställning.

En kameratyp som har en mycket speciell och avancerad optisk sökare är spegelreflexkameran. Denna kameratyp är populär bland professionella fotografer och avancerade amatörer. Figur 3.2 visar principen för en spegelreflexkamera. Förutom stillbilder kan dessa kameror ofta spela in videosekvenser.

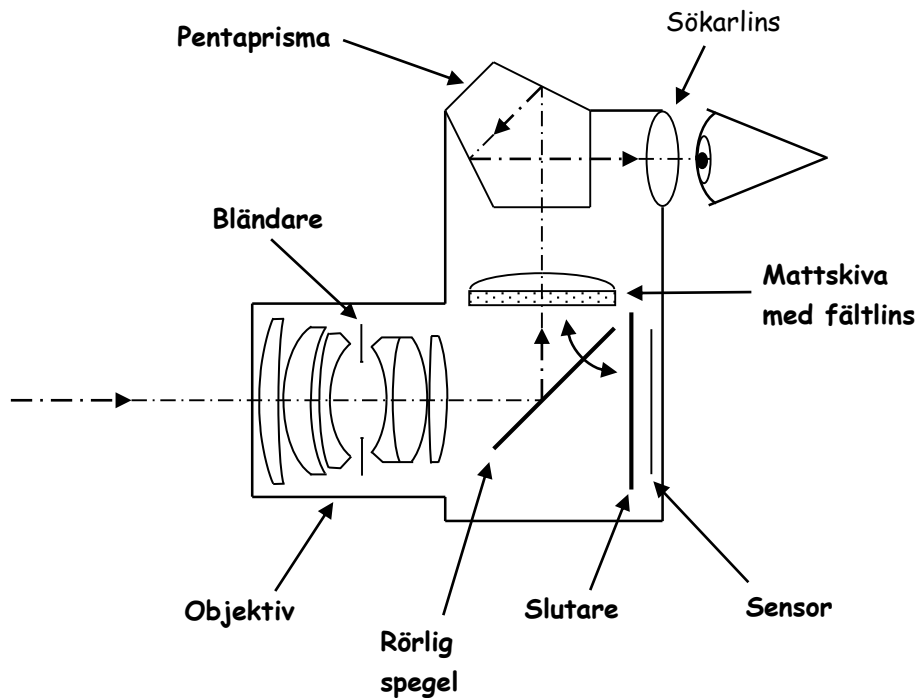


Fig. 3.2. Spegelreflexkamera i genomskärning.

I en spegelreflexkamera sitter en rörlig spegel en bit framför sensorplanet. Denna spegel befinner sig normalt i nedfällt läge, så att dess yta bildar 45° vinkel med objektivets optiska axel. Detta innebär att ljuset kommer in genom objektivet, speglas rakt upp mot taket i kamerahuset, och träffar en mattskiva som är monterad där. Mattskivan utgörs av en tunn plastskiva som diffust sprider ljus i alla riktningar. Fältlinsen gör att mer av det spridda ljuset riktas mot pentaprismat, vilket ger en ljusstarkare sökarbild. Mattskivan befinner sig på samma avstånd från objektivet som sensorn. Får vi en skarp optisk bild på mattskivan, så vet vi alltså att vi även kommer att få en skarp bild på sensorn vid exponeringen. När vi tittar in i sökaren ser vi (via en del optik) en bild av mattskivan. När avståndsställningen ändras kommer objektivet att flyttas närmare eller längre bort från sensorplanet. I sökaren kan vi då direkt se hur skärpan i den optiska bilden förändras, och på så sätt säkerställa optimal fokusering. Genom att den optiska bilden först faller på en mattskiva, vilken vi sedan betraktar med vårt öga, så kommer skärpeinställningen inte att bli beroende av ögats ackommodation (avståndsställning) eller eventuella synfel, t.ex. närsynhet. Den utomordentliga skärpekontrollen man på detta sätt får är speciellt värdefull vid närbildsfotografering, eftersom skärpeinställningen då måste vara mycket exakt.

När man trycker på avtryckaren för att exponera en bild med en spegelreflexkamera, så fälls spegeln upp mot mattskivan i kamerahusets tak. Ljuset fortsätter då istället fram mot sensorplanet som dock fortfarande är täckt av ridåslutaren (som beskrivs i kapitlet om slutare). En kort stund efter att spegeln fällts upp mot mattskivan, öppnas ridåslutaren och släpper fram ljus till sensorn under så lång tid som ställts in, varefter den återigen täcker för sensorn. Strax därefter fälls spegeln åter ner i 45-gradersläget. Eftersom det är nödvändigt att spegeln fälls upp vid exponeringen, så innebär det att bilden i sökaren försvinner precis i exponeringsögonblicket. Detta kan möjligen uppfattas som lite störande, men de flesta tycker nog att det är ett lågt pris att betala om man beaktar de fördelar som spegelreflexkameran erbjuder.

LCD-skärmen på baksidan av många kamerahus fyller på flera sätt samma funktion som den optiska sökaren i en spegelreflexkamera. Man ser t.ex. om bilden på sensorn är skarp eller inte. En fördel med LCD-skärmar är att de i mörka miljöer ger en ljusstarkare bild än en optisk sökare. En nackdel är däremot att de ofta har otillräcklig ljusstyrka vid utomhusfotografering i dagsljus pga. ströljus från omgivningen, plus att en viss eftersläpning i bildvisningen är vanlig. LCD-skärmar har också en begränsad dynamik vad gäller att samtidigt visa både mörka och ljusa motivpartier. Det finns emellertid kameror som förutom LCD-skärm på baksidan har en så kallad elektronisk sökare (electronic viewfinder på engelska). Dessa kameror påminner till funktionen om spegelreflexkameror, men saknar spegel mattskiva och pentaprisma. När man tittar genom sökarlinsen (se Fig. 3.2) ser man istället en liten elektronisk bildskärm som påminner om en LCD-skärm. Härigenom undviker man mycket av ströljusproblematiken med LCD-skärmar, men fortfarande kan den begränsade dynamiken vara en nackdel. Elektroniska sökare har emellertid utvecklats kraftigt på senare år, och uppfyller idag högt ställda krav.

Vissa kameror, så kallade systemkameror, har utbytbara objektiv, från extrema vidvinklar till långa teleobjektiv. Dessutom kan de utrustas med mellanringar och bälgar för närbildsfotografering. Spegelreflexkameror är alltid systemkameror. Systemkameror med elektronisk sökare brukar kallas spegellösa systemkameror.

Den kameratyp som till vardags används mest är antagligen mobilkameran, dvs den kamera som är inbyggd i mobiltelefoner. Förutom att ta stillbilder kan dessa kameror ofta spela in videosekvenser. En mobilkamera måste vara mycket liten och lätt, vilket ställer speciella krav på konstruktionen. Sensorn är mindre än i de flesta andra kameratyper, och linserna i objektivet är ofta asfäriska gjutna plastlinser. Eftersom mobiltelefoner är så tunna används ibland en periskop-konstruktion som gör att ljuset vinklas 90°, och sedan färdas i telefonens längdriktning där det finns mera plats. Fokuseringen sker automatiskt (se kap. 24) genom så kallad intern fokusering, dvs yttersta linsen ligger fast och man förskjuter de inre linselementen relativt varandra. Förskjutningarna kan göras t.ex. piezoelektriskt eller med MEMS teknik (vi går inte in på detta här, den intresserade kan googla). Ofta finns flera kameror i samma telefon så att man både kan fotografera andra och ta egna "selfies". Zoomobjektiv är inte så vanligt förekommande på grund av att dessa är mer utrymmeskrävande (men de förekommer på en del avancerade telefoner). Det är vanligare att bygga in flera kameror, t.ex. både vidvinkel- och telekameror, och sedan utnyttja "digital zoom", dvs zoomning i den digitala bilden. Detta ger dock sämre resultat än ett zoomobjektiv. Bildkvalitén för mobilkameror har genomgått en dramatisk förbättring med tiden, men fortfarande begränsas den av

restriktionerna vad gäller storleken. Att bara öka pixelantalet hjälper inte efter att man nått en viss gräns. Pixlarna blir så små att de registrerar för få fotoner, och optikens teoretiska upplösningsgräns spelar också in. För att kompensera för detta tillämpar ofta kamerafabrikanterna en mycket kraftig kosmetisk bildbehandling som gör att bilderna ser bra ut vid en första anblick, men om man zoomar in och tittar på detaljer ser bilderna ofta konstgjorda och nästan animerade ut. Så brukar det inte se ut i bilder tagna med andra, större kameratyper.

Som avslutning på detta kapitel ska vi titta lite på en gammaldags kameratyp som kallas ateljékamera, och som huvudsakligen används med film (men det finns även digitala varianter). Ateljékameror används inte så ofta idag, men erbjuder unika möjligheter att påverka bildresultatet. I en ateljékamera är objektiv och filmhållare (ofta för bladfilm) monterade på två plattor åtskilda av en ljustät bälg, se Fig. 3.3. Karakteristiskt för ateljékameran är att man har ett mycket stort antal frihetsgrader vad gäller att förskjuta och vrida objektiv- och filmplattorna i förhållande till varandra. Detta ger t.ex. möjlighet att undvika ”störtande linjer” vid arkitekturfotografering, samt att åstadkomma skärpeplan som ligger på ”snedden” (se kap. 6.7). För att kunna studera skärpeinställning, motivutsnitt med mera i en ateljékamera ersätts filmkassetten med en mattglasskiva vilken studeras med ögat, ofta under en svart duk för att skärma av ströljus.

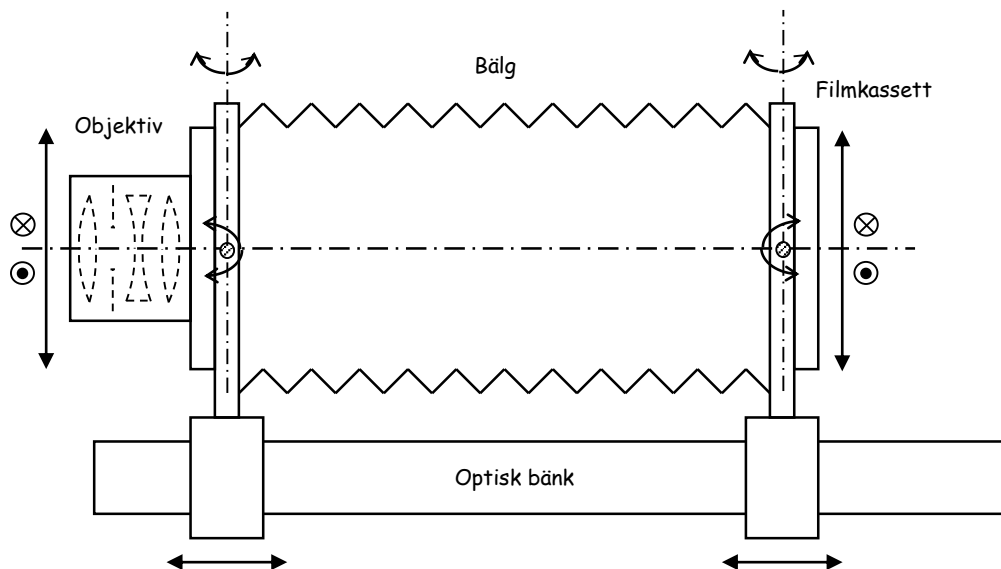


Fig. 3.3. Ateljékameran har ett mycket stort antal frihetsgrader vad gäller justering av objektiv- och filmplan. Både translation och rotation i snart sagt alla riktningar är möjliga. Objektiv och filmkassett kan lätt bytas så att olika filmformat, eller en elektronisk sensor, kan utnyttjas.

4. Bländaren och slutaren

4.1. Bländaren

Bländaren och slutaren används i en kamera för att få en korrekt exponering av sensorn. Bländaren sitter oftast monterad mellan linselementen i objektivet och består av ett antal tunna metall-lameller som kan skjutas in mer eller mindre långt i objektivöppningen. Därvid bildas en central öppning vars diameter kan varieras, Fig. 4.1.

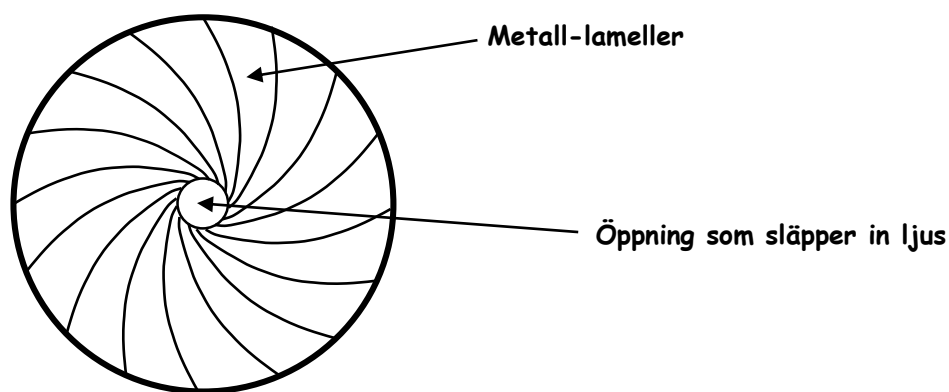


Fig. 4.1. Irisbländare.

När bländaröppningens diameter varieras, så varieras också det ljusflöde som passerar in genom objektivet. Detta påverkar i sin tur belysningsnivån i den optiska bilden som projiceras på sensorn. Hur hög belysningen blir beror på det så kallade **bländartalet**, F , som definieras genom formeln $F = \frac{f}{D}$, där f är objektivets

brännvidd och D bländaröppningens diameter*. På ett fotografiskt objektiv kan man typiskt variera bländartalet mellan ca. 2 och 22. Att ändra bländartalet med "ett steg" innebär att det ändras med faktorn $\sqrt{2}$, t.ex. från 5.6 till 8. Detta innebär att bländaröppningens area ändras med en faktor $\frac{1}{2}$, vilket gör att belysningen på sensorn också ändras med en faktor $\frac{1}{2}$. Observera att högre bländartal ger lägre belysning. Det lägsta bländartal som kan ställas in (motsvarande maximal bländaröppning) kallas objektivets ljusstyrka och behandlas närmare i kap. 9. De bländartal som kan ställas in på kameran omfattar alltid de så kallade hela stegen, t.ex. 4, 5.6, 8, 11, 16, samt oftast även halv- och tredjedelssteg. Det valda bländartalet anges på kamerans display, t.ex. som F5.6. Begreppet bländartal kommer att beröras ytterligare i kapitlet om fotometri.

* Egentligen motsvarar D den så kallade ingångspupillen, se Fig. 9.8.

4.2. Centralslutaren

Slutarens funktion är att se till att sensorn belyses under en viss bestämd tid, exponeringstiden. En typ av slutare som under många år använts för detta ändamål är centralslutaren, Fig. 4.2. Centralslutaren sitter ofta inbyggd mellan linselementen i objektivet, eller alldeles bakom objektivet. Denna typ av slutare har varit vanlig i enklare typer av kameror, typ kompaktkameror, som inte har utbytbara objektiv. Dessutom har den använts i ateljékameror. Under senare år har den tekniska utvecklingen av digitalkameror gjort att centralslutare har ersatts med så kallade ”elektroniska” slutare i kompaktkameror och kameramobiler, se kap. 4.4. Detta har gjort att centralslutaren har blivit ovanligare.

Precis som bländaren består centralslutaren av ett antal metall-lameller som kan dras isär och skjutas ihop. Innan exponeringen är dessa lameller helt ihopskjutna så att inget ljus kan passera in genom objektivet. När man trycker på avtryckaren dras lamellerna isär helt och hållet från objektivöppningen, så att ljus kan passera in och exponera sensorn. Efter en tid som bestäms av den inställda exponeringstiden så skjuts lamellerna åter ihop och exponeringen avslutas.

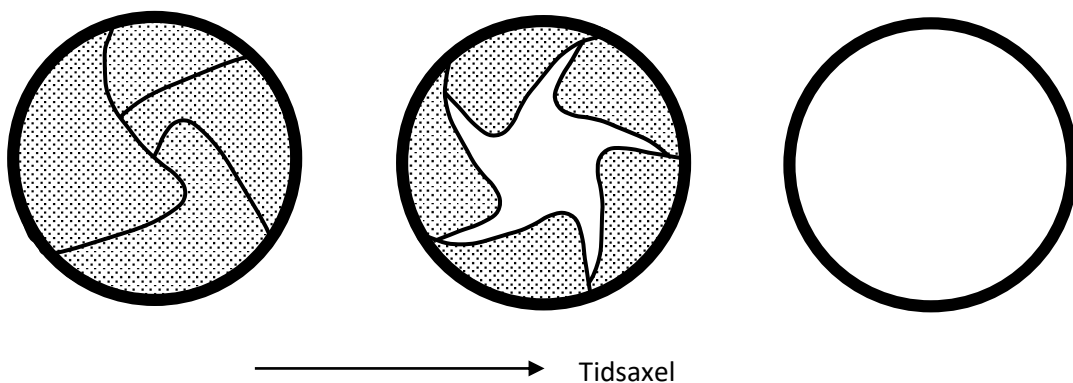


Fig. 4.2. Öppningsförloppet hos en centralslutare.

Typiska exponeringstider vid normal fotografi är $1/30 - 1/500$ sekund, vilket innebär att lamellerna mycket snabbt måste accelereras och bromsas in. Det är svårt att åstadkomma mycket korta exponeringstider med en centralslutare, och den kortaste tiden ligger ofta på $1/500$ s. Förutom denna begränsning vad gäller korta tider, så har centralslutaren nackdelen att den effektiva exponeringstiden kommer att bero på inställt bländartal. Detta beror på att den tid det tar att dra isär och skjuta ihop lamellerna inte är försumbar vid korta exponeringstider. Detta gör att belysningen successivt ökar under öppningsfasen, därefter är den konstant under en viss tid, samt minskar därefter successivt under stängningsfasen, Fig. 4.3.

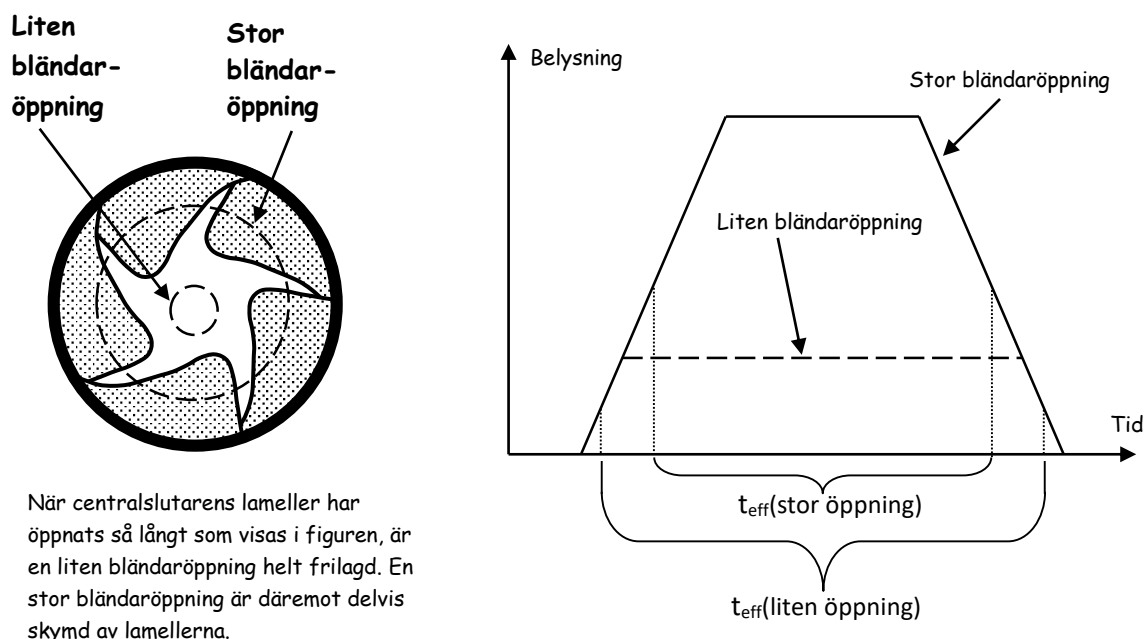


Fig. 4.3. Illustration av orsaken till att effektiva slutartiden varierar med bländaröppningen för en centralslutare.

Ju mindre bländaröppning, desto snabbare friläggs hela bländaröppningen, och därmed uppnås maximal belysning snabbare. Denna effekt motverkas av att öppningen som skapas av slutarlamellerna är "stjärnformig" under öppnings- och stängningsfaserna. Men detta kan inte helt kompensera för denna effekt. Eftersom sensorn har en viss marginal för felexponering kan detta "fel" hos centralslutaren oftast försummas.

4.3. Ridåslutaren

I kameror med utbytbara objektiv, så kallade systemkameror, används vanligen inte centralslutare utan ridåslutare, även det en gammal uppfinning. Ridåslutaren är fördelaktig av två skäl. Dels gör den att man inte behöver bygga in en slutare i varje objektiv som används (vilket sparar vikt och pengar), dels skyddar slutaren sensorn från damm när man tar bort objektivet från kamerahuset.

Ridåslutaren sitter monterad omedelbart framför sensorn och består av två ridåer, vanligen uppbyggda av metall-lameller. Före exponeringens början täcker ena ridån sensorn, medan den andra ligger ovanför, Fig. 4.4. Vid exponeringen rör sig den ridå som täckte sensorn vertikalt nedåt, och täcker av sensorn så att den kan exponeras för ljus. Den andra ridån, som tidigare låg ovanför sensorn, börjar efter en viss fördröjning att röra sig nedåt och täcker för sensorn igen efter att den inställda exponeringstiden nåtts*. Hastigheten varmed ridåerna rör sig ligger ofta

* I många kameror rör sig ridåerna nerifrån och uppåt, vilket naturligtvis går lika bra.

runt 5-10 m/s, och är oberoende av exponeringstiden. Korrekt exponeringstid erhålls genom val av den fördröjning med vilken ridå nummer två börjar röra sig. Vid korta slutartider (gränsen går ofta vid ca. $1/250$ s) börjar ridå nummer två att täcka för den del av sensorn där exponeringen startade, innan ridå nummer ett har hunnit fullborda sin rörelse (Fig. 4.4). Detta innebär att sensorn exponeras genom att en springa som släpper in ljus rör sig över sensorn under exponeringen. När man ändrar exponeringstiden påverkas inte springans hastighet, utan det är springans bredd som ändras - ju mindre bredd, desto kortare exponeringstid. Ridåslutare i digitala systemkameror kan komma ner i tider runt $1/8000$ sekund (vilket innebär en springbredd av ca. 1 mm). Efter exponeringen täcker ridå 2 sensorn, medan ridå 1 ligger nedanför. Innan nästa bild tas rör sig ridåerna så att ursprungsläget, dvs. att ridå 1 täcker sensorn, återställs. I praktiken kommer inte ridåernas hastighet att vara helt konstant under exponeringsförloppet, eftersom det tar viss tid accelerera upp till full hastighet. Sålunda är ridåernas hastighet lite lägre i början av förloppet, vilket kompenseras genom att springan i detta skede är lite smalare och sedan successivt görs bredare.

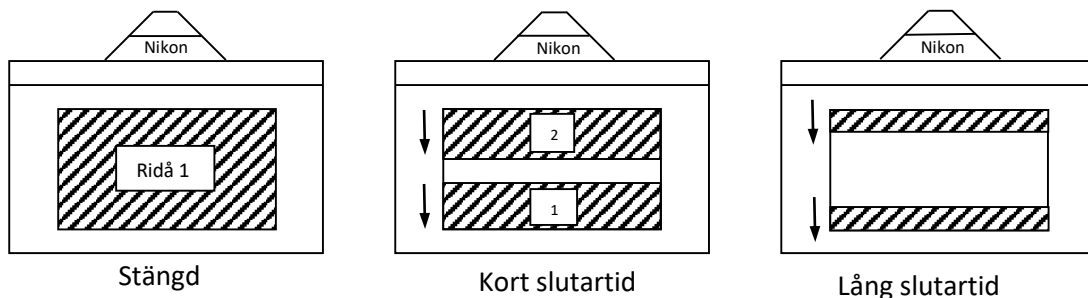


Fig. 4.4. Funktionen hos en ridåslutare.

Vid användning av ridåslutare exponeras inte hela sensorytan samtidigt, vilket ibland kan ge vissa problem (eller lustiga effekter) när man fotograferar snabbt rörliga föremål. Typexemplet är roterande flygplanspropellrar eller helikopterrotorer, vilka ofta ser krökta ut som bananer på bilden. En annan sak att tänka på är de problem som kan uppstå vid blyxtfotografering. Blyxttiden är oftast mycket kort ($\approx 1/1000$ s). Om man använder en kort exponeringstid kommer endast den smala del av sensorn där springan befinner sig att exponeras. Därför måste man vid blyxtfotografering använda så lång exponeringstid, ofta runt $1/250$ s, att ridå nummer ett hinner slutföra sin rörelse med att täcka av hela sensorn innan ridå nummer två börjar täcka för igen. Digitalkameror brukar ha en spärr som gör att man inte kan ställa in för kort tid när blixten är aktiverad. Den i kameran inbyggda blyxtsynkroniseringen ombesörjer att blixten tänds i rätt ögonblick, dvs. när hela sensorn är avtäckt och kan ta emot ljus. Det finns emellertid "intelligenta" blyxtar som gör att man kan använda blyxt även vid de allra kortaste slutartiderna för en ridåslutare. Istället för en enda blyxt kommer då ett helt tåg av blyxtar under hela den tid som den smala slutarspringan rör sig fram över sensorn.

4.4. Elektroniska slutare

I digitalkameror är det inte alltid nödvändigt att ha en mekanisk slutare. Den ljuskänsliga sensorn (se kap. 14) kan nämligen ofta styras rent elektriskt så att den påbörjar och avslutar en exponering. Man brukar då säga att kameran har en ”elektronisk” slutare. Fördelar med elektroniska slutare är att de är helt tysta, ger ökad driftsäkerhet samt minskad kostnad. Om mekanisk slutare behövs eller inte beror på vilken typ av sensor som sitter i kameran. För de vanligaste typerna av sensorer som används i konsumentprodukter (interline transfer CCD eller CMOS, se kap. 14), kan man i princip klara sig utan mekanisk slutare. Därför brukar lite enklare kameror, typ mobilkameror och kompaktkameror, sakna mekanisk slutare. Lite mer avancerade digitalkameror, typ systemkameror, brukar dock vara utrustade med en mekanisk ridåslutare. I kap. 14.4 beskrivs hur elektroniska slutare fungerar, och varför avancerade kameror fortfarande har en mekanisk slutare.

Det kan nämnas att en kombination av elektronisk och mekanisk slutare även kan utnyttjas för centralslutare, vilket gör det möjligt att blyxfotografera ner till så korta tider som 1/2000 s. Detta utnyttjas i Hasselblad X1D, lanserad 2016.

5. Optisk avbildning

5.1. Tunn lins

Objektivet är en av de viktigaste komponenterna i en kamera. En förståelse av dess uppbyggnad, funktion och prestanda är därför mycket väsentlig för att på ett bra sätt kunna utnyttja sin kamera. Dessa frågor kommer att beröras i ett flertal kapitel, varav detta kommer att behandla den grundläggande beskrivningen av optisk avbildning. Vi ska börja med det enklaste fallet, nämligen avbildning i en tunn lins, Fig. 5.1.

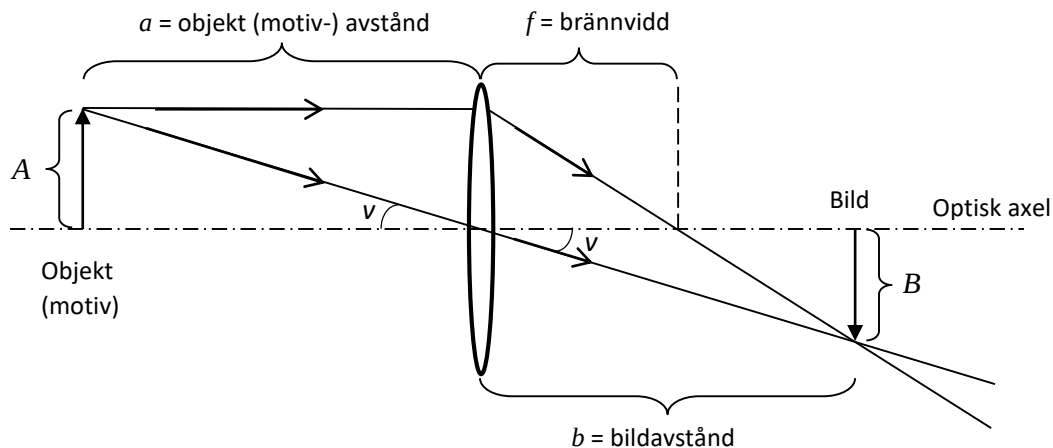


Fig. 5.1. Strålgång i tunn lins. Objektavståndet a i figuren är dock inte detsamma som avståndet som visas på objektivets avståndsställningsskala, se Fig. 10.4. Det sistnämnda avståndet anger totalt avstånd mellan motivet och sensorplanet.

Med tunn lins avses att linsens tjocklek är försumbar jämfört med brännvidden. Vid fotografering är det av intresse att veta var den (reella) optiska bilden kommer att hamna längs optiska axeln. Det är ju där som sensorn ska befinna sig. Intressant är också att veta avbildningsskalan, dvs. hur stor bilden är jämfört med motivet. För detta ändamål gör man en strålgångskonstruktion enligt figuren ovan. För att få fram bildens läge och storlek behöver man finna skärningspunkten i bildrymden av två strålar som utgår från en punkt i motivet. Ofta använder man då en stråle som passerar genom linsens centrum, samt en som infaller parallellt med optiska axeln mot linsen. Den förra kommer att gå igenom linsen utan att brytas eller sidoförskjutas, alltså precis som om linsen inte fanns (detta förutsätter att linsen är tunn). Den andra strålen bryts av linsen så att den korsar optiska axeln på brännvidds avstånd (det är så brännvidd definieras). Vid skärningspunkten mellan dessa båda strålar i bildrymden hamnar den optiska bilden. Med kännedom om linsens brännvidd kan man skalenligt rita upp strålgången på ett papper och helt enkelt mäta upp bildens läge samt avbildningsskalan. Vanligare är dock att man gör en frihandsteckning som illustration, och sedan använder linsformeln för beräkningar. Linsformeln ger oss sambandet mellan motivavstånd, bildavstånd och brännvidd:

$$\frac{1}{a} + \frac{1}{b} = \frac{1}{f}. \quad (5.1)$$

Känner man två av dessa storheter kan man räkna ut den tredje. Bildavståndet kan alltså lätt räknas ut om man känner motivavstånd och brännvidd. För oändligt motivavstånd blir $b = f$, och för ändliga motivavstånd gäller $b > f$. Vid fokuseringen (manuell eller automatisk) i en kamera, så förflyttas linsen relativt sensorn så att bildavståndet, b , blir korrekt (uppfyller ekv. 5.1) för det aktuella motivavståndet. Avbildningsskalan, M (efter engelskans magnification), ges av

$$M = \frac{B}{A} = \{\text{p.g.a. likformiga trianglar}\} = \frac{b}{a} \quad (5.2)$$

Ofta vill man i fotografin undvika att ha formler innehållande b , utan man vill hellre ha a (som lätt kan mätas eller uppskattas) och f (som står på objektivet). För detta syfte kan man göra omskrivningen

$$M = \frac{b}{a} = \left\{ b = \frac{af}{a-f} \text{ fås ur linsformeln} \right\} = \frac{f}{a-f}. \quad (5.3)$$

Ofta gäller vid fotografering att $a \gg f$ vilket medför att $M \approx \frac{f}{a}$, vilket är en mycket användbar approximation. Om det är bildstorleken B , snarare än avbildningsskalan M , som är av intresse kan den för stora motivavstånd beräknas ur $B = f \tan v$, där v är synvinkeln under vilken motivet ses, Fig. 5.1 (vid kortare motivavstånd måste f bytas mot b). Detta var ju enkelt, eller hur? Tyvärr är det i praktiken så att man mycket sällan kan approximera fotografiska objektiv med tunna linser.

5.2. Objektiv

Ett typiskt objektiv kan ha en uppbyggnad enligt Fig. 5.2. Den komplicerade uppbyggnaden är till för att minska avbildningsfelen, vilka kommer att beskrivas i ett senare kapitel.

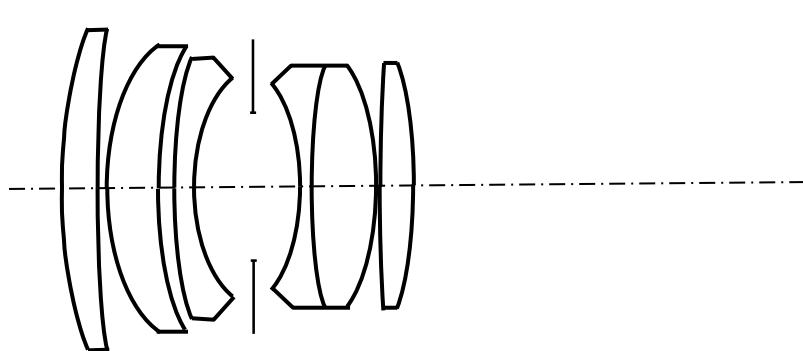


Fig. 5.2. Exempel på uppbyggnad av ett fotografiskt objektiv (55 mm, $f/1.8$).

Avståndet längs optiska axeln mellan främsta och bakersta linsytan är i många fall större än brännvidden. Approximationen tunn lins faller naturligtvis platt i ett sådant fall. Man kan naturligtvis tänka sig att betrakta varje linselement i objektivet som en tunn lins och räkna på varje linselement för sig. Detta blir dock ganska komplicerat när man har många linselement, och dessutom kan de individuella linselementen ibland vara ganska tjocka. I praktiken gör man därför vanligen så att man inför begreppet huvudplan, vilket kraftigt underlättar beräkningar på optisk avbildning i objektiv. Vi betraktar ett linssystem med ett godtyckligt antal linser (som inte behöver vara tunna) enligt Fig. 5.3, där bara första och sista linsytorna är utritade.

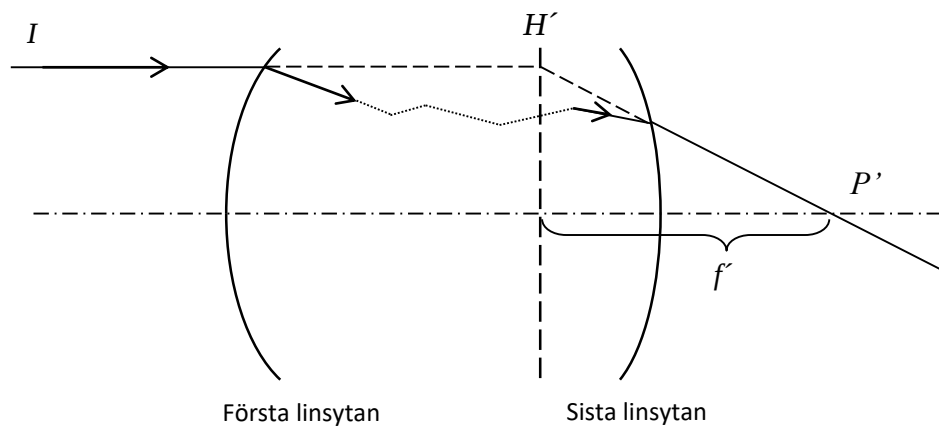


Fig. 5.3. Definitionen av huvudplan och brännvidd.

Vi låter en ljusstråle, I , infalla från vänster parallellt med optiska axeln. Denna ljusstråle kan brytas ett flertal gånger i olika linselement (strålens väg är markerad med prickade linjer) och bryts slutligen en sista gång i den sista linsytan, varefter den korsar optiska axeln i punkten P' . Vi drar nu en tänkt förlängningsstråle till infallande strålen såsom den skulle ha fortsatt om inte objektivet funnits (streckad linje). Likaså drar vi en rätlinjig förlängningsstråle bakåt från den utgående strålen. Genom skärningspunkten mellan dessa två förlängningsstrålar, och vinkelrätt mot optiska axeln, ritar vi en rät linje betecknad H' i figuren. Denna linje markerar huvudplanets läge i objektivet. Finessen med att införa huvudplanet är att, som synes i figuren, effekten av alla linsers brytning av infallande strålen är precis densamma som om vi hade haft en enda tunn lins placerad vid huvudplanet. Brännvidden f' för den tunna linsen motsvarar avståndet mellan huvudplanet och punkten P' , vilket definitionsmsätt är objektivet brännvidd. Fantastiskt, då kan vi alltså beräkningsmässigt ersätta ett objektiv med en tunn lins, eller? Nja, riktigt så enkelt är det inte, vilket man inser om man istället skickar in en ljusstråle, II , som går parallellt med optiska axeln från höger mot vänster, Fig. 5.4.

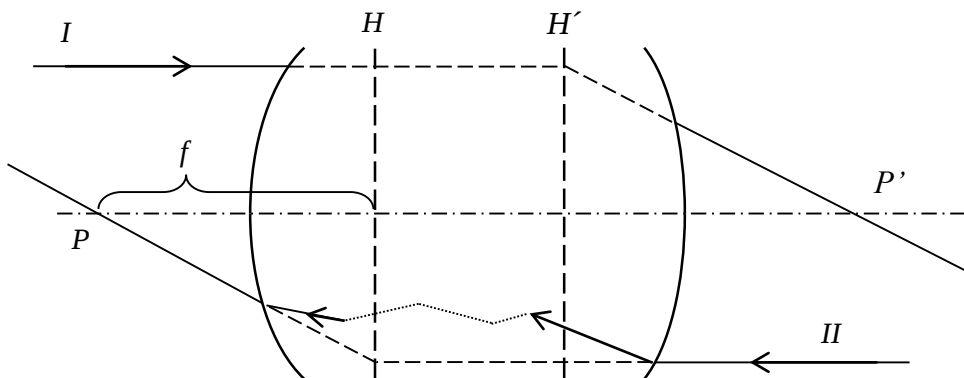


Fig. 5.4. Exempel på hur de två huvudplanen kan vara belägna i ett objektiv.

På samma sätt som för stråle I kan vi nu ta fram ett huvudplan H och brännvidd f för denna stråle. Om medierna till vänster och höger om objektivet har samma brytningsindex, så är $f' = f$. Vi ska i fortsättningen anta att så är fallet, eftersom det vanligen är luft på bägge ställena. Tyvärr medför detta inte att huvudplanen sammanfaller, utan dom kan i praktiken ligga klart åtskilda. Hur ska vi nu rita upp strålgången i ett objektiv med användande av huvudplan och brännvidd? Jo, det görs enligt figuren nedan.

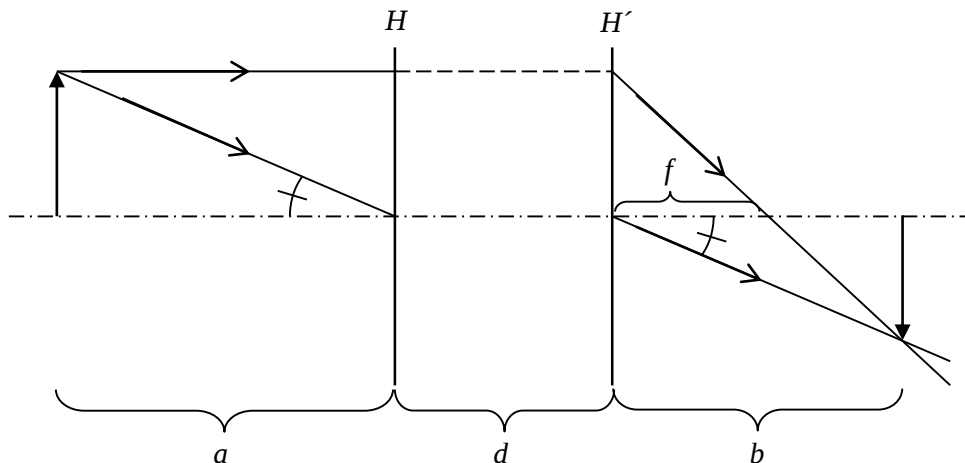


Fig. 5.5. Strålgångskonstruktion i ett objektiv med utnyttjande av huvudplan.

Vi ser att strålgången mycket liknar strålgången i en tunn lins. Enda skillnaden är att vi nu "skurit itu" den tunna linsen i två delar som vi särat på avståndet d , dvs. avståndet mellan huvudplanen. Mellan dessa båda halvor färdas ljusstrålarna alltid horisontellt. Om vi definierar objektavståndet a som avståndet mellan motiv och huvudplan H , samt bildavståndet som sträckan mellan H' och bilden, så kommer den enkla linsformeln att gälla även för ett objektiv. På samma sätt gäller att

avbildningsskalan $M = \frac{b}{a}$. Känner man huvudplanens lägen och brännvidden så är en strålgångsberäkning för ett objektiv alltså lika enkel som för en tunn lins. I praktiken kan man ofta räkna på objektiv som om de vore tunna linser, dvs enligt Fig. 5.1, utan att införa några fel. Enda fallet då man måste ta hänsyn till objektivets tjocklek, och därmed införa huvudplan, är då det totala avståndet mellan motiv och sensor (inbegripande d i Fig. 5.5) är av intresse.

Huvudplanens lägen kan variera en hel del beroende på objektivtyp. Ofta ligger de någonstans mellan första och sista linsytan, men de kan mycket väl ligga ute i luften både till vänster och höger om objektivet. De bägge huvudplanens lägen kan också vara omkastade så att H' ligger till vänster om H . Detta ger en strålgång enligt Fig. 5.6.

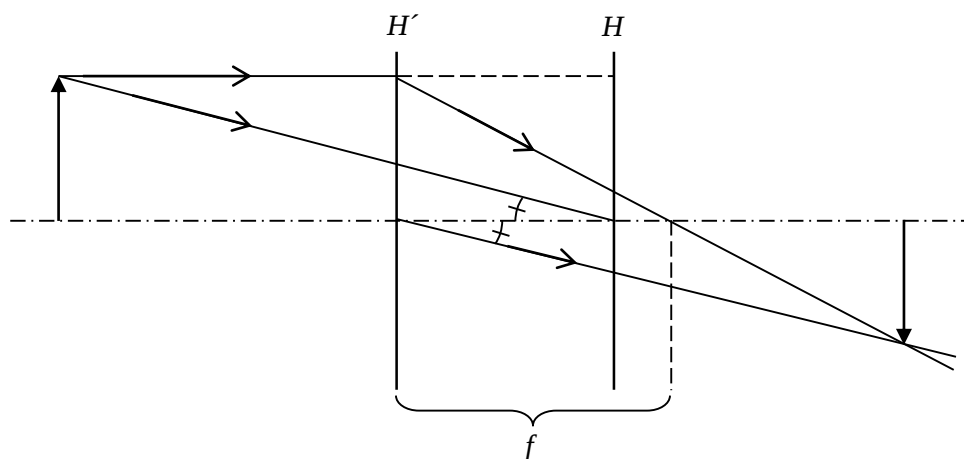


Fig. 5.6. Strålgångskonstruktion vid omkastade huvudplan.

Vi ser att inkommande strålen som är parallell med optiska axeln alltid bryts i H' och skär optiska axeln på avståndet f till höger om H' . Strålen genom linsens centrum riktas alltid mot skärningspunkten mellan H och optiska axeln, varefter den förskjuts horisontellt, och fortsätter i samma riktning från skärningspunkten mellan H' och optiska axeln.

5.3. Bildvinkel

Inom fotografi använder man ibland begreppet bildvinkel. Man talar t.ex. om vidvinkelobjektiv (se kap. 6). I Fig. 5.7. illustreras vad som menas med bildvinkel. Bildvinkeln beror på brännvidden* och sensors diagonalmått. Exempelvis så får vi för en så kallad fullformatsensor (24 mm x 36 mm storlek) diagonalmåttet $D = 43$ mm. Med brännvidden $f = 50$ mm får vi då en bildvinkel av 47° . Bildvinkeln

*Man brukar titta på fallet att motivavståndet är stort, vilket innebär att bildavståndet är ungefär lika med brännvidden.

kan variera från över 90° för extrema vidvinkelobjektiv ner till $2-3^\circ$ för extrema teleobjektiv. Mer om detta i nästa kapitel.

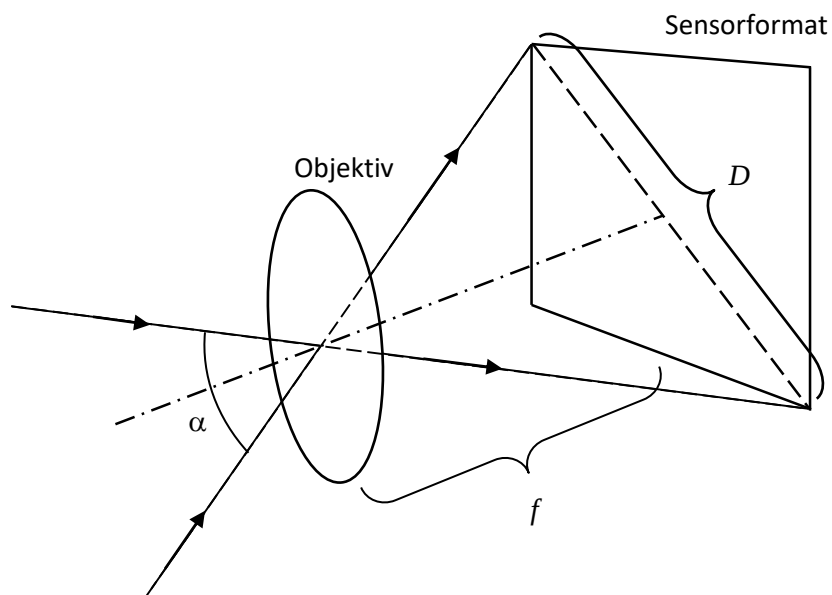


Fig. 5.7. Bildvinkeln, α , för ett objektiv anger hur stort vinkelintervall i motivrymden som ryms inom sensorn diagonalmått.

6. Objektivtyper

6.1. Inledning

Enklare kameror har fast optik med en enda brännvidd, eller ett zoomobjektiv med ganska begränsat omfång. Detta är oftast tillfyllt för enklare fotografering, typ minnesbilder från resan eller fotografering av släkträffar. För mer avancerad fotografering behöver man dock ofta variera brännvidden en hel del. Varför? Jo, helt enkelt för att få en avbildningsskala som passar motivet och fotograferingsavståndet. Vill man t.ex. fotografera en fågel på stort avstånd (den blir kanske rädd och flyger iväg om man går närmare) så behövs lång brännvidd för att inte avbildningsskalan ska bli så liten att fågeln bara ser ut som en prick på bilden. Den motsatta situationen uppstår om man vill fotografera ett stort föremål på kort avstånd. Då behövs en liten avbildningsskala för att hela motivet ska komma med på bilden. För att ge stor flexibilitet vad gäller objektivval så har mer avancerade kameror utbytbara objektiv. Sådana kameror kallas systemkameror. Beroende på objektivets brännvidd i förhållande till diagonalmåttet på sensorn så delar man in objektiverna i typerna vidvinkel, normal och tele.

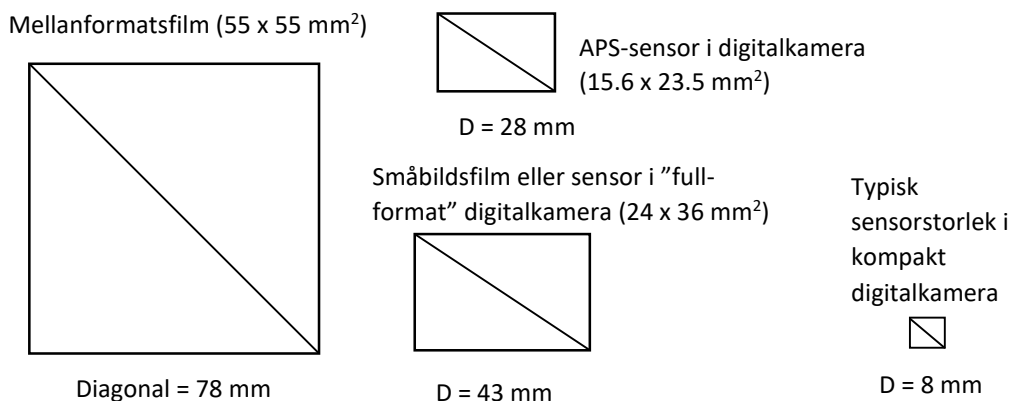


Fig. 6.1. Exempel på diagonalmått för olika film- och sensorformat (ej skalenlig figur).

6.2. Normalobjektiv

Man brukar anse att normaloptik ska ha en brännvidd som är ungefär lika med diagonalmåttet på sensorn. Detta ger en total bildvinkel av 53°. Detta betyder för en fullformatssensor ca. 43 mm brännvidd. Av tradition brukar man dock anse att 50-55 mm är "normal" brännvidd för fullformatskameror. För digitalkameror gäller att sensorstorleken oftast är mindre än fullformat. En typisk kompaktkamera har ofta en sensor med ett diagonalmått av endast 8 mm, varför normalbrännvidd blir ca. 8 mm. Många systemkameror har sensor vars area är ungefär hälften av en fullformatssensor. För dessa blir normalbrännvidden ca. 30 mm. Mer avancerade

amatörkameror samt rena proffskameror har vanligen fullformatsensor, eller i vissa fall ännu större. Om man å andra sidan betraktar en gammaldags ateljékamera för 9 cm x 12 cm bladfilm, så hamnar normalbrännvidden runt 150 mm.

6.3. Vidvinkel- och teleobjektiv

Vidvinkeloptik brukar man anse ska ha en brännvidd $< 0.9 \times$ diagonalmåttet på sensorn, och tele ska ha en brännvidd som är $> 1.4 \times$ diagonalmåttet. Detta svarar mot bildvinklar $> 60^\circ$ respektive $< 40^\circ$. Det bör betonas att det inte finns några strikta definitioner på detta, så gränserna kan vara lite flytande i litteraturen.

Ett teleobjektiv består i sin enklaste form av två linselement, ett positivt och ett negativt, Fig. 6.2. Genom detta byggsätt kommer huvudplanet, H' , att ligga ute i luften framför objektivet, som därigenom blir kortare och lättare att hantera. I verkligheten används ofta 4 – 5 linselement i teleobjektiv. Detta krävs för att få ner avbildningsfelen, men själva grundprincipen är den som visas i figuren.

Ett exempel på motsatta byggsättet, dvs. att huvudplanet ligger bakom objektivet, Fig. 6.3b, finner man ofta i vidvinkelobjektiv avsedda för spegelreflexkameror. Objektiv av denna typ kallas retrofokusobjektiv. Anledningen till detta byggsätt är att man måste ha tillräckligt stort utrymme mellan bakre linselementet och den rörliga spegeln som finns i spegelreflexkameror.

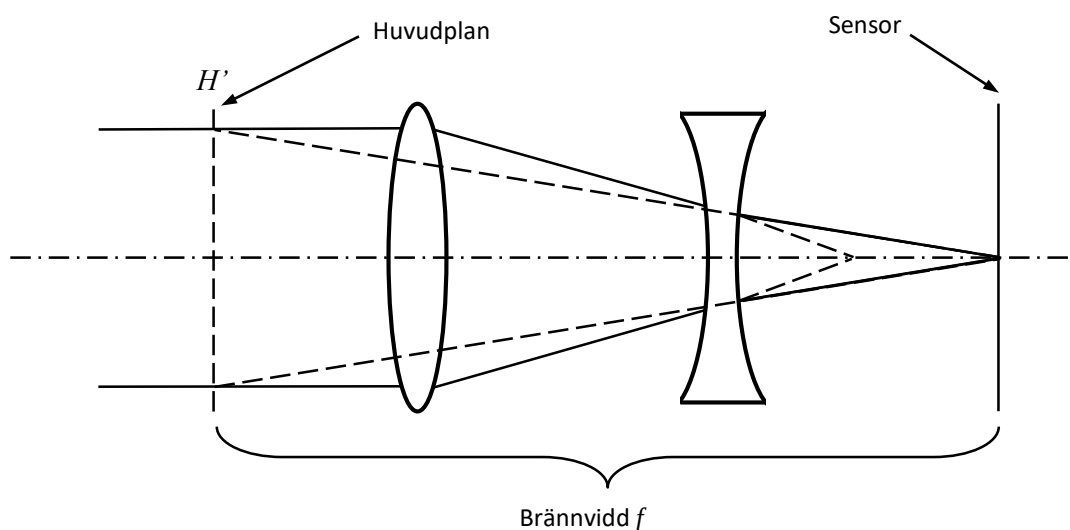


Fig. 6.2. Principiella uppbyggnaden av ett teleobjektiv.

6.4. Makroobjektiv

Förutom vidvinkel, normal och tele förekommer så kallade makroobjektiv. Dessa är speciellt avsedda för närbildsfoto och kan i princip ha vilken brännvidd som helst. Vanligast är dock att de har ungefär normalbrännvidd eller något längre (50 – 100 mm för fullformatskameror). Det som skiljer ett makroobjektiv från andra objektiv är två saker. Den första skillnaden gäller den rent mekaniska avståndsställningen, och den andra gäller optimeringen vad gäller avbildningsfelen. Avståndsställningen på ett vanligt normalobjektiv sträcker sig vanligen mellan ca. 0.5 meter och ∞ . Det som sker rent mekaniskt när man ändrar avståndsställningen är att hela objektivet flyttas närmare eller längre bort från sensorplanet (i vissa äldre objektivtyper flyttas istället linselementen relativt varandra så att brännvidden ändras). Vid inställning på ∞ motivavstånd så är avståndet mellan objektivets bakre huvudplan och sensorplanet lika med brännvidden. Vid kortare motivavstånd flyttas objektivet längre bort från sensorn. Det finns dock en gräns rent mekaniskt för hur lång denna förflyttning kan bli, och det är detta som sätter en gräns för hur nära motivet man kan gå (och därmed hur stor avbildningsskala man kan få). I ett makroobjektiv så har mekaniken ett större rörelseomfång, vilket gör att man kan fotografera på kortare avstånd. Maximal avbildningsskala är ofta 0.5 – 1 för ett makroobjektiv.

Det räcker dock inte med att införa ett större fokuseringsomfång (t.ex. med en mellanring) för att omvandla ett vanligt normalobjektiv till ett bra makroobjektiv. Anledningen är att vanliga objektiv är optiskt korrigerade för att ge bästa bildskärpa på ”stora” avstånd (vanligen ∞). Om man optimerar ett objektiv på detta sätt visar det sig tyvärr att avbildningsfelen blir stora när man fotograferar på korta avstånd. Därför har man optimerat lite annorlunda för makroobjektiven, nämligen att de ska ge bästa skärpa vid närbildsfoto (vanligen avbildningsskala 0.1). Ett sådant objektiv brukar ge tillfredsställande skärpa i hela området från ∞ avstånd till avbildningsskala 1 (dvs. då motivavståndet är dubbla brännvidden). Detta innebär emellertid stora krav på objektivkonstruktionen, och en kompromiss man ofta gör för att klara detta är att minska ner på linsdiametrarna, dvs. ljusstyrkan blir lägre (se definition på ljusstyrka i kap. 9). En annan lösning är att använda så kallade ”floating elements”. Detta innebär att linselementen förflyttas relativt varandra inne i objektivet när man ändrar avståndsställningen. På så sätt kan objektivets konstruktion förändras beroende på om man fotograferar på stora eller korta avstånd. Makroobjektiv med floating elements kan tillverkas med högre ljusstyrkor, men priset blir naturligtvis också högre.

Vid extrem närbildsfotografering med avbildningsskala > 1 så brukar man få bättre resultat om objektivet monteras bakfram, dvs. den ända som vanligen är vänd mot sensorn ska istället peka mot motivet. Detta kräver en särskild adapter.

6.5. Zoomobjektiv

Objektiv med variabel brännvidd kallas zoomobjektiv. Dessa har blivit mycket populära under senare år, och är nästan standard på digitalkameror. Zoomomfånget, dvs. förhållandet mellan längsta och kortaste brännvidd, är ofta i området 3 – 5 (men vissa objektiv har betydligt större omfång). Detta gör att man kan gå från vidvinkel

till måttligt tele utan att byta objektiv. Ofta har zoomobjektiven också ett makroläge som tillåter närbildsfotografering. Med tanke på denna mångsidighet kan man fråga sig varför man överhuvudtaget använder objektiv med fast brännvidd. Det finns flera skäl till detta. För det första kan man ibland behöva använda extrema vidvinkel- eller teleobjektiv. Inget zoomobjektiv klarar ett så stort zoomomfång med god bildkvalitet. För det andra så ger zoomobjektiv nästan alltid en generell lägre bildkvalitet än objektiv med fast brännvidd. Zoomobjektiven har ofta både mer ströljus, vilket ger sämre bildkontrast, och sämre skärpa. Slutligen är zoomobjektiven större och tyngre än objektiv med fast brännvidd, samt har oftast lägre ljusstyrka. Zoomobjektiv innehåller oftast fler linselement än objektiv med fast brännvidd (ofta 10 linselement eller fler). Detta är nödvändigt dels för att man utnyttjar förflyttningar av linselement för att ändra brännvidden, dels för att det behövs många linselement för att reducera avbildningsfelen. Det sistnämnda försvåras av att felen måste vara låga för alla brännviddsvärden som objektivet kan ge. De många linselementen innebär att antireflexbehandlingen måste hålla god kvalitet för att inte ströljuset ska bli alltför besvärande (men som nämnts är det ändå oftast högre än i andra objektiv). En enkel illustration av hur förskjutningar av linselement kan ge en variabel brännvidd visas i Fig. 6.3 som återger ett förenklat zoomobjektiv med bara tre linser.

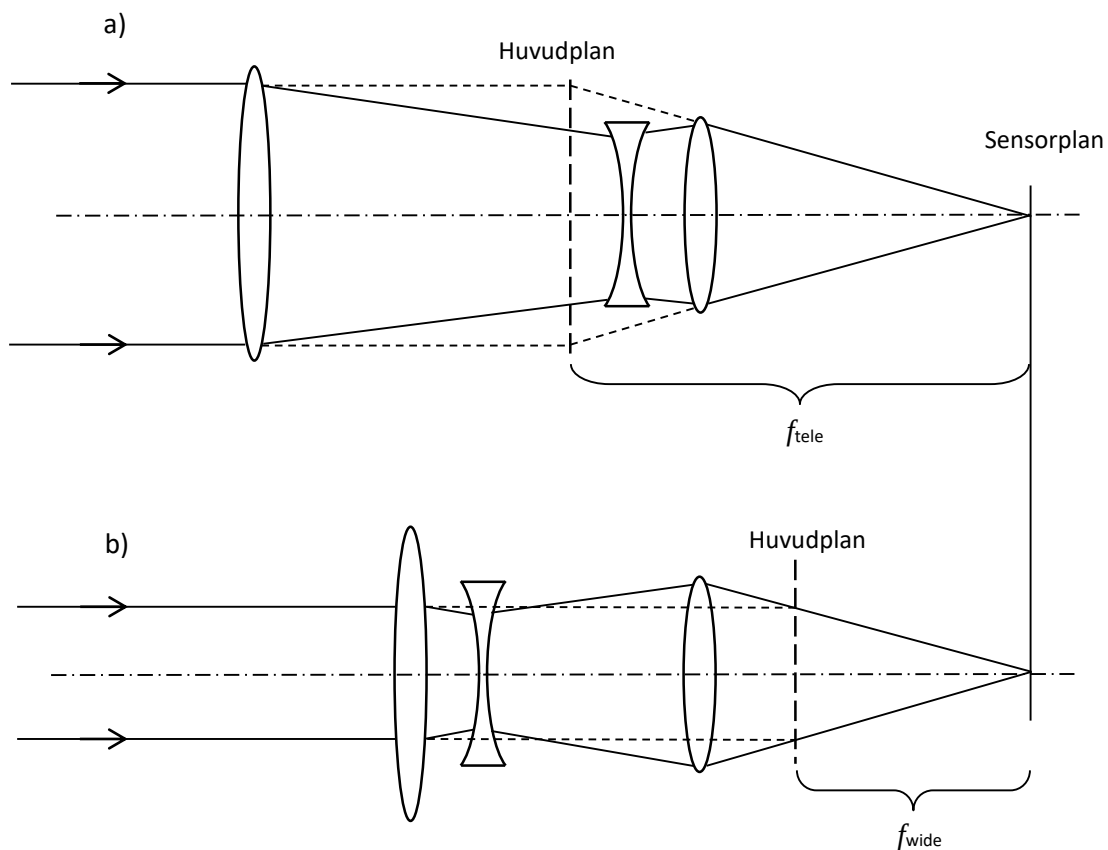


Fig. 6.3. Schematisk framställning av principen för ett zoomobjektiv. a) Teleläge, b) Vidvinkelläge.

I övre delen av Fig. 6.3 visas teleläget, och i den undre vidvinkelläget. Det senare erhålls genom att förskjuta de två vänstra linselementen mot varandra, medan det tredje linselementet ligger fixt. Zoomobjektiv brukar vara konstruerade så att bländartalet är konstant när brännvidden förändras. Detta uppnås genom att vid zoomning så ändras ingångspupillens diameter utan att den mekaniska bländaröppningen behöver förändras*. Principen för hur det fungerar syns i Fig. 6.3. Vid den kortare brännvidden, f_{wide} , så utnyttjas bara en mindre del av frontlinsens yta till att samla in ljus. Effektiva diametern minskar alltså samtidigt som brännvidden minskar, och kvoten mellan de två kan därför hållas konstant. Även avståndsställningen bevaras vid zoomning, vilket gör objektiven enkla att använda.

6.6. Ekvivalent brännvidd

En praktisk svårighet vad gäller brännviddsangivelser på digitalkameror, är att det finns många olika storlekar på sensorer. Dessutom är sensorstorleken ofta okänd för användaren. För att förstå vad t.ex. zoomobjektivet på 14.3 – 71.5 mm omfång innebär för kameran SONY DSC-R1, så måste man först ta reda på att sensorn har storleken 14.4 mm x 21.5 mm. Sedan får man jämföra de olika brännvidderna med diagonalmåttet för sensorn. Detta är krångligt, och eftersom de flesta (åtminstone lite äldre) fotografer är vana vid kameror med småbildsfilm (som i storlek motsvarar en fullformatsensor, dvs. 24 mm x 36 mm), så anger många digitalkameror sin brännvidd som "(1)35 mm equivalent" eller liknande (135, 35 och FX är beteckningar som betyder för fullformat). Den ovan nämnda SONY-kameran anger sålunda sitt brännviddsomfång till 24 – 120 mm "equivalent", vilket betyder att man vid kortaste brännvidd (14.3 mm) får med lika mycket av motivet som med en fullformatskamera utrustad med ett 24 mm objektiv (dvs. kraftig vidvinkel). Motsvarande jämförelse kan göras vid andra brännvidder. För en fotograf som är van vid fullformatskameror, så innebär dessa ekvivalenta brännviddsangivelser en stor hjälp eftersom man vet precis vilka brännvidder som svarar mot vidvinkel, normal och tele.

Som nämndes i inledningen på detta kapitel så utnyttjar man olika brännvidder för att kunna variera avbildningsskalan. Det visar sig emellertid att olika brännvidder inte bara ger olika avbildningsskala, utan också olika djupintryck i bilden. Detta kallas för bildens perspektiv och avhandlas i nästa kapitel.

6.7. "Tilt/shift" objektiv

Förutom de redan nämnda objektivtyperna finns ett flertal specialobjektiv för olika ändamål. Här ska en sådan typ beskrivas som på engelska går under beteckningen "tilt/shift lens". Innan vi går in på beskrivningen av detta objektiv, bör emellertid påpekas att alla möjligheter som det medger också kan uppnås med en ateljékamera (se kap. 3). En ateljékamera ger dessutom ännu större fotografiska möjligheter än ett "tilt/shift"-objektiv. Ateljékameror är dock stora, klumpiga och allmänt opraktiska att jobba med, i synnerhet ute i fält. Därför är det en stor fördel att ha ett

* Se Fig. 9.8 ang. begreppet ingångspupill.

”tilt/shift”-objektiv som ger de viktigaste av ateljékamerans fördelar, och som kan monteras på en systemkamera.

Ett ”tilt/shift”-objektiv är mycket användbart för arkitekturfotografering, samt när man vill att kamerans skärpeplan ska ligga på snedden. Låt oss först titta lite på arkitekturfotografering och problem som detta medför.

En ofta förekommande situation är att man står nere på marken och vill fotografera en hög husfasad. För att få med hela fasaden tvingas man ofta luta kameran snett uppåt. Exempel på en sådan bild ses i Fig. 6.4a. Man stöter emellertid på två problem när man lutar kameran uppåt. Ett problem är att man får så kallade störtande linjer i bilden, dvs byggnaden ser ut att smalna av uppåt och räta vinklar blir inte längre räta. Ett annat problem är att byggnadens nedre och övre delar befinner sig på olika avstånd från kameran. De kan därför inte samtidigt avbildas med optimal skärpa (jfr linsformeln i ekv. 5.1, och Fig. 10.1).

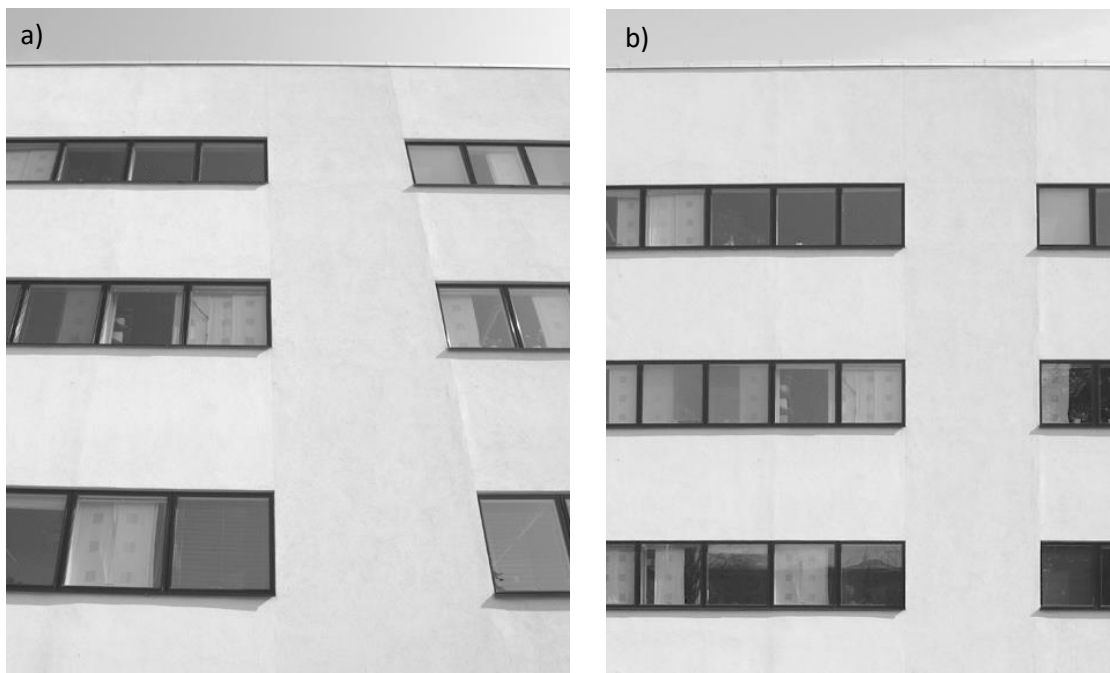


Fig. 6.4. a) ”Störtande linjer” har uppstått på grund av att kameran har lutats uppåt vid fotograferingen. Detta innebär att husfasaden återges distorderad, eftersom avbildningsskalan avtar ju längre upp i bilden man kommer. b) Störtande linjer kan undvikas genom att förskjuta objektivet uppåt relativt kamerahuset, så att optiska axel pekar vinkelrätt in mot husfasaden. I Fig. 6.5 visas strålgången i de bägge fallen.

Botemedlet mot både störtande linjer och oskärpa är att inte peka kameran uppåt, utan låta objektivets optiska axel vara vinkelrät mot husfasaden. Men då fick vi ju inte med hela husfasaden, eller hur? Det är just här som specialobjektivets ”shift”-funktion kommer in. Denna funktion innebär att man kan förskjuta objektivet i

höjddled relativt kamerahuset. Ett vanligt rörelseomfång brukar vara ca. 10-15 mm både uppåt och nedåt för en kamera med sensor i fullformat (24 mm x 36 mm). Om man låter objektivets optiska axel peka rakt in mot husfasaden, men höjer upp objektivet, så kan man få med hela husfasaden på höjden, se Fig. 6.5. Dessutom kommer hela fasaden att ligga i fokus. Enkelt och elegant! Objektiv med shift-funktion behöver konstrueras lite annorlunda, inte bara rent mekaniskt utan även optiskt. Det beror på att shift-funktionen innebär att man behöver utnyttja en större bildyta, och därmed större bildvinkel än man annars gör. Därför krävs att avbildningsfelen (kap. 15) är tillräckligt små även för dessa lite större bildvinklar.

Det kan påpekas att störtande linjer i digitala bilder kan korrigeras med datorprogram, t.ex. Photoshop. Jämfört med att använda ett tilt/shift-objektiv, innebär datormetoden två nackdelar. För det första kan den inte korrigera för att vissa motivdelar ligger ur fokus, och därför inte avbildas helt skarpt. För det andra innebär datormetoden en omräkning av pixelpositioner och pixelvärden, vilket kan ge en del förlust av bildinformation eftersom man behöver interpolera fram data som inte finns i ursprungsbilden. Bästa resultat kan alltså förväntas om ursprungsbilden är fri från störtande linjer.

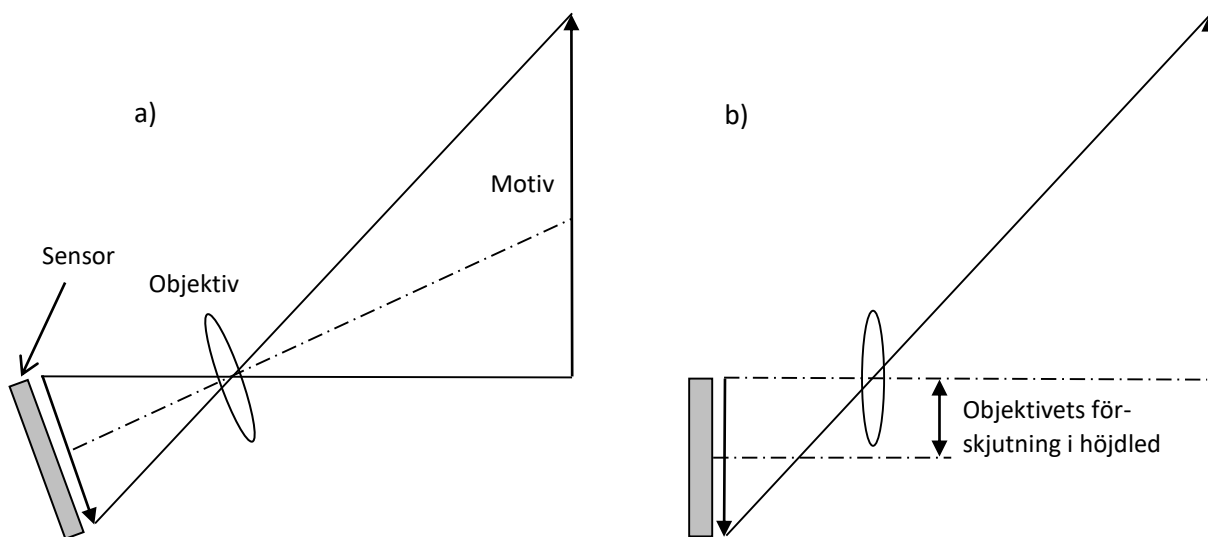


Fig. 6.5. a) Kameran lutar uppåt för att få med hela motivet på höjden. Detta ger störtande linjer, se beskrivning i texten och Fig. 6.4. Dessutom kommer bara en enda höjdnivå att avbildas skarpt. b) Genom att höja upp objektivet relativt sensorn, istället för att luta kameran, så undviker man störtande linjer. Dessutom kommer hela motivet att avbildas skarpt i bildplanet.

Förutom en förskjutning i höjddled, tillåter ett "tilt/shift"-objektiv även rotation runt en axel vinkelrät mot optiska axeln. Vanligen tillåts en rotation på ca. $\pm 10^\circ$. Detta gör att man kan åstadkomma sneda skärpeplan, vilket kan vara mycket användbart. Låt oss ta ett exempel för att belysa detta. Antag att vi står på ett långt rakt järnvägsspår, och vill ta en bild som visar hur spåret försvinner bort i fjärran, Fig.

6.6. För att kunna få skärpa både på förgrunden och bakgrunden krävs i en sådan situation ofta att man jobbar med mycket små bländaröppningar (se kap. 10 angående skärpedjup). Ibland kan man dock inte få tillräckligt skärpedjup ens med den minsta bländaröppningen. Små bländaröppningar medför också långa exponeringstider (risk för skakningsoskärpa), samt ger ofta lite sämre allmän skärpa på grund av diffraktion (se kap. 15). Det finns alltså flera skäl till att man kan vilja arbeta med lite större bländaröppningar, trots att motivet kräver stort skärpedjup.

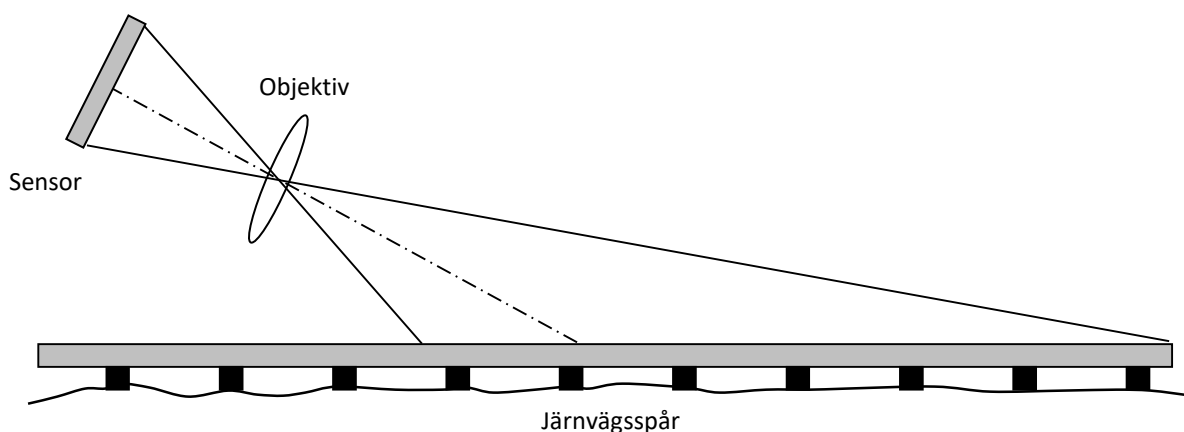


Fig. 6.6. Vid fotografering av motiv som ligger på "snedden" får man ofta problem med att skärpedjupet inte räcker till. I sådana fall är det fördelaktigt om också skärpeplanet kan fås att ligga på snedden. Detta kan uppnås med arrangemanget som visas i Fig. 6.7.

För de fall att motivet är ungefär plant, som i fallet järnvägsspår, så finns möjligheten att "tilta", dvs vrida på objektivet, se Fig. 6.7. Om man därvid uppfyller det så kallade Scheimpflug*-villkoret, så kan man visa att alla delar av motivet kommer att avbildas helt skarpt. Villkoret innebär att motivplanet, sensorplanet samt objektivets bakre huvudplan skall skära varandra längs en och samma linje. En sådan vridning kommer inte att påverka bildens perspektiv. Precis som för "shift"-funktionen som nämndes ovan, så krävs vid "tiltning" av objektivet att det är konstruerat för att klara av lite större bildvinklar än normalt.

I vissa krävande fotografiska situationer kan man samtidigt behöva använda både "shift"-funktionen och "tilt"-funktionen. Exempel på en sådan situation kan vara att man står snett framför en hög husfasad som ska fotograferas. Objektivet ställs då verkligen på prov vad gäller att klara av stora bildvinklar, och det är inte ovanligt att man ser någon felaktighet, t.ex. kraftigt ljusavfall, ute i ett bildhörn.

* Uppkallat efter den österrikiske armékaptenen Theodor Scheimpflug (1865-1911) som arbetade med att korrigera geometriska felaktigheter i flygbilder.

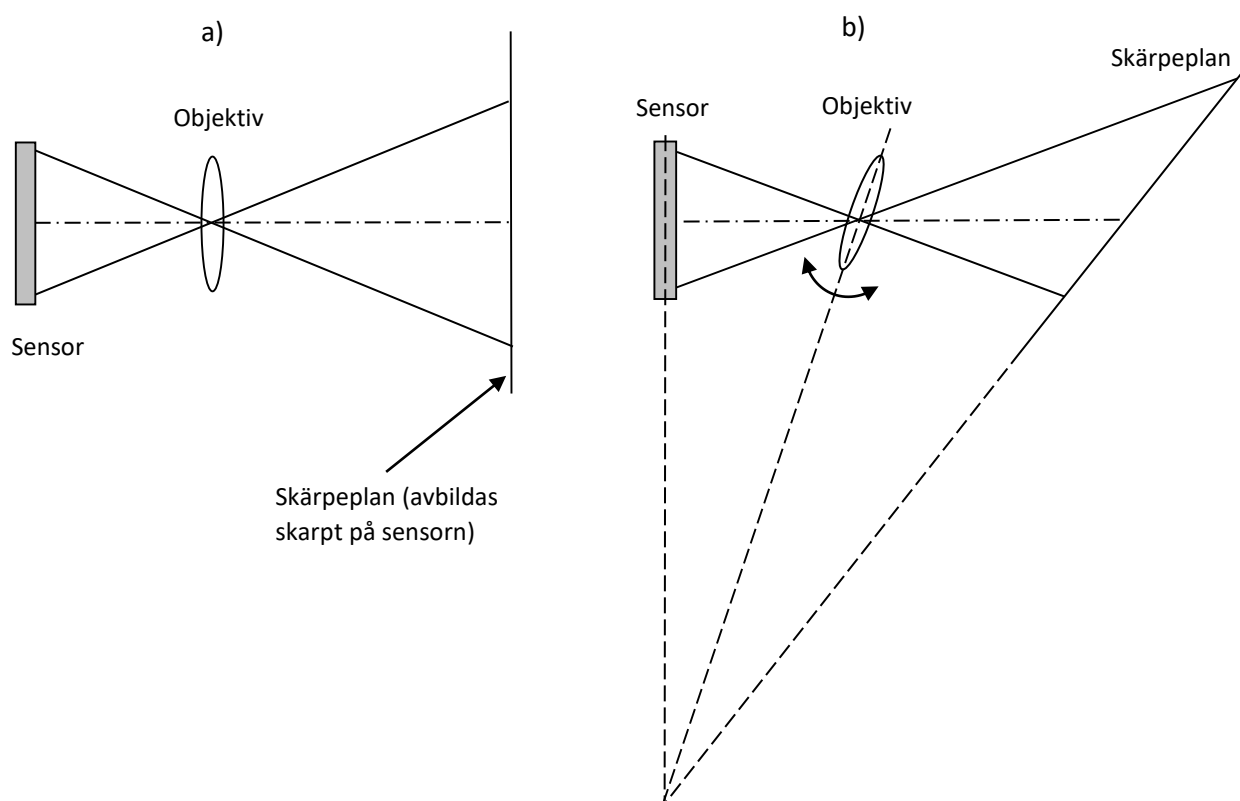


Fig. 6.7. a) Vid normal fotografering är sensorplanet, linsplanet och kamerans skärpeplan parallella med varandra. Detta innebär att man helt skarpt bara kan avbilda ett motivplan vinkelrätt mot linsens optiska axel. b) Genom att uppfylla det så kallade Scheimpflug-villkoret kan man få skärpeplanet att ligga på snedden. Villkoret innebär att motivplanet, sensorplanet och linsplanet (egentligen objektivets bakre huvudplan) ska skära varandra längs en linje. I ett "tilt/shift"-objektiv uppnås detta genom att objektivet kan roteras runt en axel genom bakre huvudplanet.

6.8. Teleförlängare

En teleförlängare (teleconverter) är en optisk enhet som monteras mellan objektivet och kamerahuset, varvid en förlängning av brännvidden erhålls. Förlängningsfaktorn är vanligen ca. 1.5 – 2. Teleförlängare används mest tillsammans med teleobjektiv, i de fall man önskar en mycket lång brännvidd. I Fig. 6.8. visas en förenklad bild av en teleförlängare monterad på en spegelreflexkamera.

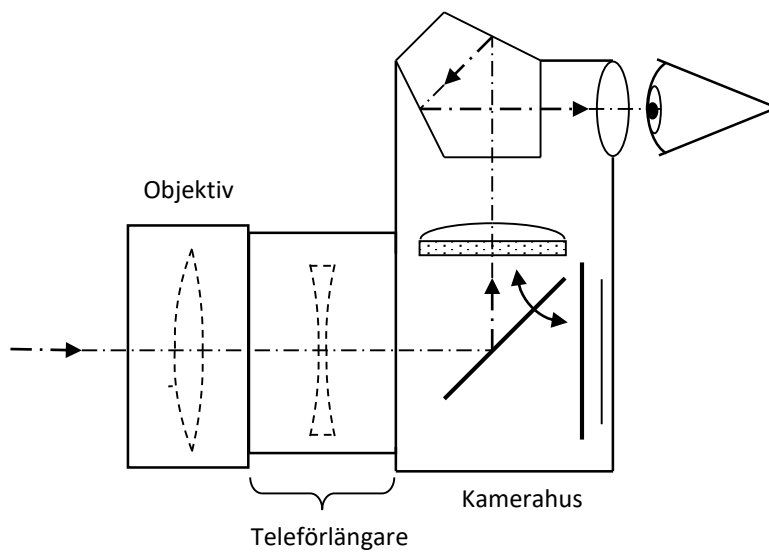


Fig. 6.8. Teleförlängare monterad mellan objektiv och kamerahus på en spegelreflexkamera. Optiken visas kraftigt förenklad, i verkligheten består både objektiv och teleförlängare av ett flertal linselement.

Teleförlängarens optiska princip är densamma som för teleobjektivet, Fig. 6.2, men med den skillnaden att den negativa linsen inte rör sig när objektivet fokuseras om, se Fig. 6.9. Teleförlängaren kommer att minska objektivets ljusstyrka med en viss faktor (vanligen ca. 2), och även bländartalen måste räknas om med samma faktor. Däremot stämmer avståndsskalan på objektivet även med en teleförlängare monterad. Eftersom objektiv med kortare brännvidd också oftast har en kortare närgräns (minsta avstånd till motivet), så kan det vara fördelaktigt att vid närbildsfoto använda ett objektiv med relativt kort brännvidd plus en teleförlängare, istället för ett objektiv med lite längre brännvidd.

För att få en bra bildkvalitet med teleförlängare krävs att denna är av hög kvalitet. Det betyder att den ofta består av 5-7 linselement för att minska avbildningsfelen som annars skulle uppstå (avbildningsfel beskrivs i kap. 15).

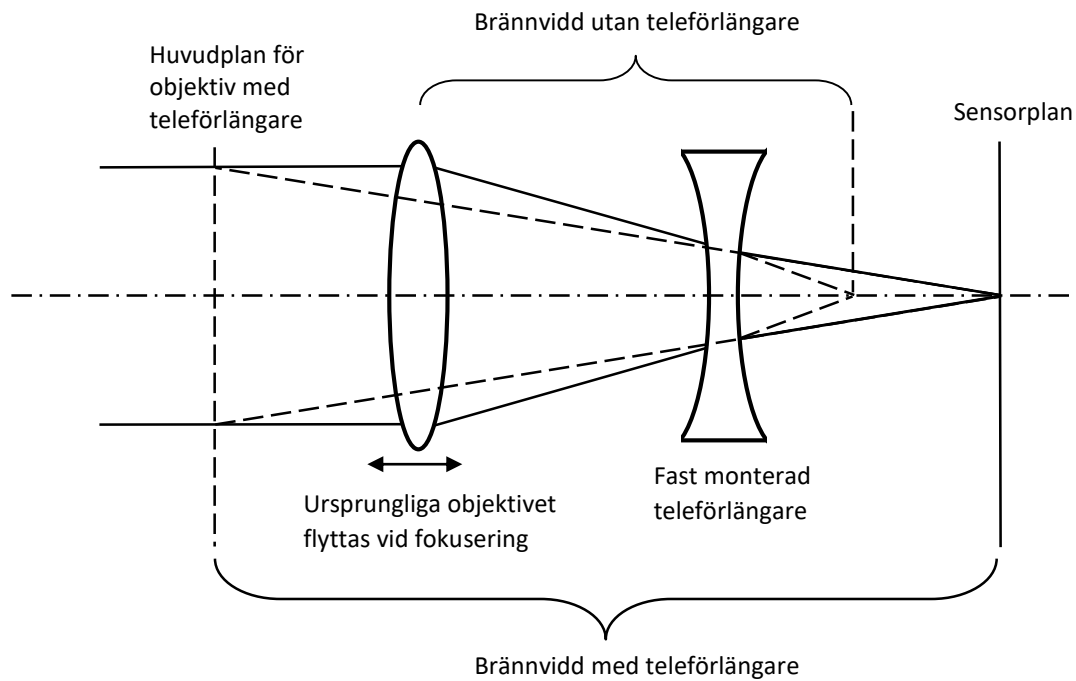


Fig. 6.9. Optiska principen för en teleförlängare är densamma som för ett teleobjektiv, jfr. Fig. 6.2. Det ursprungliga objektivet representeras av en positiv lins i denna förenklade figur, och teleförlängaren representeras av en negativ lins. Teleförlängaren är fast monterad på ett visst avstånd från sensorn, medan ursprungliga objektivet kan förflyttas längs optiska axeln för att fokusera bilden på sensorn.

7. Perspektiv

Mycket har skrivits och pratats om perspektiv vid fotografering. I denna framställning ska vi bara mycket kort beröra detta ämne ur ett fysikaliskt perspektiv (förlåt vitsen!). Det är två enkla frågeställningar vi ska titta på. Dels vad vi menar med att en bild har ett korrekt perspektiv, dels hur vi ska uppnå detta korrekta perspektiv. Det är naturligtvis inte alltid man önskar ett korrekt perspektiv (t.ex. av konstnärliga skäl), men det är vanligen en bra början att lära sig hur man ska göra för att saker och ting ska bli korrekta, sedan kan man avsiktligt bryta mot dessa regler om man vill.

Perspektiv handlar om djupintryck i en bild. Med korrekt perspektiv avser vi helt enkelt att när vi tittar på en bild, så ser vi alla föremålen som fotograferats under samma synvinklar som om vi stod på fotograferingsplatsen och betraktade motivet direkt (med ett öga; stereoverkan p.g.a. att vi har två ögon behandlas i nästa kapitel). Om detta är uppfyllt får vi alltså samma intryck av djup i motivet vare sig vi tittar på fotografiet eller direkt på motivet. Korrektare avbildning än så kan vi inte få, och detta definierar vi som korrekt perspektiv. Låt oss nu titta på hur vi kan åstadkomma ett sådant korrekt perspektiv vid den fotografiska återgivningen. Fig. 7.1 visar strålgången dels vid fotograferingstillfället, dels när vi med ögat tittar på bilden.

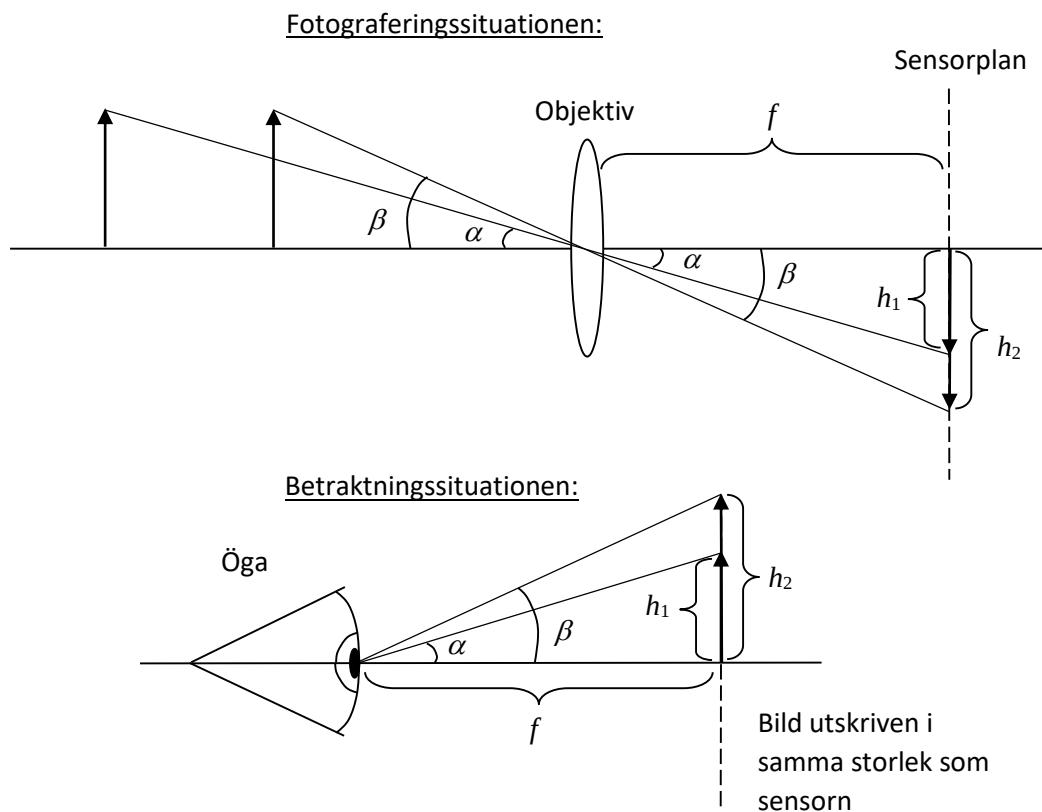


Fig. 7.1. Villkor för erhållande av korrekt perspektiv i en bild.

För enkelhets skull har vi i figuren antagit att de två objekten, illustrerade som pilar, befinner sig på stort avstånd från kameran (även om det p.g.a. utrymmesskäl inte ser ut så i figuren) och därför avbildas bägge skarpt i linsens fokalplan där sensorn befinner sig. Endast de centrala strålarna, som går obrutna genom linsens centrum har ritats ut. Vinklarna α och β är de synvinklar under vilka objekten ses från fotograferingsplatsen. Objekten kommer i sensorplanet att avbildas med höjderna h_1 och h_2 . Antag att vi skriver ut vår bild i samma storlek som sensorn har, samt därefter betraktar densamma enligt figuren. Genom att betrakta bilden på ett avstånd som motsvarar brännvidden på objektivet som användes vid fotograferingen, kommer vi att se objekten under precis samma vinklar som om vi stod på fotograferingsplatsen och tittade direkt på motivet. Korrekt perspektiv med andra ord.

Nu kan man anmärka att den beskrivna betraktningssituation inte är realistisk av två skäl. Dels är brännvidden på objektivet ofta mycket kortare än ett bekvämt betraktningsavstånd, vilket gör att vi inte med ögat kan fokusera så att vi ser en skarp bild. Detta kan emellertid lösas om vi sätter ett förstoringsglas eller en lupp (omedelbart) framför ögat. Detta är tillåtet, och kommer inte att förändra perspektivet. Den andra anmärkningen man kan göra är att den utskrivna bilden så gott som alltid är betydligt större än sensorn. Detta är emellertid lätt att ta hänsyn till i resonemanget om perspektiv. Om förstoringsgraden från sensor till slutbild är M_{kopia} , så kommer objektens höjder i figuren att bli $h_1 M_{kopia}$ respektive $h_2 M_{kopia}$ (förstoringsgraden gäller linjär förstoring, inte areaförstoring). Om vi ökar på betraktningsavståndet till $f M_{kopia}$, så kommer synvinklarna fortfarande att bli desamma som tidigare och perspektivet blir korrekt. Ofta skriver man inte ut sina bilder, utan visar dem på en datorskärm eller en vit duk. Även i detta fall gäller att korrekt betraktningsavstånd är lika med $f M_{kopia}$, där M_{kopia} i detta fall är förstoringsgraden från sensor till slutbild. Sammanfattningsvis gäller alltså:

$$s = f \times M_{kopia} \quad (7.1)$$

där s är korrekt betraktningsavstånd till bilden, f är kameraobjektivets brännvidd och M_{kopia} är förstoringsgraden från sensorn till den förstorade (eller ev. förminskade) bild som betraktas. Man inser lätt att ovanstående formel också gäller vid närbildsfotografering om brännvidden (f) ersätts med bildavståndet (b) som erhålls ur linsformeln.

Om man använder ett annat betraktningsavstånd än det som ges av ekv. 7.1, så erhålls en felaktig djupuppfattning. Därvid gäller att ett för stort betraktningsavstånd ger en överdriven djupverkan, vilket är ett välkänt fenomen för alla som betraktat bilder tagna med vidvinkelobjektiv. Motsatsen, ett för kort betraktningsavstånd, ger för liten djupverkan och därmed en ”platt” bild. Detta är också ett välkänt fenomen för alla som betraktat bilder tagna med teleobjektiv.

Låt oss ta ett enkelt exempel för att få lite realistiska siffror:

Antag att vi med en fullformatskamera (sensorformat 24 mm x 36 mm) fotograferar ett motiv på stort avstånd med ett 55 mm objektiv (typiskt "normalobjektiv" på en systemkamera). Bilden skrivs ut på ett 10 cm x 15 cm papper. Förstoringsgraden blir i detta fall 4.17, vilket ger att betraktningssavståndet $55 \times 4.17 = 230$ mm ger ett korrekt perspektiv. Detta är ett ganska normalt betraktningssavstånd, vilket gör att när man tittar på "normalstora" utskrifter av bilder tagna med "normalobjektiv" så får man ungefär korrekt perspektiv (vilket naturligtvis är avsikten). Om man däremot hade fotograferat med ett 28 mm vidvinkelobjektiv eller 135 mm teleobjektiv, så skulle motsvarande korrekta betraktningssavstånd bli 12 cm resp. 56 cm. Eftersom man normalt betraktar en bild på ca. 25 cm avstånd, så innebär det att vidvinkelbilden betraktas på för stort avstånd medan telebilden betraktas på för kort avstånd. Detta kommer att yttra sig i ett felaktigt perspektiv, och medför en överdriven djupverkan i vidvinkelfallet medan djupverkan blir för liten (platt) i telefallet.

I Fig. 7.2 och 7.3 visas exempel på hur olika samma motiv kan se ut på bilder som tagits med olika brännvidder på objektivet. Bilder tagna med vidvinkeloptik betraktas ofta på för stort avstånd, och ger därmed en överdriven djupverkan. Bilder tagna med teleobjektiv betraktas däremot ofta på för kort avstånd och ger därmed "underdriven" djupverkan och uppfattas som "platta". En överdriven djupverkan kan vara önskvärd i många sammanhang, det ger t.ex. dramatiska bilder av natur och arkitektur. För porträttbilder är däremot ett lite "plattare" perspektiv ofta att föredra (se Fig. 7.3). Annars upplevs ofta näsor som stora och utstickande, ansiktet som smalt och öronen som små. I praktiken betyder det att för en fullformatskamera (sensorstorlek 24 mm x 36 mm) så kan ett 35 eller 28 mm objektiv passa för natur- och arkitekturbilder, medan ett 90 mm objektiv kan passa för porträttfoto. Men naturligtvis måste man beakta hur mycket slutbilden kommer att förstöras samt betraktningssavståndet. De angivna brännvidderna baseras på en slutbild av ca. 10 x 15 cm betraktad på ca. 25 cm avstånd.

Om en bild betraktas på felaktigt avstånd så kommer inte bara djupverkan att bli fel. Föremålen kommer också att verka geometriskt distorderade utåt bildkanten. Exempelvis så kommer sfäriska föremål att se ut som ellipser, vilket illustreras i Fig. 7.4. Detta har inget med avbildningsfelet distorsion att göra (se avsnitt 15.8), eftersom det endast uppträder för tredimensionella och inte för plana motiv. I Fig. 7.5. visas hur fenomenet uppstår. Effekten är mest påtaglig i bilder som tagits med korta brännvidder (vidvinkeloptik). Den kan därför ställa till problem när man vill skapa ett dramatiskt överdrivet perspektiv genom att använda en kort brännvidd vid fotograferingen.

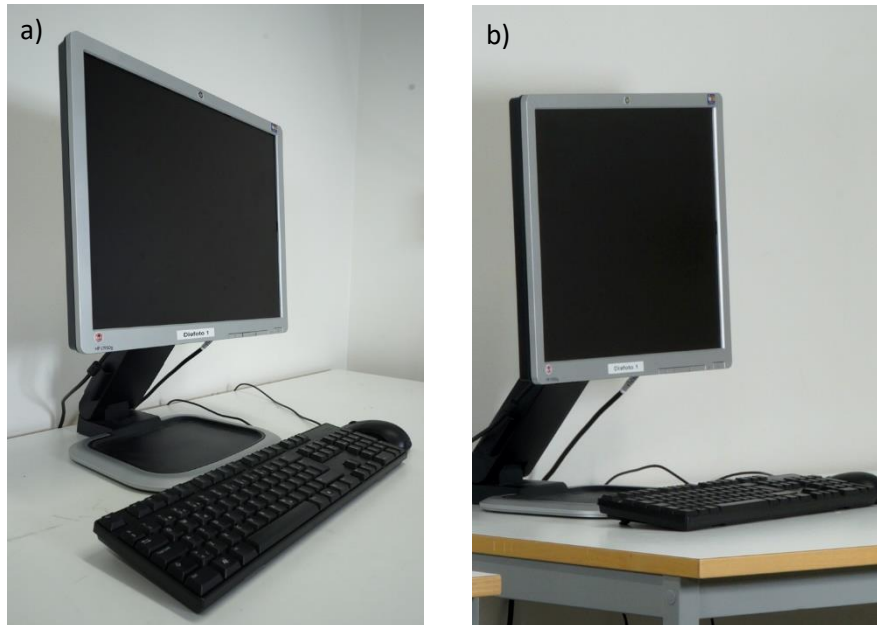


Fig. 7.2. Samma datorskärm fotograferad med brännvidd 28 mm (a) och 280 mm (b). Bilderna visas utan någon beskärning. I första fallet ser skärmen betydligt bredare ut ("widescreen") jämfört med andra fallet ("narrowscreen"). Korta brännvidder ger ett större djupintryck än långa brännvidder för samma betraktningsavstånd. Om bilderna i figuren upptar en bredd av 10 cm vardera på en datorskärm, erhålls korrekt perspektiv för betraktningsavstånd av ca. 12 cm för a-bilden och 120 cm för b-bilden.

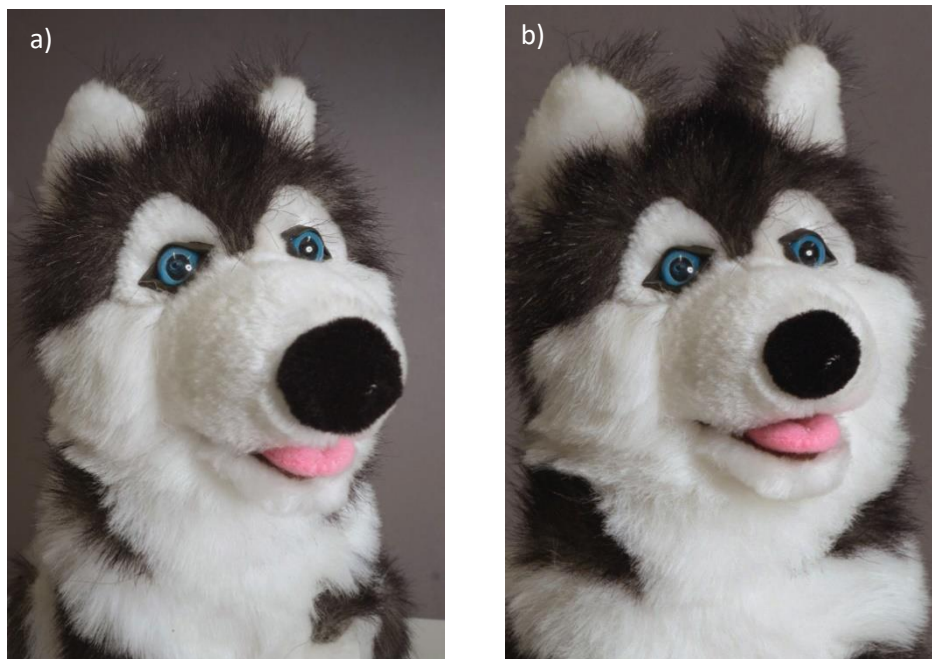


Fig. 7.3. "Människans bästa vän" fotograferad med brännvidd 28 mm (a) och 280 mm (b). Notera skillnaden i nosens storlek, bredden på huvudet och öronens storlek. På vilken av bilderna ser hunden "gulligast" ut? Även porträtt av människor görs bäst med lite längre brännvidder (se text). Bilderna visas utan någon beskärning, och korrekt perspektiv erhålls på samma sätt som enligt Fig. 7.2.



Fig. 7.4. Bild av två glödlampor tagen med vidvinkelobjektiv. Glödlampan närmast bildkanten ser deformerad ut beroende på ett perspektivistiskt fel som uppstår när bilden betraktas på felaktigt avstånd. Vid korrekt betraktningsavstånd (ca. 6 cm om bildbredden är 10 cm på en datorskärm) så försvinner felet. En förklaring till hur fenomenet uppstår ges i Fig. 7.5.

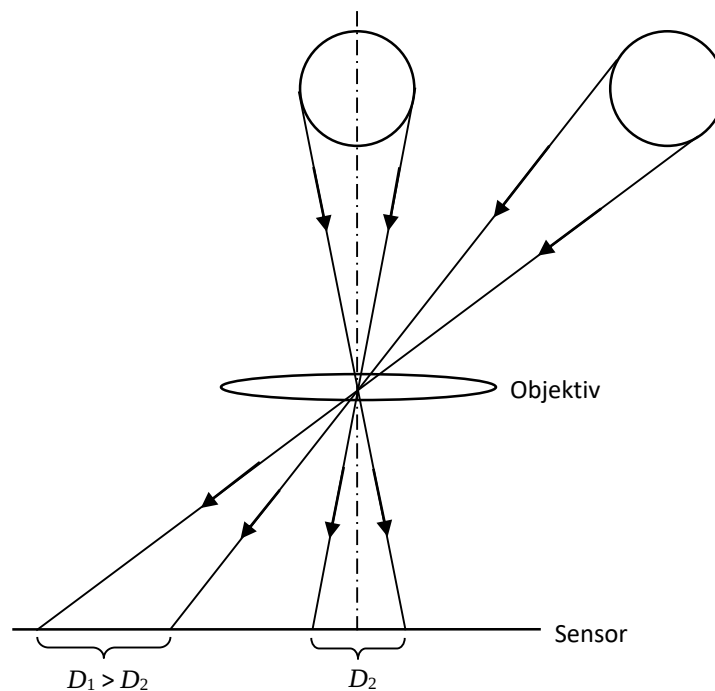


Fig. 7.5. Figuren visar hur två sfäriska föremål avbildas på kamerans sensor. Eftersom den tredimensionella verkligheten avbildas på ett plan, så kommer sfären ute i bildkanten att avbildas utdragen i sidled. Effekten förekommer vid fotografering av alla objekt som har en utsträckning i djupled, men är mest störande för enkla regelbundna former som vi vet hur dom ska se ut. Observera att effekten inte är störande, utan tvärtom önskvärd, för det fall att bilden betraktas på korrekt avstånd och med ögat rakt ovanför bildens centrum. Då kommer vi nämligen att uppleva precis samma sak som när vi stod på fotograferingsplatsen och tittade på motivet.

8. Stereofotografi

Kap. 7 behandlade perspektivet, dvs djupintrycket man får när man betraktar en bild. Det som åstadkommer djupintrycket är inbördes storleksförhållandena mellan de föremål som syns i bilden. Om vi t.ex. ser två elefanter på en bild, och den ena är mycket mindre än den andra, så tolkar vi detta så att den elefant som återges liten befinner sig längre bort. Men vi kan ju inte helt säkert veta detta, det är vår erfarenhet som säger oss att elefanter inte är pyttesmå. Om vi istället betraktar en bild med föremål som vi inte vet hur stora de kan förväntas vara, så blir det mycket svårare att uppleva djup i bilden. Vi kan t.ex. tänka oss att bilden föreställer en snårig skog. Vi ser då en massa kvistar med olika grovlekar, men vi har svårt att tolka bilden eftersom vi inte vet om en kvist är smal och befinner sig nära kameran, eller om den är grov och befinner sig längre bort. Eftersom vi har svårt att tolka bilden, känns den "orolig" att titta på. Erfarna fotografer undviker oftast sådana motiv. Om vi ute i naturen betraktar samma snårskog med egna ögon känns inte motivet alls så "oroligt". Det beror på att de flesta av oss (dock inte alla) kan göra en tredimensionell tolkning av motivet genom att vi har stereoseende. Våra två ögon ger oss något olika bilder av omgivningen eftersom de sitter 6-7 cm från varandra. Detta avstånd kallas ögonbasen. Vår hjärna har lärt sig att tolka de subtila skillnaderna i våra två näthinnebilder och översätta dessa till djupinformation. Vårt stereoseende ger oss alltså möjlighet att avgöra avståndet till olika kvistar i snårskogen, och genom att vi kan tolka vad vi ser känns inte längre motivet "oroligt". Vi skulle också kunna avslöja en bluff med en pytteliten elefant på samma avstånd som en stor elefant.

Man kan naturligtvis också fotografera i stereo. Man tar då två fotografier, ett stereopar, från lite olika positioner i sidled. Ett tredimensionellt motiv kommer då att se lite olika ut i de bägge bilderna, vilket kan användas för att få en sann djupinformation. Man kan ur sådana bildpar mäta upp tredimensionella läget av objekt ute i terrängen, vilket används inom vetenskapsgrenen fotogrammetri. Ett enkelt exempel på detta visas i Fig. 8.1.

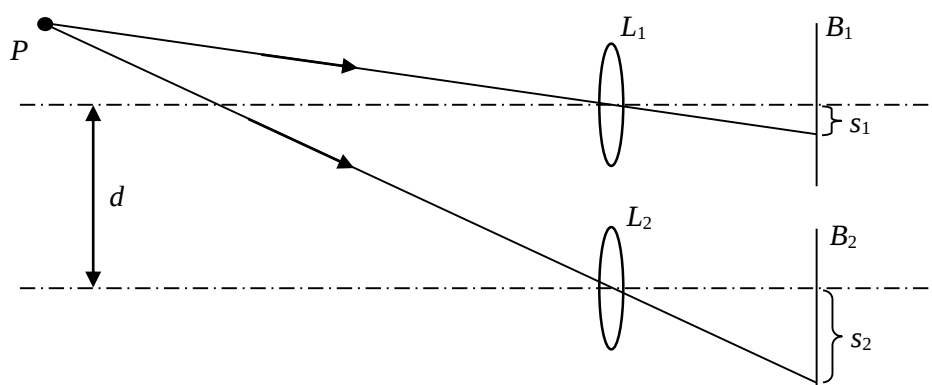


Fig.8.1. Stereofotografering . Två kameror med linserna L_1 och L_2 fotograferar samtidigt motivpunkten P . Denna avbildas på olika ställen i de två bilderna B_1 och B_2 (på avstånden s_1 resp. s_2 från bildcentrum). Genom att mäta upp s_1 och s_2 kan man beräkna P 's läge om man känner brännvidderna och stereobasen d . I stället för att använda två kameror kan man använda en enda kamera som flyttas mellan två lägen. Det finns också speciella stereokameror med två objektiv inbyggda sida vid sida.

Om man betraktar ett fotografiskt stereopar så att vänstra ögat bara ser vänster bild, och höger öga bara höger bild, kan man få ett mycket kraftfullt tredimensionellt intryck. Bilder som är helt onjutbara som monobilder kan bli fantastiskt vackra i stereo. I själva verket är stereoseendet mycket komplicerat och består av flera komponenter. Dels har vi de tidigare påtalade skillnaderna i näthinnebilderna för vänster och höger öga, dels kan våra hjärnor känna av och tolka konvergensvinkeln för våra ögon när vi fäster blicken på en viss motivpunkt, se Fig. 8.2. Vi ska inte här gå närmare in på stereoseendets mekanismer, utan vi ska bara göra en enkel betraktelse över vilka resultat vi kan förvänta oss för några olika fotograferingssituationer.

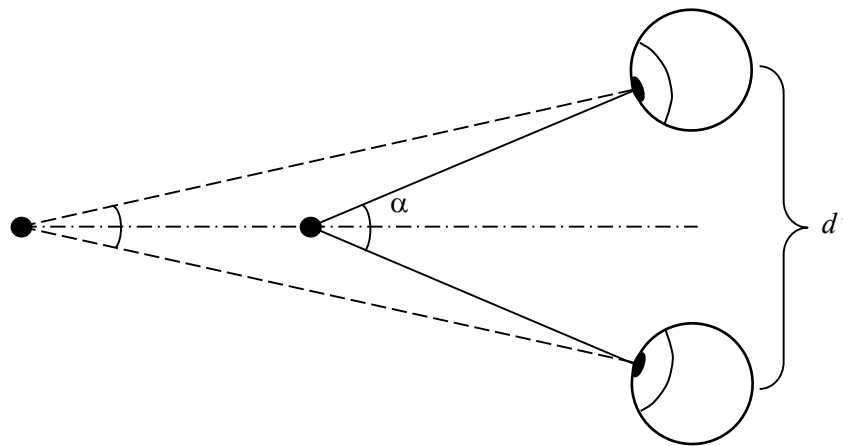


Fig. 8.2. Konvergensvinkeln α för våra ögon blir större ju närmare föremålet är som vi tittar på.

Precis som vi i Kap. 7 ställde oss frågan vad som menas med en perspektivistiskt korrekt bild, kan vi för stereofotografering fråga vad som menas med en stereoskopiskt korrekt bild. Svaret blir mycket likartat. Precis som i perspektivfallet vill vi att motivets olika delar ska återges, så att vi i bilderna ser dem under samma synvinklar som om vi stod på fotograferingsplatsen och tittade på motivet. Men i stereofallet måste vi dessutom lägga till att vi ska uppleva att deras avstånd i djupled ska stämma med verkligheten. Dessa villkor kommer att vara uppfyllda om vi fotograferar med en stereobas (d i Fig. 8.1) som är lika med ögonbasen (d' i Fig. 8.2). Bilderna ska sedan betraktas på ett avstånd som överensstämmer med brännvidderna på kameraobjektiven (om bilderna är förstörade ska betraktningsavståndet ökas med samma faktor). Avståndet mellan bildernas centrumpunkter (dvs. där optiska axeln träffade sensorplanet) ska vara lika med stereobasen (som är lika med ögonbasen). Om vi följer dessa regler kommer vi att se precis samma saker som om vi stod på fotograferingsplatsen och betraktade föremålet med bägge ögonen samtidigt. (En enda perspektivistiskt korrekt bild motsvarar det vi ser på fotograferingsplatsen om vi bara tittar med ett öga.) Naturligtvis kan man bryta mot dessa regler för att få "häftiga" bilder. Exempelvis brukar stereofotografer experimentera en hel del med olika värden på stereobasen. Men det är då bra att ha en förståelse för vad som händer om man låter stereobasen avvika från ögonbasen.

Låt oss först ta fallet att vi använder en stereobas som är större än ögonbasen vid fotograferingen, men att bilderna fortfarande betraktas så att deras centrumpunkter ligger på ögonbasens avstånd. I detta fall kommer skillnaderna mellan de två delbilderna att bli större än vid ”korrekt” fotografering (exempelvis blir skillnaden mellan s_1 och s_2 i Fig. 8.1 större). Detta tolkar vår hjärna så att föremålen ligger närmare oss än vad som är fallet. Men eftersom föremålen fortfarande har samma storlek på bilderna kommer vår hjärna att tolka det som att föremålen är mindre än dom i verkligheten är. Alltså, för stor stereobas vid fotograferingen ger stereobilder där föremålen ser mindre ut, och ligger närmare oss än vad dom gör i verkligheten. Denna effekt är påtaglig i t.ex. flygbilder av marken där man ofta använder en mycket stor stereobas (tiotals till hundratals meter). Vid stereobetraktning av sådana bilder får man en stark känsla av att titta på en liten modell av ett landskap där husen ser ut som små leksakshus. En överdriven stereobas är dock nödvändig vid fotografering på stora avstånd, eftersom ögats förmåga att uppleva stereoeffekt blir knappt märkbar på större avstånd än några tiotals meter. Om stereobasen vid fotograferingen är mindre än ögonbasen ser istället föremålen ut att ligga längre bort, och vara större än i verkligheten. Förutom att variera stereobasen, kan man experimentera med att vinkla kamerorna på olika sätt. Särskilt vid närbildsfotografering är det användbart att låta kamerornas ”tittriktningar” konvergera precis som för ögonen i Fig. 8.2.

Till sist några ord om betraktning av stereobilder, samt några bildexempel. Vanligtvis behöver man något hjälpmedel för att kunna se på ett stereopar. Om man bara lägger ett par bilder sida vid sida framför sig, så tittar man helt naturligt på samma bild med bägge ögonen. För stereobetraktning ska emellertid vänster- och högeröga titta på olika bilder. Detta kan åstadkommas med ett enkelt stereoskop, t.ex. av den typ som visas i Fig. 8.3. Har man inget stereoskop kan man istället hålla två små identiska, kraftiga förstoringsglas framför ögonen när man betraktar bilderna. Med lite träning kan också de flesta lära sig att betrakta stereobilder helt utan hjälpmedel. Det gäller då att frikoppla ögonens fokusering och konvergensvinkel (α i Fig. 8.2.). Det är också viktigt att inte avståndet mellan bildernas centrumpunkter är för stort (max. ca. 60 mm).

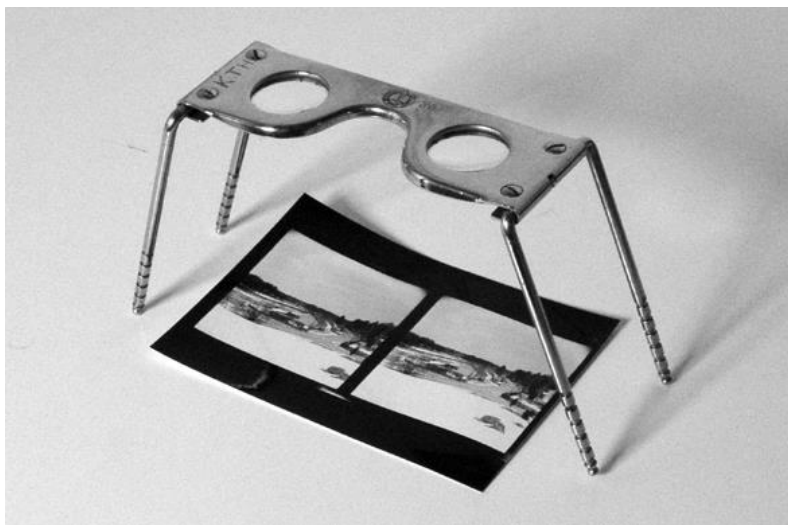


Fig. 8.3. Ett stereoskop underlättar betraktandet av stereobilder.



Fig. 8.4. Stereobild av blommande liljor.

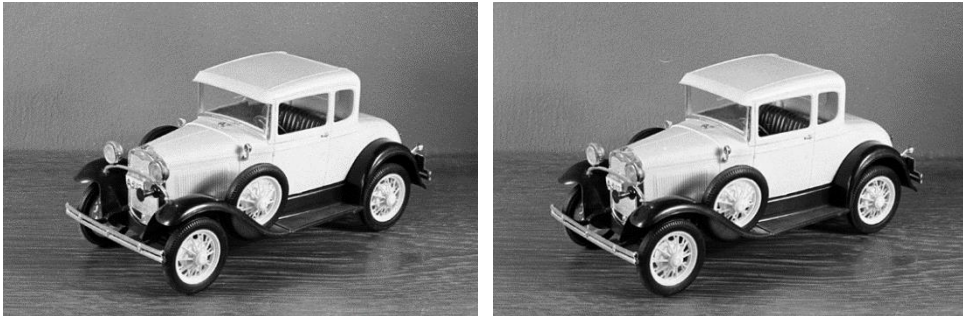


Fig. 8.5. Stereobild av modell av A-Ford.



Fig. 8.6. Stereobild av stadsbebyggelse. På grund av den stora stereobasen får man intryck av att betrakta små leksakshus.

9. Fotometri

9.1. Rymdvinkel

I detta kapitel ska vi bland annat titta på hur belysningen i kamerans sensorplan beror på faktorer som motivluminans, bländartal, avbildningsskala och bildvinkel. Innan vi ger oss i kast med detta ska vi emellertid gå igenom några grundläggande begrepp inom områdena fotometri och radiometri. Låt oss börja med begreppet rymdvinkel som definieras i Fig. 9.1.

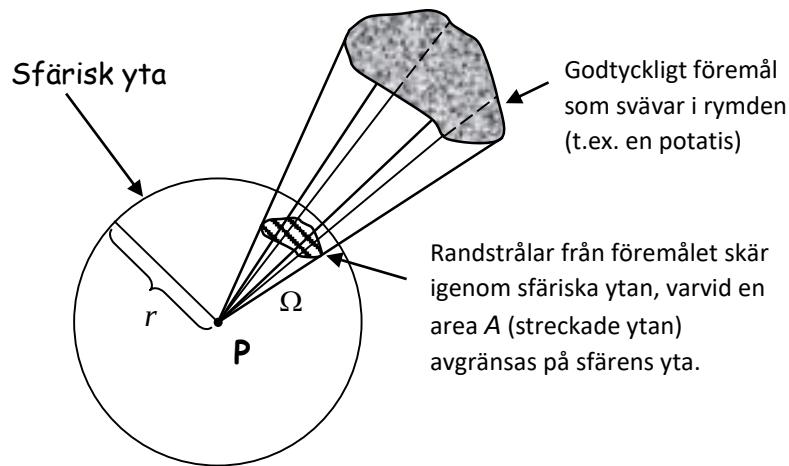


Fig. 9.1. Illustration av begreppet rymdvinkel. Den rymdvinkel, Ω , under vilken vi från punkten **P** ser föremålet definieras genom formeln $\Omega = \frac{A}{r^2}$. Största möjliga rymdvinkel är alltså 4π , vilket innebär att föremålet helt omsluter oss (när vi står på jordytan upptar himmelskupolen rymdvinkeln 2π). Enheten för rymdvinkel kallas steradian och förkortas sr.

9.2. Radians

Vi betraktar nu en svartkroppsstrålare, dvs. en kropp vars utstrålning ges av Plancks strålningslag, se läroböcker i fysik. Vi tittar på hur stor effekt, P , som strålar ut från en area A inom en viss rymdvinkel Ω , Fig. 9.2.

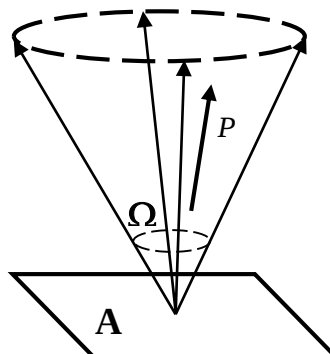


Fig. 9.2. Illustration av begreppet radians.

Denna storhet, dvs. utstrålad effekt per area- och rymdvinkelenhet, kallas **radians** och kan skrivas som $R = \frac{P}{A\Omega}$. Om vi låter storleken på yt- och rymdvinkelelementen gå mot noll får vi $R = \frac{d^2P}{dAd\Omega}$. Radiansen R ökar med temperaturen för svartkroppsstrålaren. Det visar sig emellertid att det strålas ut mindre effekt snett åt sidan än vinkelrätt ut från ytan, Fig. 9.3.

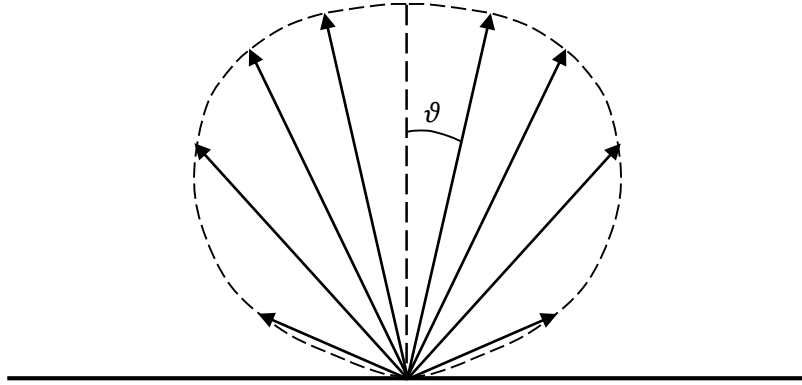


Fig. 9.3. Utstrålningens riktningsberoende.

Ovanstående formel för radiansen gäller därför bara för en strålningsriktning vinkelrätt ut från ytan. Riktningsberoendet ges av en faktor $\cos\theta$, vilket innebär att det i 45° -riktningen strålar ut bara 71 % så mycket effekt per area- och rymdvinkelenhet som det gör vinkelrätt från ytan. I en riktning parallellt med ytan har utstrålningen sjunkit till noll. För att ta hänsyn till detta riktningsberoende definierar man radiansen enligt formeln

$$R = \frac{d^2P}{dAd\Omega \cos\theta}. \quad (9.1)$$

För en svartkroppsstrålare beror R bara på kroppens temperatur enligt formeln $R = 1.80 \times 10^{-8} \times T^4$ [$\text{Wm}^{-2}\text{sr}^{-1}$], där T är absoluta temperaturen.

9.3. Irradians

För instrålad effekt mot en yta inför vi begreppet **irradiance**,

$$I = \frac{dP}{dA} \quad (9.2)$$

Effekten dP är integrerad över alla infallsvinklar för vilka strålning faller in mot ytelementet dA .

9.4. Ljusflöde och ljusstyrka

Inom fotografien brukar man oftast använda fotometriska enheter snarare än radiometriska. Enda skillnaden mellan dessa enheter är att de fotometriska är viktade med avseende på ögats spektrala känslighet. Fotometriska motsvarigheten till strålningseffekten P (ibland kallad strålningsflöde) är sålunda **ljusflödet** Φ , definierat genom

$$\Phi = \int_0^{\infty} P_{\lambda}(\lambda) V(\lambda) d\lambda, \quad (9.3)$$

där $P_{\lambda}(\lambda)$ är strålningseffekt per våglängdsintervall och $V(\lambda)$ är ögats spektrala känslighetskurva, Fig. 9.4. Sorten för ljusflöde kallas lumen (förkortat lm). En 60 W glödlampa (eller 10 W LED-lampa) har ett ljusflöde av ca. 700 lm.

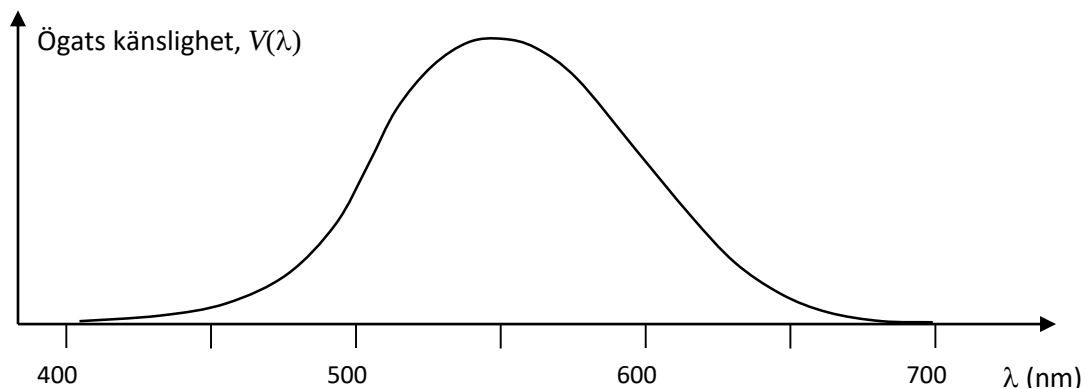


Fig. 9.4. Ögats spektrala känslighetskurva.

För riktade ljuskällor, typ spotlights, kan det vara intressant att veta hur stort ljusflöde som sänds ut per rymdvinkelenhet, snarare än det totala ljusflödet. Utsänt ljusflöde per rymdvinkelenhet kallas "ljusstyrka" och har sorten candela (förkortas cd) som är samma sak som lumen/steradian.

9.5. Luminans

Vi kan nu införa storheten **luminans**, L , vilken är den fotometriska motsvarigheten till radianzen R . Vi byter helt enkelt ut strålningseffekt mot ljusflöde, vilket ger

$$L = \frac{d^2\Phi}{dA d\Omega \cos\theta} \quad (9.4)$$

För en svartkroppsstrålarare får vi, precis som för radianzen, att L är en konstant vars värde endast beror på kroppens temperatur. Solytan (temperatur 5800 K) har en luminans av ca. $1.9 \times 10^9 \text{ lm m}^{-2} \text{sr}^{-1}$. På jordytan gäller då att en vit (80% reflektion), diffust reflekterande yta, som belyses av solen stående i zenit, har en luminans $L \approx 2.5 \times 10^4 \text{ lm m}^{-2} \text{sr}^{-1}$.

9.6. Belysning

Vi ska nu titta på den fotometriska motsvarigheten till storheten irradians, I , som ju betecknar infallande strålningseffekt per ytenhet. Genom att byta ut effekt mot ljusflöde får vi storheten **belysning**, som säger oss hur stort infallande ljusflöde (integrerat över alla riktningar) som vi har per ytenhet. Vi ska beteckna belysning med bokstaven E , och får då att

$$E = \frac{d\Phi}{dA}, \quad (9.5)$$

där Φ betecknar ljusflöde och A area. Sorten för belysning blir lumen/m², vilket även kallas **lux**. Som exempel kan nämnas att en 60 W glödlampa (eller 10 W LED-lampa), som sitter en halvmeter över ett bord, ger en belysning av 200 – 300 lux på bordsytan.

9.7. Exponering

Vi kan nu passa på att införa begreppet **exponering**, H , som är produkten av belysningen i sensorplanet, E , och exponeringstiden t .

$$H = E \cdot t \text{ [luxs]} \quad (9.6)$$

9.8. Radiometriska och fotometriska storheter

Nedan återfinns små sammanställningar i tabellform av radiometriska och fotometriska storheter. Radiometriska storheter används för att beskriva strålade energi, effekt, effekttäthet mm. De grundläggande fysikaliska enheterna joule, watt etc. används.

Storhet	Enhet
Strålningsflöde (-effekt)	W
Radians	W m ⁻² sr ⁻¹
Irradians	W m ⁻²

Såsom redan nämnts motsvarar de fotometriska storheterna de radiometriska. Enda skillnaden är att de fotometriska storheterna är viktade med avseende på det mänskliga ögats spektrala känslighetskurva. Sålunda kommer våglängder runt 550 nm att ha den högsta viktfaaktorn, medan kortare och längre våglängder får en lägre faktor (ju längre bort från 550 nm desto lägre). Våglängder utanför det synliga området får viktfaaktorn noll. De fotometriska storheter som svarar mot de tre radiometriska ovan ges av tabellen på nästa sida. Inom fotografin använder man som regel fotometriska storheter.

Storhet	Enhet
Ljusflöde	lm (lumen)
Luminans	$\text{lm m}^{-2} \text{sr}^{-1}$
Belysning	lm m^{-2} (lux)

9.9. Lambertreflektor (matt yta)

I verkligheten händer det ofta att man fotograferar motiv som inte är så varma att de glöder! Det vanligaste är ju att man fotograferar reflekterat ljus från motivet, och därför ska vi nu införa begreppet matt reflekterande yta. I verkligheten är långt ifrån alla ytor matta. Många är blanka eller halvmatta, vilket betyder att de fungerar som (mer eller mindre bra) speglar för infallande ljus, Fig. 9.5.

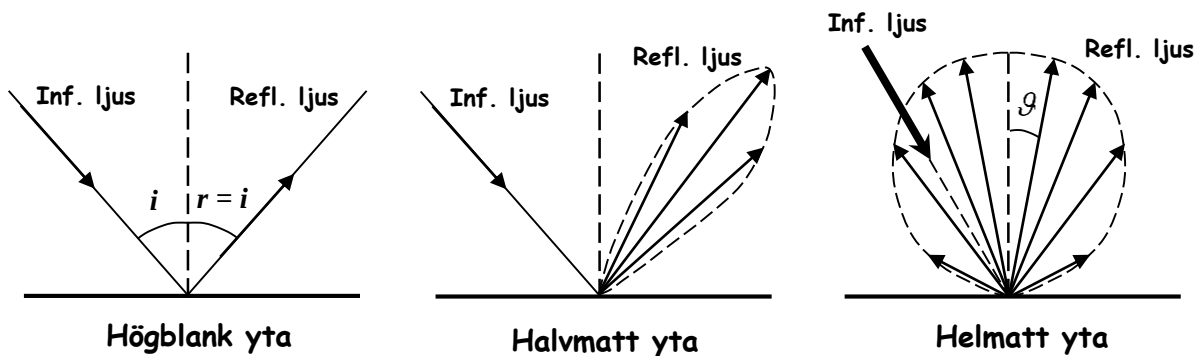


Fig. 9.5. Illustration av reflektion från ytor med olika grad av matthet.

För den helmatta ytan gäller att reflekterat ljus sprids symmetriskt runt ytnormalen. Spridningen är kraftig, och inget spår finns i det utgående ljuset av från vilken riktning det infallande ljuset kommer. En sådan perfekt helmatt yta kallas ibland för Lambertreflektor. Dess luminansvärde, L , kommer att bero enbart på hur kraftigt den belyses samt på dess reflektionskoefficient. Sålunda är L riktningsoberoende, dvs. den är *ej* en funktion av Θ . Detta innebär att det ljusflöde, $d^2\Phi$, som sänds ut från ytelementet dA under rymdvinkeln $d\Omega$ i en vinkel Θ från ytnormalen, Fig. 9.6, ges av (se ekv. 9.4) $d^2\Phi = L \cdot dA \cdot d\Omega \cdot \cos\Theta$, där $L = \text{konstant}$.

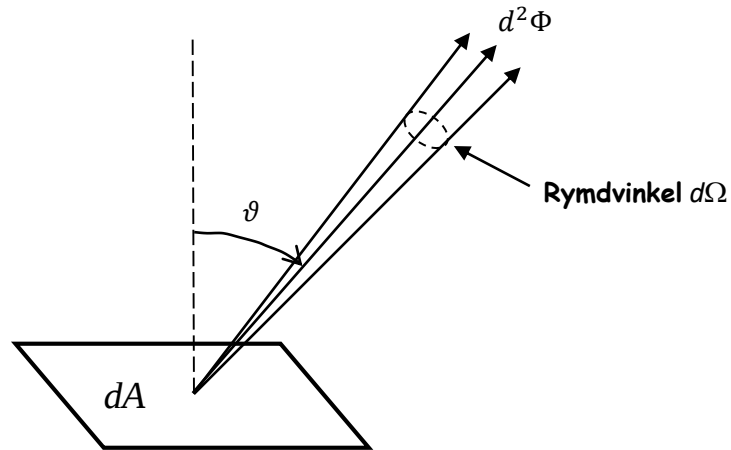


Fig. 9.6. Illustration av ljusflöde som sänds ut inom en infinitesimal rymdvinkel, och i en riktning ϑ mot normalen.

9.10. Belysningen i kamerans bildplan

Nu ska vi härleda ett uttryck på hur hög belysningen i sensorplanet blir om vi avbildar ett motiv med en viss given luminans L . (Vi antar att objektivet är korrekt fokuserat, så att det optiska bildplanet sammanfaller med sensorplanet.) Eftersom luminansen oftast är olika för olika delar av motivet, ska vi titta på ett litet ytelement vars luminans antas konstant över arean dA . Vi kan inte heller anta att motivet är plant, och därför ska vi anta att ytelementet bildar en vinkel α med ett plan vinkelrätt mot optiska axeln, Fig. 9.7. Vidare antar vi att det avbildade ytelementets centrumpunkt ligger på ett vinkelavstånd φ från optiska axeln. Vi antar att det avbildade ytelementet är en perfekt matt yta, dvs. en Lambertreflektor. Avbildningen antas ske med en tunn lins som har diametern D och brännvidden f . I verkligheten kommer D att utgöras av bländarens diameter, snarare än linsdiametern, utom vid full bländaröppning. Dessutom behöver man för verkliga objektiv göra en strålgång med två huvudplan såsom visades i kapitlet om optisk avbildning. Detta påverkar dock inte slutresultatet, och därför väljer vi att betrakta fallet med tunn lins.

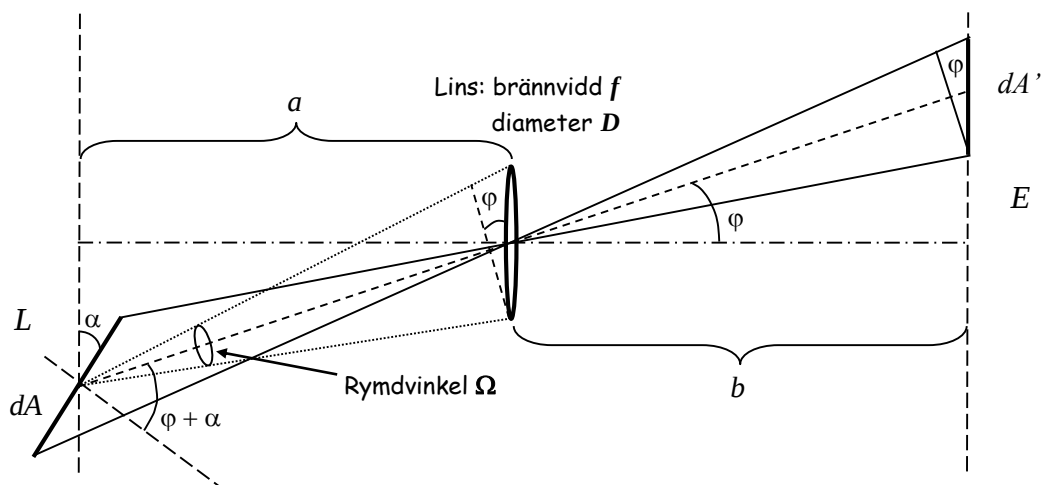


Fig. 9.7. Schematisk figur för beräkning av belysningen i bildplanet.

Enligt tidigare har vi sambandet $L = \frac{d^2\Phi}{dAd\Omega \cos\vartheta}$. Om vi antar att $D \ll a$ (vi ska titta på generella fallet senare) så kan vi anta att Ω är en mycket liten rymdvinkel. Vi kan då direkt stoppa in Ω istället för $d\Omega$ i formeln för L ovan. Vi får då att ljusflödet in genom linsen ges av

$$d\Phi = L \cdot dA \cdot \Omega \cdot \cos(\varphi + \alpha) = L \cdot dA \cdot \frac{\pi D^2 \cos\varphi}{4 \left(\frac{a}{\cos\varphi} \right)^2} \cdot \cos(\varphi + \alpha) \quad (9.7)$$

Belysningen, E , i bildplanet blir

$$E = \frac{d\Phi}{dA'} = \left\{ \text{ur fig. erhålls } \frac{dA' \cos\varphi}{b^2} = \frac{dA \cos(\varphi + \alpha)}{a^2} \right\} = \frac{L \pi \cos^4 \varphi}{4 \left(\frac{b}{D} \right)^2} \quad (9.8)$$

Vi ser här att det är fyra faktorer som påverkar belysningen på sensorn, nämligen motivluminans, L , linsdiameter, D , bildavstånd, b och vinkelavståndet till optiska axeln, φ . Om vi tittar på det viktiga specialfallet $\varphi = 0$ (dvs. i bildcentrum) och $a \gg f$ (fotografering på "stora" avstånd, vilket medför att $b \approx f$), så får vi

$$E \approx \frac{L\pi}{4 \left(\frac{f}{D} \right)^2} \cdot \text{Kvoten } \frac{f}{D}, \text{ dvs. linsens brännvidd dividerad med bländaröppningens}$$

diameter, är lika med **bländartalet**, F . Vi får då

$$E \approx \frac{L\pi}{4F^2} \quad (9.9)$$

dvs. **belysningen i bildcentrum kommer vid fotografering av avlägsna motiv bara att bero av motivluminansen och bländartalet**. Vi har direkt proportionalitet mot luminansen, och omvänd proportionalitet mot kvadraten på bländartalet.

Bländarens diameter, och därmed bländartalet F , varieras med hjälp av tunna metall-lameller såsom beskrevs i kap. 4. De hela bländartalen ges av följande serie ... 2, 2.8, 4, 5.6, 8, 11, 16 22 ... Mellan intilliggande tal skiljer det en faktor $\sqrt{2}$ (avrundat till två siffrors noggrannhet). Detta innebär att belysningen i sensorplanet ändras med en faktor 2 eller $\frac{1}{2}$ när vi går ett bländarsteg, dvs. ett steg nedåt eller uppåt längs skalan. Låga bländartal ger hög belysning, medan höga tal ger låg belysning. Lägsta respektive högsta bländartal som kan erhållas varierar mellan olika objektiv.

Som sagt gäller att bländartalet ges av formeln $F = \frac{f}{D}$. För ett verkligt fotografiskt objektiv gäller att D -värdet som insätts i formeln inte ska vara lika med bländarens verkliga mekaniska diameter. Istället ska man sätta in storleken på den så kallade ingångspupillen, Fig. 9.8.

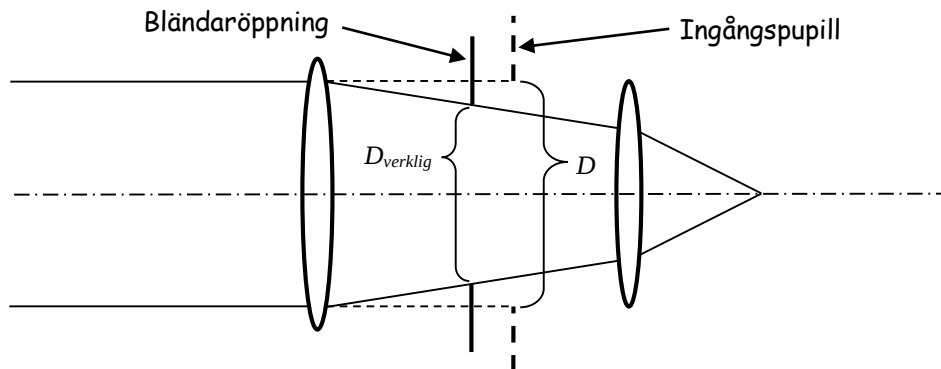


Fig. 9.8. Illustration av begreppet ingångspupill.

Anledningen till att $D \neq D_{\text{verklig}}$ är att ljuset bryts av ett eller flera linselement innan det når bländaröppningen. Den ljussuppsamlande förmågan hos objektivet svarar därför mot en öppning av D och inte mot den verkliga öppningen D_{verklig} . Detta tar naturligtvis objektivkonstruktörerna hänsyn till när de bestämmer objektivets bländartal.

9.11. Objektivets ljusstyrka

Lägsta bländartalet, som kallas **objektivets ljusstyrka**, bestäms av linsdiametern. Det lönar sig ju inte att göra D i Fig. 9.8 större än linsdiametern, vi får inte in mera ljus för det. Om största användbara ingångspupill är D_{max} , får vi att ljusstyrkan är $\frac{f}{D_{\text{max}}}$. Observera att låga tal på ljusstyrkan betyder att vi kan få en hög belysning i bildplanet. Därför säger man att objektiv är ljusstarka om de har ett lågt värde på $\frac{f}{D_{\text{max}}}$. Dumt kan det tyckas, men detta är gammal inarbetad standard. Vanliga

ljusstyrkor för fotografiska objektiv ligger i intervallet 1.4 – 8. De högre talen (dvs. låg ljusstyrka) brukar gälla för teleobjektiv med långa brännvidder. Bra ljusstyrka är värdefullt om man vill fotografera vid dåliga ljusbetingelser, t.ex. vid gryning och skymning. Fotografering vid låga bländartal innebär emellertid nackdelar i form av litet skärpedjup och allmänt sämre bildskärpa (se senare kapitel). När man ändrar bländartalet så pratar man ofta om att man ”bländar upp” eller ”bländar ner”. Med blända upp menar man att bländaröppningens storlek ökar, vilket innebär att bländartalet blir lägre. Motsatsen, att blända ner, innebär att bländaröppningen görs mindre, vilket medför högre bländartal.

9.12. Belysningen i bildplanet vid närbildsfoto

Om inte motivavståndet är mycket större än brännvidden, så kan vi inte längre anta att $b \approx f$. Belysningen i bildplanets centrum ges då av (se ekv. 9.8)

$$\begin{aligned}
 E &= \frac{L\pi}{4\left(\frac{b}{D}\right)^2} = \frac{L\pi}{4\left(\frac{b}{f}\right)^2 F^2} = \left\{ \text{Ekv. 5.1} \Rightarrow \frac{b}{f} = \frac{a}{a-f} \right\} = \\
 &= \frac{L\pi}{4F^2} \cdot \left(\frac{a-f}{a}\right)^2 = E_0 \cdot \left(1 - \frac{f}{a}\right)^2
 \end{aligned} \tag{9.10}$$

där E_0 är belysningen vid ∞ motivavstånd. Ur formeln ser vi att belysningen blir lägre ju kortare motivavståndet är. Vid $a = 20f$ så är $E = 0.90 \cdot E_0$, och vid $a = 2f$ är $E = 0.25 \cdot E_0$. För praktisk fotografering kan man säga att belysningens beroende av motivavståndet kan försummas så länge detta är minst en storleksordning större än brännvidden. Däremot blir effekten stor när motivavståndet närmar sig brännvidden. Formeln för E kan också uttryckas i termer av avbildningsskalan $M = \frac{b}{a}$ (Ekv. 5.3). Detta ger

$$E = \frac{L\pi}{4\left(\frac{b}{f}\right)^2 F^2} = \left\{ \text{Ekv. 5.1 + 5.3} \Rightarrow \frac{b}{f} = 1 + M \right\} = \frac{L\pi}{4F^2(1+M)^2} = \frac{E_0}{(1+M)^2} \tag{9.11}$$

Ur denna formel ser vi att närbildsfoto i skala 1:1 ger en sensorbelysning som bara är en fjärdedel av vad den blir om samma motiv fotograferas på stort avstånd.

9.13. Ljusavfall mot bildkanten

Ur ekv. 9.8 framgår att belysningen i bildplanet avtar med ökande infallsvinkel φ . Vinkelberoendet ges av faktorn $\cos^4 \varphi$, och därför brukar man kalla denna effekt för **$\cos^4$ – avfallet** (ibland kallas effekten ”naturliga ljusavfallet”). $\cos^4$ -avfallet finns illustrerat i Fig. 9.9. Effekten märks mest när man använder vidvinkelobjektiv. Exempelvis så medför $\cos^4$ – avfallet att ljusnivån ute i bildhörnet bara är 39 % av centrumvärdet när man använder ett 28 mm objektiv på en fullformatskamera (se Fig. 6.1). För ett normalobjektiv med 50 mm brännvidd är motsvarande värde 71 %, och för ett 135 mm tele 95 %. I Fig. 9.10 visas en extrem vidvinkelbild som tydligt uppvisar problemet med ljusavfall mot bildkanten.

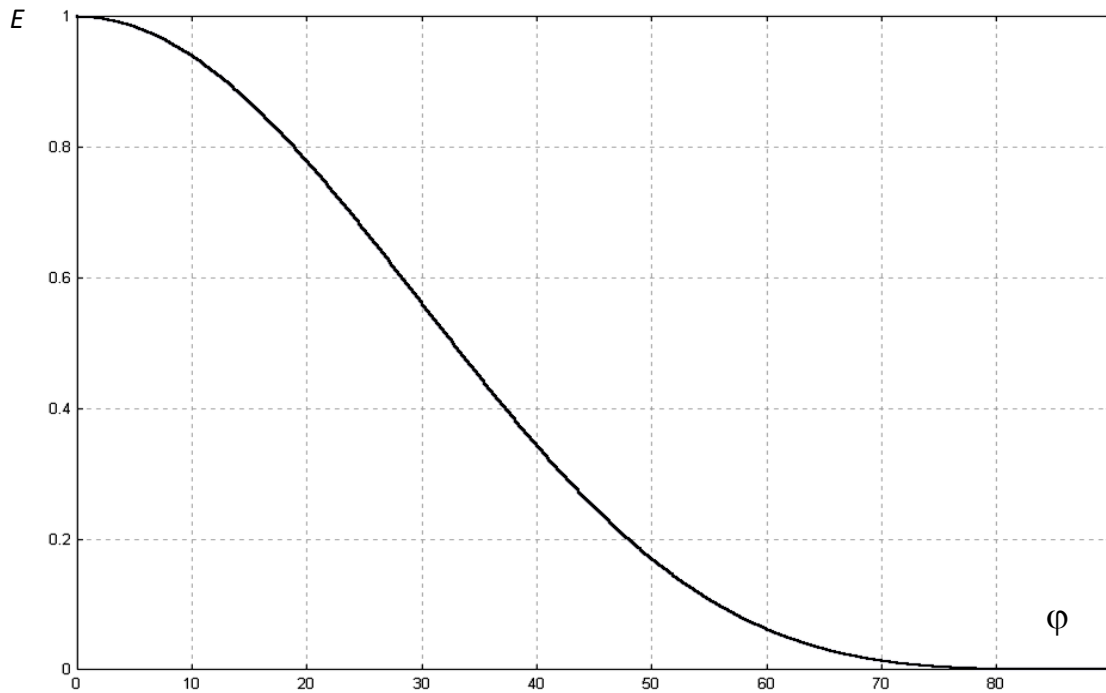


Fig. 9.9. Illustration av hur belysningen i sensorplanet avtar som funktion av bildvinkeln ϕ på grund av det så kallade "naturliga ljusavfallet". Belysningen är proportionell mot $\cos^4\phi$.



Fig. 9.10

Extrem vidvinkelbild tagen med hålkamera (se kap. 2). Sensorytan (film) var ca. $50 \times 70 \text{ mm}^2$. Med en "brännvidd" av 40 mm, ger detta en bildvinkel av 46° ute i hörnen. På grund av naturliga ljusavfallet blir därför belysning i hörnen bara lite drygt 20 % av vad den är i bildcentrum.

$\cos^4$ – avfallet kan tyckas vara ett allvarligt problem, men i praktiken kan man tolerera en hel del ljusavfall utan störs av det. Men det är uppenbart att för extrema vidvinklar blir $\cos^4$ – avfallet ett problem. Det finns ju till och med så kallade fish-eye objektiv där vinkeln φ kan bli praktiskt taget 90° . Detta skulle innebära att belysningen blir noll ute i bildkanten! Tittar man på en fish-eye bild så ser man emellertid knappast något ljusavfall alls ute i periferin. Hur kan det komma sig? Jo, extrema vidvinklar, varav fish-eye typen är ett paradexempel, uppvisar distorsion. Detta innebär att avbildningsskalan minskar när man går från centrum ut mot periferin. Detta innebär att det låga ljusflödet som passerar objektivet vid stora infallsvinklar kommer att koncentreras till en allt mindre yta ju större vinkeln är. Sålunda kan belysningen, dvs. ljusflödet per ytenhet, hållas praktiskt taget konstant. Förutom distorsion finns några andra knep man kan ta till för att minska ljusavfallet. Man kan t.ex. när man konstruerar ettektiv ta med belysningsjämnheten som ett av de avbildningsfel man strävar efter att minimera genom att kombinera flera olika linselement. En annan variant är att framför sensorn ha ett filter som absorberar mest i mitten och mindre utåt kanterna. Man kan också kompensera för ljusavfallet vid efterbearbetningen av bilderna. Detta medför dock en viss försämring av bildkvalitén, t.ex. en ökad brusnivå, speciellt om kompensationen är kraftig.

Om man mäter upp ljusavfallets storlek på verkliga objektiv så finner man, speciellt vid stora bländaröppningar, att ljusavfallet utåt bildkanterna ofta blir större än vad som ges av faktorn $\cos^4 \varphi$, se exempel i Fig. 9.11. Detta extra ljusavfall beror på en effekt som kallas **vinjettering**. Vinjettering beror på att hela bländaröppningen inte kommer att fyllas med ljus vid snett infall och stora bländaröppningar, se Fig. 9.12. Genom att använda mindre bländaröppning, dvs högre bländartal, minskas vinjetteringen.

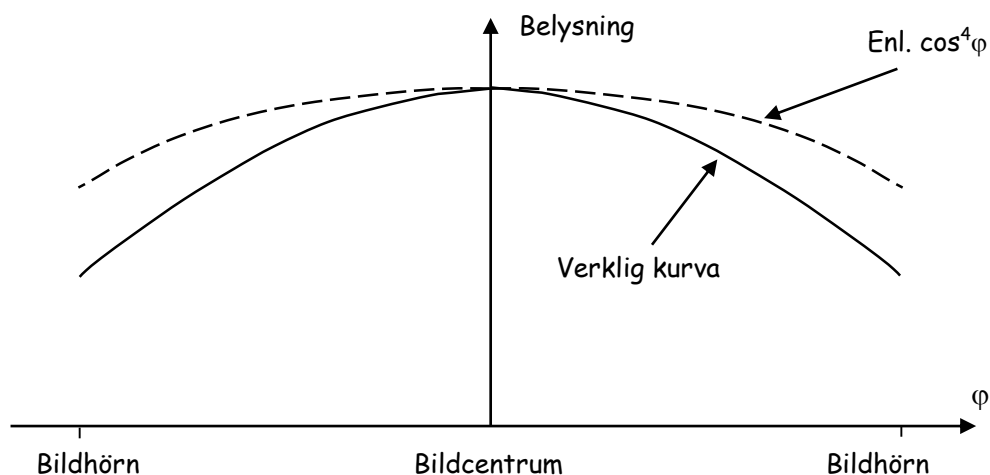


Fig. 9.11. Ljusavfallet i ettektiv gör att belysningen blir lägre i bildkanten. Oftast avtar belysningen snabbare än vad som ges av funktionen $\cos^4 \varphi$. Detta beror på så kallad vinjettering, se Fig. 9.12.

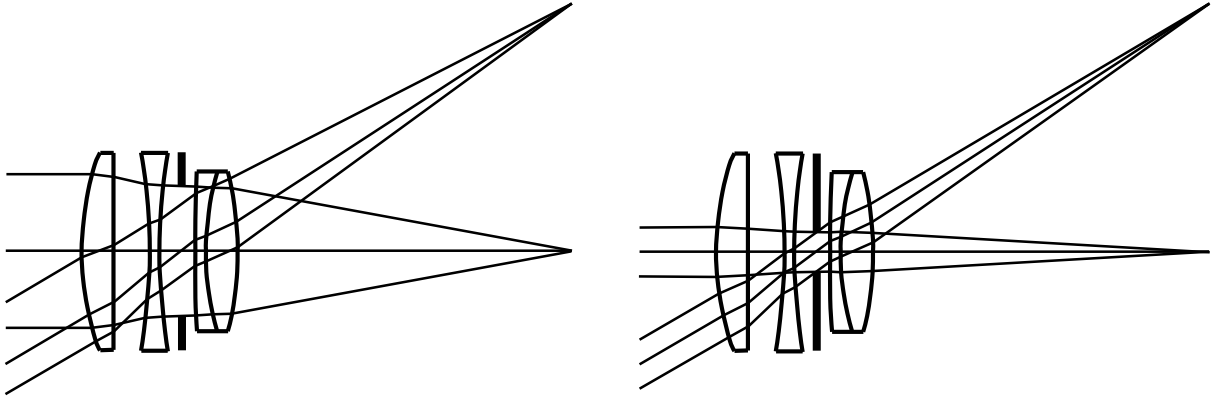


Fig. 9.12. Vid bländartal 4 (vänstra figuren) och snett ljusinfall kommer inte hela bländaröppningen att fyllas med ljus. Detta kallas vinjettering, och orsakas av linsernas begränsade diameter. Vid bländartal 11 (högra figuren) uppstår ingen vinjettering, vilket gör att ljusavfallet mot bildkanten blir lägre än vid bländartal 4.

9.14. Belysningen i bildplanet vid extremt låga bländartal

Innan vi lämnar detta kapitel om fotometri, ska vi titta lite på de förhållanden som råder vid mycket låga bländartal. Enligt ekv. 9.9 gäller att belysningen E i bildplanet är omvänt proportionell mot bländartalet, $\frac{f}{D}$, i kvadrat. Rent teoretiskt finns inget

som hindrar att $\frac{f}{D}$ blir hur litet som helst (i praktiken finns naturligtvis en sådan

begränsning). Teoretiskt finns det därför inget som hindrar att E blir hur hög som helst, kan tyckas. Formeln bygger emellertid på antagandet att rymdvinkeln Ω är liten, vilket inte stämmer vid låga bländartal. Dessutom gäller vid låga bländartal en något annorlunda formel för avbildningsskala än den som antagits vid härledningen av ekv. 9.9. För den intresserade följer här en beräkning av ett generellt uttryck för belysningen i bildcentrum som inte förutsätter låga bländartal. Resultatet visar att belysningen i bildplanet aldrig ens teoretiskt kan bli högre än π gånger motivluminansen.

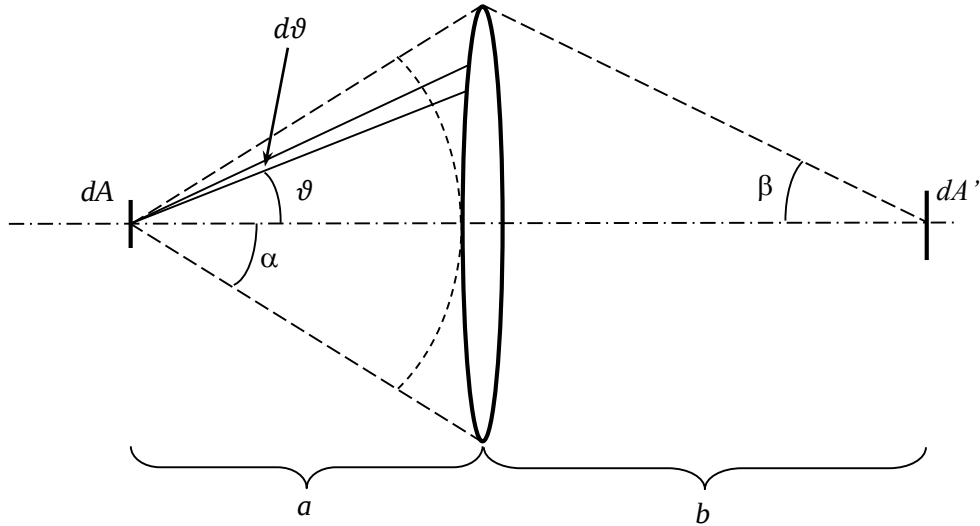


Fig. 9.13. Schematisk figur för härledning av belysningsformel som är giltig även vid mycket låga bländartal.

Vi kan nu inte längre anta att allt ljus som lämnar ytan dA och färdas mot linsöppningen bildar samma vinkel ϑ mot optiska axeln. Därför får vi ta bidraget till ljusflödet som kommer från det infinitesimala vinkelintervallet $d\vartheta$, se figur, och integrera från $\vartheta = 0$ till $\vartheta = \alpha$. Bidraget från intervallet $d\vartheta$ ges av

$$\begin{aligned} d^2\Phi &= \{\text{enl. def. på } L \text{ för Lambertreflektor}\} = L \cdot dA \cdot d\Omega \cdot \cos\vartheta = \\ &= L \cdot dA \cdot \frac{a \cdot d\vartheta \cdot 2\pi \cdot a \cdot \sin\vartheta}{a^2} \cdot \cos\vartheta = L \cdot dA \cdot 2\pi \sin\vartheta \cos\vartheta d\vartheta \end{aligned}$$

Detta integrerat över vinkel ϑ ger

$$d\Phi = L \cdot dA \cdot \pi \int_{\vartheta=0}^{\vartheta=\alpha} 2 \sin\vartheta \cos\vartheta d\vartheta = L \cdot dA \cdot \pi \cdot \sin^2\alpha. \text{ Avbildningsskalan styrs av det}$$

så kallade sinusvillkoret, se läroböcker i optik, vilket ger att $dA \cdot \sin^2\alpha = dA' \cdot \sin^2\beta$. Genom att kombinera de bägge ekvationerna ovan får vi

$$E = \frac{d\Phi}{dA'} = L \cdot \pi \cdot \sin^2\beta \quad (9.12)$$

Eftersom $\sin\beta \leq 1$ så gäller att $E_{\max} = \pi L$. $\sin\beta$ kallas objektivets numeriska apertur. Inom mikroskopi används denna term ofta. Ovanstående uttryck för E övergår för små vinklar (dvs. då $\sin\alpha \approx \tan\alpha$) till vårt tidigare härledda

$$E = \frac{L\pi}{4\left(\frac{f}{D}\right)^2}.$$

9.15. Ströljus

Populärt uttryckt kan man säga att ströljus är en "ljusdimma" som mer eller mindre jämnt lägger sig över bildplanet. Fenomenet är mycket komplext och därmed svårt att beskriva (eller i varje fall kvantifiera) i form av enkla modeller och formler. En vanlig orsak till ströljus är reflexer i objektivets linselement. Trots att linserna är belagda med antireflexskikt, kan man inte undvika att en liten del av ljuset reflekteras istället för att transmittas vid varje glas/luft övergång. För ett objektiv med många linselement är inte detta försumbart. Det reflekterade ljuset kan studsas och länkas av i riktning upprepade gånger, och kan så småningom hamna någonstans i bildplanet. Speciellt allvarligt blir detta om man riktar kameran i ungefär samma riktning som solen. Även om vinkeln mot solen är så stor att solskivan inte avbildas på sensorn, så kan snett infallande strålar genom första linselementet studsas via objektivets innervägg och slutligen hamna på sensorn. (I fotografins barndom då ingen antireflexbehandling fanns var regeln enkel: Fotografera alltid med solen i ryggen!) Motivet inverkar också på ströljusnivån. Om det innehåller stora ljusa partier och bara små mörka partier, kan man förvänta sig mera ströljus än om de mörka partierna överväger. Kamerahusets och objektivets innerväggar bör vara mattsvarta för att minska ströljusnivån. För att kvantifiera ströljusnivån används ibland den så kallade ströljusfaktorn, S , som definieras genom formeln

$$S = \frac{\text{Motivets kontrastomfång} \left(= \frac{L_{\max}}{L_{\min}} \right)}{\text{Belysningsomfånget i bildplanet} \left(= \frac{E_{\max}}{E_{\min}} \right)} \quad (9.13)$$

Utan ströljus får vi $S = 1$, och med ströljus $S > 1$ (ju mer ströljus, desto större S). I en verklig situation kan S typiskt vara 2 – 3. Ströljusets inverkan är speciellt märkbar i mörka motivpartier, vilket framgår av Fig. 9.14. Visserligen finns ströljuset också med i ljusa motivpartier, men dess inverkan är där procentuellt sett så mycket mindre att den knappt syns i bilden. Men vid svåra atmosfäriska förhållanden (dis, dimma) kan man även för höga värden på motivluminansen få en märkbar påverkan av ströljus.

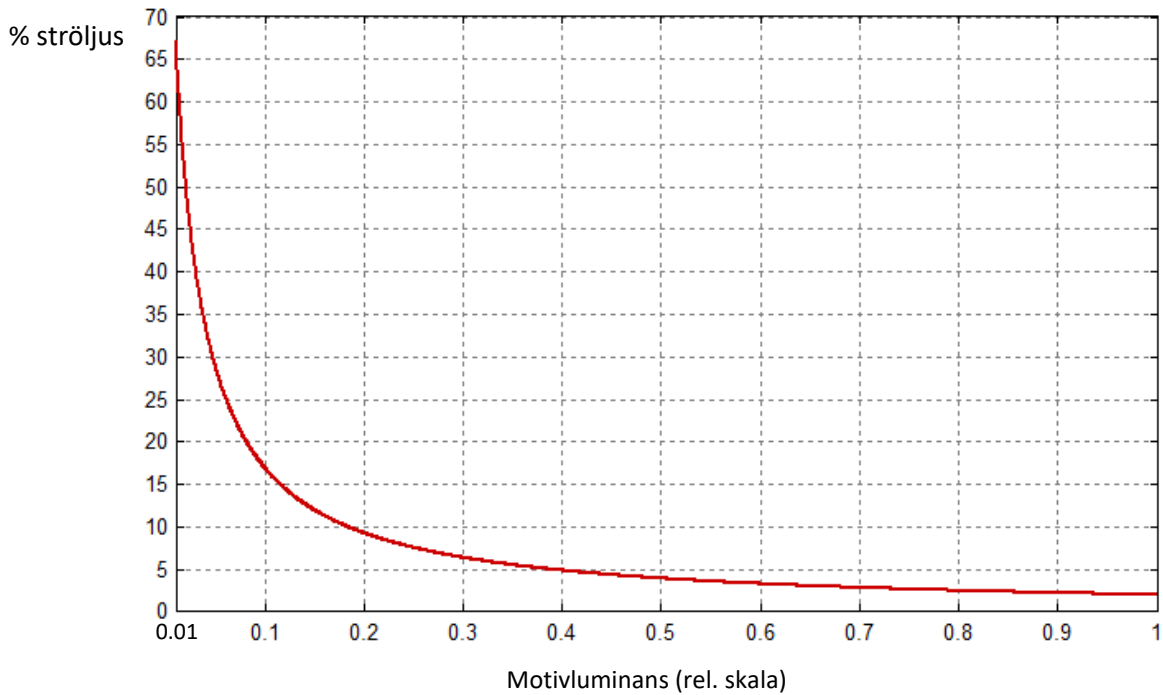


Fig. 9.14. Exempel på hur ströljuskivån i en bild kan variera som funktion av motivluminansen. I det valda exemplet antas motivluminansen variera mellan 0.01 i mörka partier och 1.0 i ljusa partier. Motivets kontrastomfång är alltså 100. Ströljuskivån varierar mellan 2% i ljusa partier och 67% i mörka partier. Belysningsomfånget i bildplanet blir då 34, vilket ger en ströljuskfaktor av 3.0 (ekv. 9.13).

10. Skärpedjup

10.1. Inledning

Skärpedjup är ett mycket viktigt begrepp inom fotografin. Figur 10.1 visar två bilder av samma motiv, men med olika skärpedjup. Ofta vill man ha ett så stort skärpedjup som möjligt så att motivets alla delar avbildas skarpt. Det finns dock situationer där man vill begränsa skärpedjupet, t.ex. för att få bort en störande bakgrund vid porträttfotografi. I detta kapitel ska vi titta på vilka faktorer som bestämmer skärpedjupet, och hur man som fotograf kan påverka detta.

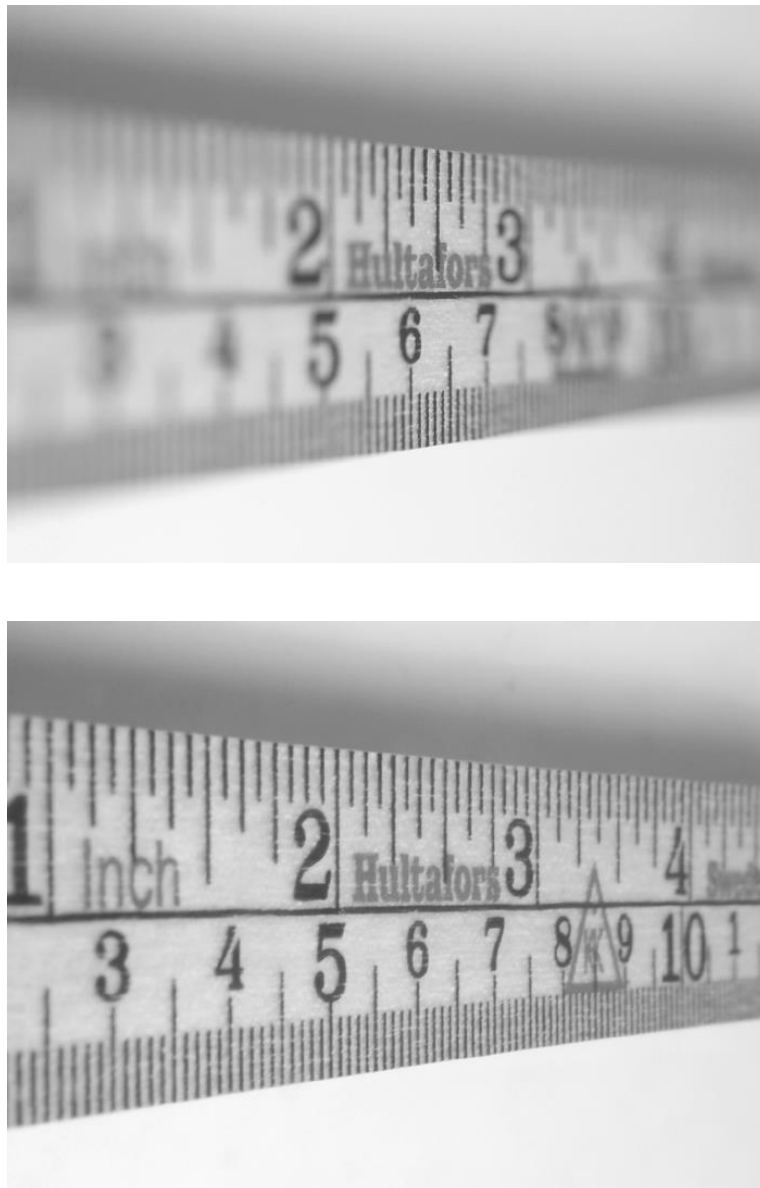


Fig. 10.1. Samma motiv fotograferat under identiska förhållanden, utom att bländartalet var 3.5 i övre bilden och 22 i den undre.

Enligt linsformeln, ekv. 5.1, kan man för ett visst bildavstånd och en viss brännvidd bara avbilda föremål skarpt på ett enda avstånd. I praktiken kommer dock föremål på något större och något mindre avstånd också att avbildas med acceptabel skärpa, Fig. 10.2. Avståndet mellan det avlägsnaste och närmaste föremålet som avbildas tillräckligt skarpt kallas skärpedjupet.

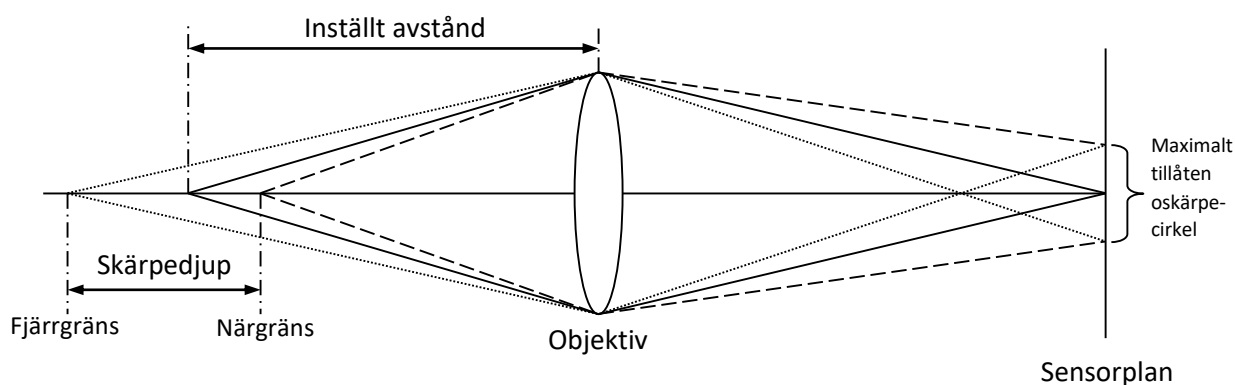


Fig. 10.2. Strålgång som illustrerar begreppet skärpedjup.

Den maximalt tillåtna oskärpecirkeln beror på en mängd faktorer såsom bildens betraktningsavstånd och förstöringsgrad, sensorns egenskaper, hur kritisk betraktaren är etc. För "normalt" bruk kan en acceptabel oskärpecirkel för en fullformatsensor (24 mm x 36 mm) anses vara cirka 30 μm . Denna siffra bygger på att om bilden förstoras till 10 cm x 15 cm (ca. en faktor 4) så kan man vid betraktning på 25 cm avstånd knappast uppfatta någon oskärpa.

Oskärpecirkelns diameter minskar när bländaröppningens diameter görs mindre (dvs när bländartalet ökar). I Fig. 10.3 illustreras denna effekt.

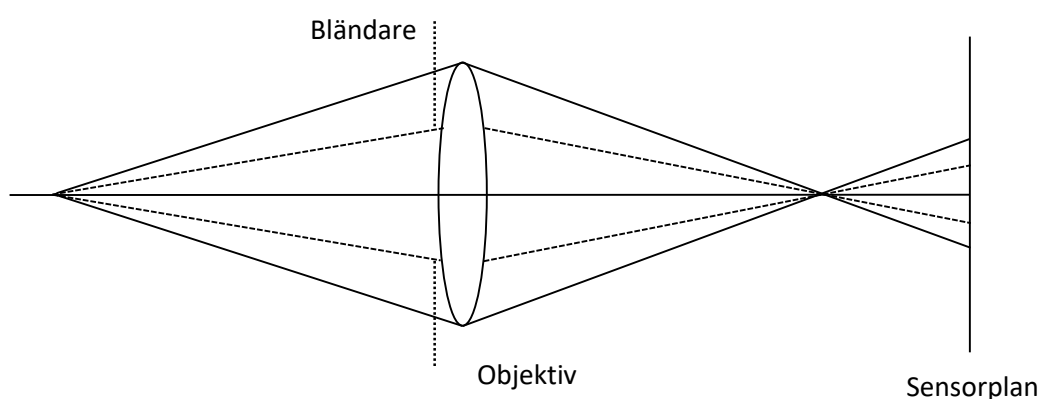


Fig. 10.3. En defokuserad punktkälla kommer att avbildas som en oskärpecirkel i sensorplanet (heldragna linjer). Om en bländare införs i strålgången (streckade linjer), så kommer oskärpefläcken att minska i storlek. Ju mindre bländaröppning, desto mindre oskärpefläck. Dessutom kommer oskärpefläcken att vara en geometrisk projektion av bländaröppningen. Om bländaröppningen är t.ex. 6-kantig, så kommer även oskärpefläcken att ha samma form.

Att en mindre bländaröppning ger mindre oskärpefläck innebär att skärpedjupet ökar vid högre bländartal. Som hjälp att bestämma skärpedjupet finns på vissa objektiv en skärpedjupsskala, Fig. 10.4.

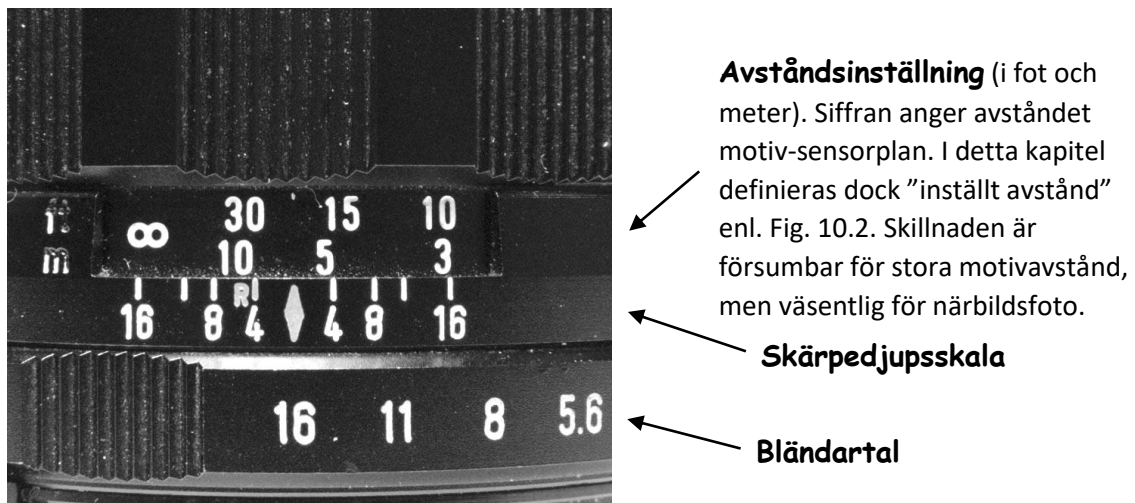


Fig. 10.4. Skärpedjupsskala på objektiv. I detta fall är bländaren inställd på 16 och avståndet på ca. 6 meter. På skärpedjupsskalan kan man då avläsa att skärpedjupet sträcker sig från 3 meter till oändligheten.

10.2. Hyperfokalinställning

Ofta vill man ha ett skärpedjup som sträcker sig från oändligheten till så nära kameran som möjligt. Detta kallas hyperfokalinställning, och innebär att fjärrgränsen placeras i oändligheten. Fig. 10.4 illustrerar hyperfokalinställning vid bländartal 16.

Hyperfokalinställningen kan även utan alltför stort besvär räknas fram för godtycklig brännvidd, bländartal och tillåten oskärpecirkel. Detta kan vara användbart om man t.ex. ställer extra höga krav på skärpa, exempelvis när man gör kraftiga förstoringar av bilden, eller om man har en digitalkamera med liten sensor. I sådana fall kan en oskärpecirkel på 30 μm i sensorplanet vara oacceptabel.

Exempel: Beräkna hyperfokalinställningen för ett 50 mm objektiv vid bländartal 8 och max. oskärpecirkel 20 μm . Hur nära kameran kommer närgränsen att ligga?

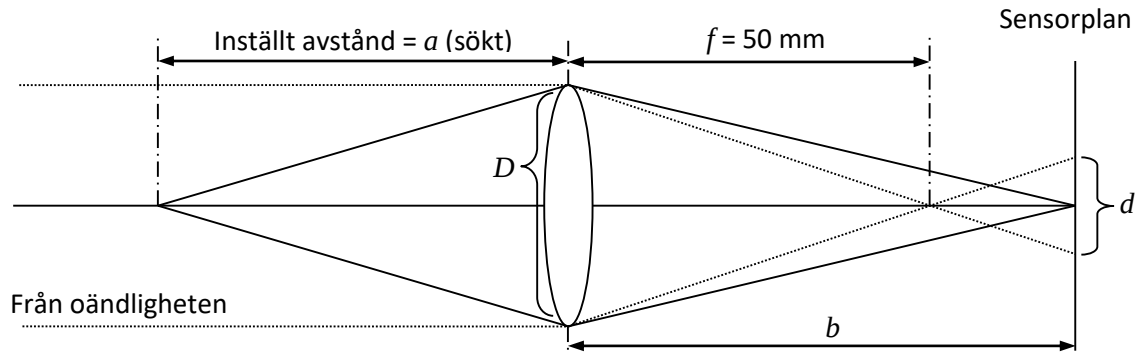


Fig. 10.5. Hyperfokalinställning.

$$\begin{cases} \frac{1}{a} + \frac{1}{b} = \frac{1}{f} & (10.1) \\ \frac{D}{f} = \frac{d}{b-f} & (10.2) \end{cases}$$

$$D = \frac{f}{\text{bländartal}} = \frac{50 \times 10^{-3}}{8} = 6.25 \times 10^{-3} \text{ m}, \quad d = 20 \times 10^{-6} \text{ m}.$$

$$(10.2) \Rightarrow b = f \left(1 + \frac{d}{D} \right) \text{ insättes i (10.1)}$$

$$\Rightarrow a = f \left(1 + \frac{D}{d} \right) = 50 \times 10^{-3} \left(1 + \frac{6.25 \times 10^{-3}}{20 \times 10^{-6}} \right) = 15.7 \text{ m}.$$

Genom att ställa in avståndet på 15.7 m kommer alltså fjärrgränsen för skärpedjupet att ligga i oändligheten. Var hamnar närgränsen (a' i Fig. 10.6)?

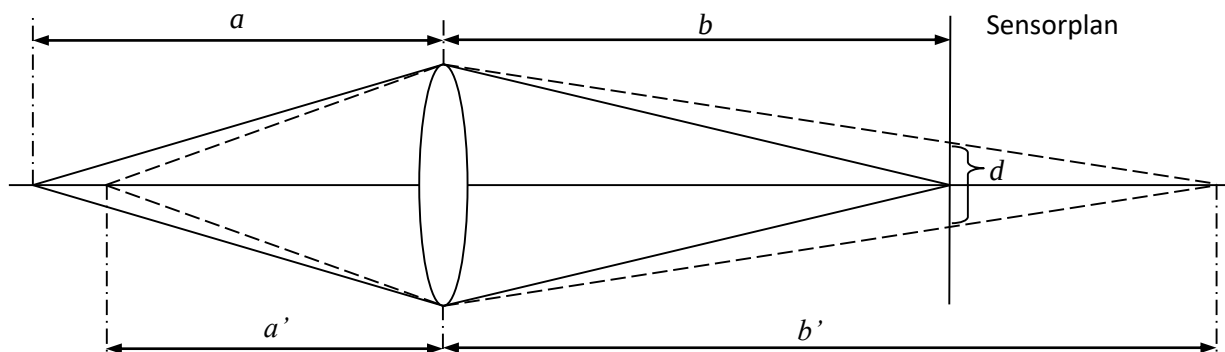


Fig. 10.6. Beräkning av närgräns vid hyperfokalinställning.

$$\left. \begin{array}{l} \frac{1}{a'} + \frac{1}{b'} = \frac{1}{f} \quad (10.3) \\ \frac{D}{b'} = \frac{d}{b'-b} \quad (10.4) \end{array} \right\} \begin{array}{l} (10.4) + (10.2) \Rightarrow b' = \frac{f\left(1 + \frac{d}{D}\right)}{1 - \frac{d}{D}} \text{ insättes i (10.3)} \\ \Rightarrow a' = \frac{f\left(1 + \frac{D}{d}\right)}{2} = \frac{a}{2} \quad (7.8 \text{ m}) \end{array} \quad (10.5)$$

Vi ser alltså att vid hyperfokalinställning gäller alltid att närgränsen är lika med halva inställda avståndet. Med användande av ovanstående generella formler är det lätt att bestämma hyperfokalinställningen för godtyckliga brännvidder, bländartal samt krav på oskärpecirklar.

Låt oss dröja ett ögonblick vid ekv. 10.5. Den ger oss närgränsen vid hyperfokalinställning, vilket innebär att allt bortom a' avbildas (tillräckligt) skarpt. Ett lågt värde på a' är alltså ekvivalent med stort skärpedjup. Vi har att

$$a' = \frac{f}{2} \left(1 + \frac{D}{d}\right) = \left\{ \frac{D}{d} \gg 1 \text{ nästan alltid} \right\} \approx \frac{f D}{2d} = \left\{ \text{bländartal}, F = \frac{f}{D} \right\} = \frac{f^2}{2dF} \quad (10.6)$$

a' beror på brännvidd, tillåten oskärpefläck och bländartal. Det kvadratiska beroendet på brännvidden innebär att skärpedjupet ökar kraftigt när vi minskar brännvidden. Likaså ser vi att höga bländartal och stor tillåten oskärpefläck ger stort skärpedjup.

En intressant frågeställning är hur skärpedjupet påverkas av sensors storlek. Antag att vi har, säg, en fullformatskamera (24 mm x 36 mm sensor) och en kompaktkamera med 1/2.3" sensor (4.55 mm x 6.17 mm), bägge utrustade med normaloptik så att samma bildvinkel erhålls i bägge fallen. Om man fotograferar med samma bländartal med dessa kameror, vilket ger samma belysning i sensorplanet, kommer då skärpedjupet att bli detsamma? Vid besvarandet av denna fråga måste vi naturligtvis ta hänsyn till att kompaktkamerabilderna kommer att behöva förstoras mer, vilket medför att vi måste kräva en mindre oskärpefläck. Likaså måste vi ta hänsyn till att normaloptikens brännvidd ökar med sensorformatet. Man brukar lite grovt räkna med att en oskärpefläck med 125 μm diameter är acceptabel när man betraktar slutbilden på 25 cm avstånd. Låt oss använda detta värde, och att slutbilden är i storlek 10 x 15 cm. Eftersom sensorer har lite olika höjd/breddförhållande, så antar vi att bilden förstoras upp så mycket att sensors långsida kommer att uppta 15 cm (ev. bildrest skärs helt enkelt bort). Under dessa förutsättningar, låt oss bestämma närgränsen vid hyperfokalinställning enligt ekv. 10.6 för några olika sensorformat. Resultatet finns sammanställt i tabellen på nästa sida. Låga tal svarar mot stort skärpedjup.

**Närgräns i meter för skärpedjupet vid hyperfokalinställning
med normaloptik (f = sensordiagonalen) och olika bländartal.**

<u>Sensorformat</u>	<u>Brännvidd (mm)</u>	<u>Bländartal</u>								
		2	2.8	4	5.6	8	11	16	22	32
		<u>Närgräns (m)</u>								
1/2.3" (4.55 mm x 6.17 mm)	7.7	2.9	2.1	1.5	X	X	X	X	X	X
APS-C (25.1 mm x 16.7 mm)	30	11	7.7	5.4	3.8	2.7	2.0	1.3	X	X
Fullformat (24 mm x 36 mm)	43	15	11	7.7	5.5	3.9	2.8	1.9	1.4	X
Mellanformat (55 mm x 55 mm)	78	33	24	17	12	8.3	6.0	4.1	3.0	2.1

De rutor som är markerade med X innebär att bländartalet är för högt för det använda sensorformatet; diffraktionen kommer att ge en oacceptabelt stor oskärpefläck på en 10 cm x 15 cm bild. Kortaste närgränsen blir densamma för alla formaten, nämligen ca. 1.4 m (avvikelserna från detta värde i tabellen beror på att bara hela bländartal finns med). Denna kortaste närgräns uppnås dock vid olika bländartal för de olika formaten – ju större format desto högre bländartal. Belysningen i sensorplanet är omvänt proportionell mot bländartalet i kvadrat (men oberoende av brännvidden). Det innebär att för att få samma skärpedjup med ett stort sensorformat som med ett litet, så måste man använda ett högre bländartal vilket kräver längre exponeringstid, med åtföljande risk för rörelseoskärpa. Ibland vill man dock ha ett litet skärpedjup för att framhäva en viss del av motivet. Detta kan man alltså lättast göra med en kamera med stor sensor (se även kap. 10.4).

10.3. Allmänna skärpedjupsfallet

Vi har nu behandlat fallet hyperfokalinställning, dvs. när skärpedjupets fjärrgräns ligger i oändligheten. I många fotografiska situationer är detta inte nödvändigt, eller ens önskvärt. Det betyder att man kan behöva titta på ett mer generellt fall vad gäller skärpedjup. Med samma typ av geometriska betraktelser som gjordes för att få fram formeln för hyperfokalinställning, kan man betrakta ett mer generellt fall. Om motivavstånd, brännvidd, bländartal och tillåten oskärpecirkel är givna, kan man beräkna skärpedjupets fjärr- och närgränser. Om vi inför beteckningen $a_{\max}$ för fjärrgränsen och $a_{\min}$ för närgränsen får vi följande formler

$$a_{\max} = \frac{a_0}{1 - \frac{Fd}{f} \left(\frac{a_0}{f} - 1 \right)} \quad (10.7)$$

$$a_{\min} = \frac{a_0}{1 + \frac{Fd}{f} \left(\frac{a_0}{f} - 1 \right)} \quad (10.8)$$

där f = brännvidden, F = bländartalet, d = maximalt tillåten oskärpecirkel och a_0 är inställt motivavstånd. Formlerna bygger på geometrisk optik, vilket innebär att

diffractionseffekter ej är medtagna. I det generella fallet gäller, precis som i fallet hyperfokalinställning, att skärpedjupet ökar vid korta brännvidder, höga bländartal och stora oskärpecirklar. När skärpedjupet är litet, t.ex. vid närbildsfoto, ligger när- och fjärrgränserna ungefär symmetriskt kring det inställda motivavståndet. Vid större skärpedjup däremot ligger huvuddelen av skärpedjupet bortom det inställda avståndet. Detta illustreras i Fig. 10.7, vilken visar när- och fjärrgränser som funktion av inställt motivavstånd för det speciella fallet att $f = 50$ mm, $F = 5.6$ och $d = 30$ μm (typiska data för en fullformatskamera).

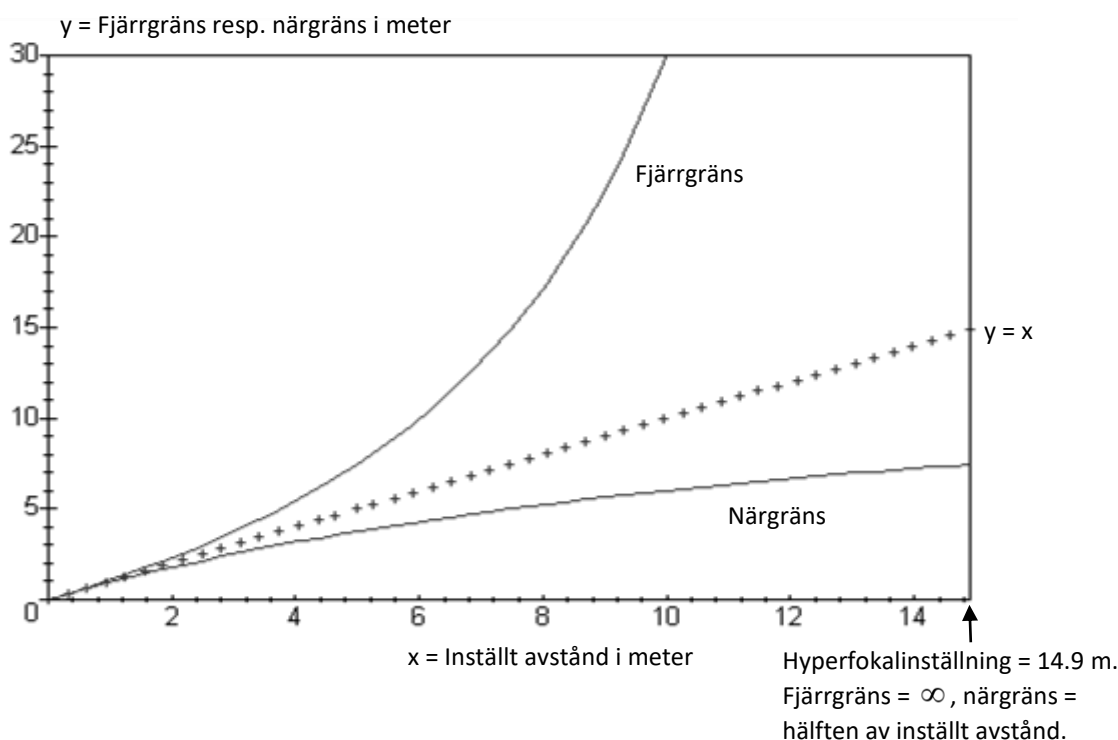


Fig. 10.7. När- och fjärrgränser för skärpedjupet vid $f = 50$ mm och bländartal 5.6.

Som vi nu har sett är ämnet skärpedjup inte speciellt svårt att behandla rent matematiskt. Man kan ganska lätt få fram formler som beskriver t.ex. fjärr- och närgränser. I praktiskt fotograferingsarbete kan det dock vara lite bökigt att jobba med dessa formler. Därför är en skärpedjupsskala på objektivet en mycket värdefull hjälp. Tyvärr saknas en sådan skala ofta på objektiv till moderna kameror. En anledning till detta är att autofokus har blivit standard, och att man därför har slopat såväl avstånds- som skärpedjupsskala.

När man utnyttjar en skärpedjupsskala är det bra att ha i åtanke att den ofta ger en lite för optimistisk bedömning av skärpedjupet. Speciellt om man tänker göra kraftiga förstoringar måste man tillämpa snävare gränser. Ett gott råd man hör från fotografer ibland, är att man visserligen kan använda skärpedjupsskalan på objektiven, men att man bör använda de skärpedjupsmarkeringar som gäller för 1-2 stegs större bländaröppning än den man använder. Detta innebär att om kameran indikerar att skärpedjupet räcker precis vid användande av bländartal 4, så bör man istället använda 5.6 eller 8.

10.4. Angående kort skärpedjup

I många fotografiska situationer eftersträvar man ett kort, snarare än ett långt, skärpedjup. Man vill kanske suddas ut en ful och störande bakgrund när man tar en porträttbild. Eller man vill kanske framhäva någon detalj i motivet som ligger på ett visst avstånd. I sådana fall är det viktigt att man kan erhålla ett tillräckligt kort skärpedjup. En illustration av detta syns i Fig. 10.8.



Fig. 10.8. Porträttbilder tagna under identiska förhållanden, förutom att bländartalen (och exponeringstiderna) är olika. Vänstra bilden togs med bländartal 36, och högra bilden med 5.6. Brännvidden var 70 mm, och kameran var utrustad med en fullformatssensor.

I fotolitteraturen diskuteras ofta vikten av att ha en kamera med stor sensor för att kunna erhålla kort skärpedjup, och att det är svårt eller omöjligt att uppnå detta med kompaktkameror (eller kameramobiler) som har små sensorer. Låt oss titta lite kvantitativt på detta. Totala skärpedjupet ges av $a_{max} - a_{min}$ i formeln 10.7 och 10.8. Detta ger följande formel för skärpedjupet

$$a_{max} - a_{min} = \frac{2a_0 F d f^2 (a_0 - f)}{f^4 - F^2 d^2 (a_0 - f)^2} \quad (10.9)$$

där a_0 = motivavstånd, F = bländartal, d = maximalt tillåten oskärpecirkel och f = brännvidd. Detta är ett ganska komplicerat uttryck att arbeta med i det generella fallet. Låt oss därför titta på ett typiskt fall där kort skärpedjup är av intresse, t.ex. porträttfoto.

Låt oss först anta att vi arbetar med en fullformatskamera (sensorstorlek 24 mm x 36 mm), med användande av brännvidd 85 mm samt en rimlig tillåten oskärpecirkel (30 μ m). Då ger formel 10.9 ett totalt skärpedjup av ca. 7 cm vid fotografering på 1.5 meters avstånd med bländartal 4. Detta skulle kunna vara typiska förhållanden vid porträttfotografering, och innebär att bara själva ansiktet avbildas med full skärpa. Låt oss sedan anta att vi använder en typisk kompaktkamera (sensorstorlek ca. 4 x 5 mm) med brännvidd 12.5 mm (vilket motsvarar 85 mm på fullformatskameran, dvs lika mycket av motivet kommer med vid fotografering på samma avstånd). Tillåten oskärpecirkel sätter vi nu till 4.5 μ m, vilket motsvarar 30

um för fullformat eftersom kompaktkamerabilden förstoras mera på grund av den mindre sensorn. Kompaktkameran ger då ett totalt skärpedjup av drygt 0.5 meter vid samma motivavstånd och bländartal som för fullformatskameran. Nästan en tiopotens större skärpedjup för kompaktkameran alltså. För att få lika kort skärpedjup som med fullformatskameran, så skulle kompaktkameran behöva använda bländartal 0.55! Detta är naturligtvis omöjligt (se kap. 9.11 angående objektivs ljusstyrka). Dessutom är användande av mycket låga bländartal ofördelaktigt på grund av ökade avbildningsfel, samt risk för överexponering vid stark omgivningsbelysning.

Nu kanske någon invänder att längre brännvidder ger mindre skärpedjup, eftersom detta är något som många fotografer erfarenhetsmässigt upplevt. Skulle ökad brännvidd på kompaktkameran minska skärpedjupet? Tyvärr så medför längre brännvidd att man måste öka motivavståndet a_0 för att få med lika mycket av motivet på bilden (samma avbildningsskala). Effekterna av förlängd brännvidd och ökat motivavstånd tenderar att ta ut varandra, så att skärpedjupet förblir praktiskt taget konstant (däremot förändras perspektivet, se kap. 7). Tumregeln att längre brännvidder ger mindre skärpedjup gäller bara om motivavståndet är konstant. Faktum kvarstår alltså att för vårt tänkta fall med porträttfoto, så ger en kamera med stor sensor ett kortare skärpedjup.

Trots att den längre brännvidden inte ger en märkbar ändring i skärpedjupet i ovanstående exempel, så visar det sig att brännvidden spelar roll för hur suddiga motivdelar utanför skärpedjupsgränserna avbildas. Om vi tänker oss att vi har en punktkälla på oändligt avstånd (t.ex. en stjärna på himlen, eller approximativt en gatlykta på stort avstånd) och undersöker med hur stor oskärpecirkel den kommer att avbildas i kamerans sensorplan så erhålls följande enkla formel

$$d' = \frac{fM}{F} \quad (10.10)$$

där d' är diametern i den oskarpa bilden av punktkällan, f = brännvidden, M = avbildningsskalan och F = bländartalet. Formel 10.10 kan härledas med samma metod som för hyperfokalinställningen i kap. 10.2. Vid konstant bländartal och avbildningsskala, vilket vi hade i fallet med porträttfoto som vi tittade på tidigare, så blir alltså oskärpefläckens diameter d' från avlägsna föremål direkt proportionell mot brännvidden. Detta kan vi alltså tolka som att det som ligger utanför skärpedjupet blir kraftigare utsuddat vid långa brännvidder än vid korta. Man kan visa att detta gäller generellt, inte bara för motivdelar på oändligt avstånd.

Förutom hur suddigt det som ligger ur fokus är, så kommer också suddighetens karaktär att påverka bildintrycket. Hittills har vi enbart beaktat geometrisk optik, samt cirkulära bländaröppningar, men verkligheten ser inte riktigt ut så. Detta kommer att beröras i kap. 10.5.

10.5. "Bokeh"

När man använder ett kort skärpedjup så kommer ofta stora delar av bildytan att vara suddig. Karaktären, dvs utseendet, av denna suddighet kallas i fotokretsar för bokeh (från det japanska ordet för "suddig"). I olyckliga fall kan man få en oangenäm bokeh. Antag att vi har en stark, nästan punktformig ljuskälla i bakgrunden långt bortom skärpedjupets fjärrgräns. Formen på den oskarpa bilden av ljuskällan kommer då att avspegla bländaröppningens form (se Fig. 10.3). Bländaröppningar är vanligen ungefär cirkulära (i praktiken ofta 5-, 6-, 8- eller 9-kantiga), men det finns också exempel på trekantiga eller fyrkantiga öppningar. De sistnämnda kan ge störande bokeh, eftersom vi av tradition är vana att se cirkulära oskarpefläckar. Det kan också vara störande om det finns hack i bländarens kant, typ sågtänder, vilket kan förekomma i en del bländare. Rent generellt anser man oftast att en bländaröppning som är så nära en perfekt cirkel som möjligt (med jämna kanter) är att föredra.

Förutom bländaröppningens form, kommer även optiska avbildningsfel (som beskrivs i kap. 15) att påverka den bokeh som man får. Även diffraktionseffekter (som också beskrivs i kap. 15) kan spela in, speciellt för motivdelar som ligger bara lite utanför skärpedjupet. Låt oss se hur avbildningsfelen i praktiken kan inverka. Vi antar att bländaröppningen är en praktiskt taget perfekt cirkel, och att motivet är en punktkälla som ligger ur fokus nära bildcentrum. För detta fall har vi två avbildningsfel som spelar in, sfärisk aberration och kromatisk (längd)aberration. Kännedom om avbildningsfelen är inte nödvändiga i detta kapitel, eftersom vi bara ska ge några illustrationer på hur de kan inverka.

För en perfekt lins, dvs utan några avbildningsfel, skulle en defokuserad vit punktkälla avbildas som en perfekt cirkel med jämn vit belysning innanför cirkelns rand (och noll utanför), jämför Fig. 10.2. I verkligheten kommer belysningen inte att vara helt jämn, och dessutom kommer randen att uppvisa färg effekter. Dessa effekter är mer markanta på objektiv med sämre korrektion för avbildningsfel. Vi kan se exempel på detta i Fig. 10.9.

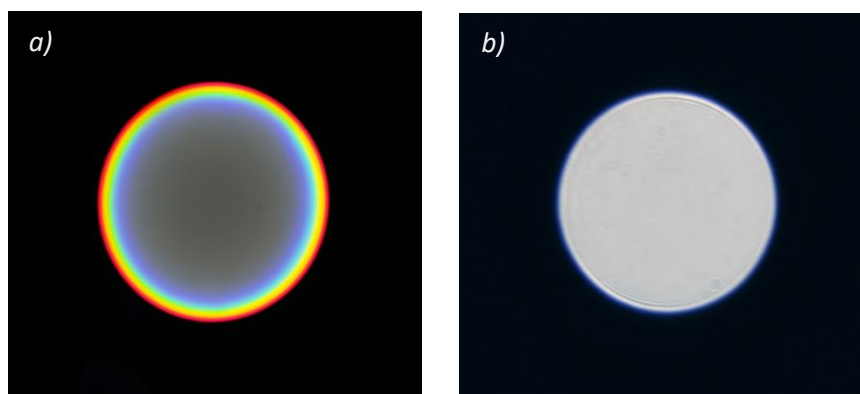


Fig. 10.9. a) Exempel på oskarpecirkelns utseende för ett objektiv med dålig korrektion för sfärisk och kromatisk aberration, samt b) ett objektiv med betydligt bättre korrektion. Effekten av sfärisk och kromatisk aberration är att belysningsfördelningen är ojämn samt att kanterna är färgade.

Det är uppenbart att en oskärpefläck med utseende som i Fig. 10.9 *a)* kan verka mycket störande i en bild som innehåller starka ljuskällor som ligger ur fokus (t.ex. gatlyktor på kvällen). I ett sådant fall är en oskärpefläck enligt figuren *b)* klart att föredra. Nu bör understrykas att *a)*-bilden visar ett *mycket* dåligt objektiv bara för att tydliggöra effekterna, och att en oskärpefläck enligt *b)*-bilden är vad man har rätt att vänta sig av ett fotografiskt objektiv av god kvalitet. Men eftersom sfäriska aberrationen beror både på objektivets ljusstyrka och det använda bländartalet, så kan oskärpefläckens utseende variera något också mellan olika objektiv av god kvalitet, och även för ett och samma objektiv beroende på bländartal. Oskärpefläckens utseende kan dessutom bero på om den ligger framför eller bakom fokuspunkten. Allt detta kan förefalla ganska komplicerat, men i praktiken brukar aberrationer ge en måttlig påverkan på oskärpefläcken. Den mest påtagliga påverkan (särskilt vid kraftig defokusering) ges istället av bländaröppningens geometriska form – är bländaröppningen 5-kantig, så blir oskärpefläcken 5-kantig etc.

Hittills har vi bara talat om oskärpefläckens utseende nära bildcentrum. Vad händer när man betraktar situationen ute vid bildkanten? Enligt geometrisk optik så borde inte oskärpefläcken ändra vare sig storlek eller utseende när man rör sig från bildcentrum ut mot bildkanten. I praktiken kan man förvänta sig en viss påverkan, eftersom avbildningsfelen blir större ute i bildkanten. Dessutom kan vinjettering (kap. 9.13) medföra att oskärpefläckens form ändras på grund av skuggningseffekter vid stora bildvinklar. Men, återigen, den största påverkan brukar ges av bländaröppningens geometriska form.

Om man lite kort ska sammanfatta innehållet i detta delkapitel som handlar om bokeh, så skulle det kunna vara:

- Bländarens geometriska form är viktig. Försök få den så cirkulär som möjligt, och med jämna kanter.
- Objektiv med låg ljusstyrka (dvs lägsta bländartal är ganska högt!) har som regel mindre avbildningsfel, och ger därmed en mer optimal belysningsfördelning i oskärpefläcken. Detta leder till en snyggare bokeh.

Ett stort problem vad gäller bokeh är att man inte kan kvantifiera den på något meningsfullt sätt. Vi befinner oss därför i den tekniska fotografins träskartade utmarker, där inget riktigt fotfäste finns. Subjektiva åsikter är det som gäller, vilket i värsta fall kan resultera i ett oändligt käbbel utan att man kommer någonstans.

11. Exponeringsmätare

11.1. Inledning

Sensorn i en kamera måste exponeras lagom mycket för att man ska få en bra bildkvalitet. Som vi såg i kap. 9 så är exponering $H = E \times t$, där E är belysningen på sensorn och t är exponeringstiden. För att reglera belysningen används bländaren, och tiden bestäms av slutaren. Men hur ska bländare och slutare ställas in så att exponeringen blir lagom? Det är här exponeringsmätaren kommer in. Exponeringsmätarens uppgift är att mäta upp hur hög belysningen är på sensorn. Med kännedom om detta, samt hur hög exponering som är lagom för den aktuella sensorn, kan kameran räkna ut vad som är en lämplig exponeringstid. Om den inställda tiden är för kort signalerar kameran underexponering, och om den är för lång signaleras överexponering. Genom att koppla ihop exponeringsmätaren med kamerans bländar- och tidsinställningar kan man få kameran att automatiskt ställa in värden som ger korrekt exponering (mer om detta i kap. 24).

Ofta fungerar sensorn i en digitalkamera också som exponeringsmätare. Men i digitala spegelreflexkameror sitter sensorn bakom en stängd ridåslutare fram till att en bild exponeras. Dessutom blockerar den nedfällda spegeln ljuset. Istället finns i spegelreflexkamerorna en speciell sensor som används både för den automatiska fokuseringen (se kap. 24) och för exponeringsmätning.

11.2. Viktning

Olika delar av ett motiv kan vara olika ljusa, och alla delar av motivet är kanske inte lika viktiga. För att ta hänsyn till detta kan man som fotograf ställa in olika viktningssmetoder för exponeringsmätningen. I fig. 11.1 visas tre sådana viktningssmetoder.

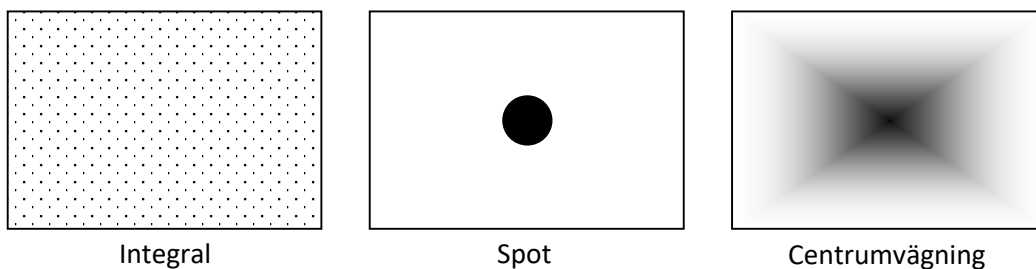


Fig. 11.1. Exempel på hur exponeringsmätare kan vikta över bildytan.

Integralmätning (ovanlig numera) innebär att mätvärdet baseras på medelbelysningen över bildytan. Vid spotmätning tas bara hänsyn till belysningen i bildcentrum. Centrumvägning innebär att mätaren tar mest hänsyn till belysningen i centrum, och viktar ner värdena alltmer ju längre ut i bildkanten man kommer. Förutom dessa varianter finns t.ex. centrum/botten-vägning (intressantaste motivdelarna finns ofta i nederkant) och avancerade mätrutiner där kamerans processorer bestämmer exponering utgående från bildinnehåll.

För att ta ett konkret exempel på nyttan av att använda rätt viktning, kan vi tänka oss att vi är på fjällsemester och ska fotografera en solbadande kamrat med vackert brunbränt ansikte mot en stor, bländande vit snöbakgrund. Antag att kameran arbetar med integral- eller centrumvägning. Eftersom ansiktet kanske upptar bara en mindre del av totala bildytan kommer den vita snöbakgrunden att dominera i mätvärdet som vi avläser. Följer vi den rekommenderade exponeringen som mätaren ger, kommer vi därför att få en medelgrå ton i snöpartierna och kamratens ansikte kommer att se ut som en svart fläck utan några detaljer alls. Detta beror på att exponeringsmätaren inte förstår vad snö är, och att den ska vara bländvit. Den kan bara gissa att motivet i medeltal är mellangrått (för så ser medelmotivet ut som den är kalibrerad för). Om vi istället använder spotmätning är vi mycket bättre rustade att klara av ovanstående fotograferingssituation. I så fall kan vi helt enkelt se till att mätningen utförs bara på kamratens ansikte, varvid mätaren inte kommer att känna av den vita snöbakgrunden. Kamratens ansikte kommer då att bli korrekt exponerat, medan bilden blir kraftigt exponerad och vit i bakgrunden (och det stämmer ju med verkligheten).

Det man får komma ihåg vid användning av en spotmätare, är att mätaren inte kan förstå om den lilla del av motivet som vi mäter på är ett ljust föremål (en marmorstaty) eller ett mörkt föremål (en svart panter). Om man följer mätarens rekommendation kommer dessa att återges lika ljusa på bilden. Kort sagt, man behöver alltså både veta hur mätaren fungerar och dessutom använda huvudet när man tolkar mätarens rekommendation för exponering. Ett hjälpmedel som många professionella fotografer använder är en matt jämngrå skiva med 18% reflektion. Skivan placeras i motivet, t.ex. precis framför en fotomodell. Exponeringsmätningen görs mot skivan, och man ställer in exponeringen efter mätarens rekommendation. Därefter tar man bort skivan och exponerar. Tanken här är att 18% reflektion motsvarar medelreflektionen hos ett ”typiskt” motiv, och det är just detta som exponeringsmätare är kalibrerade för att återge korrekt. Om motivet i medeltal är mörkare eller ljusare än så får vi en lägre respektive högre ljusnivå i bilden. Men detta är ju precis vad vi vill ha! Ett mörkare motiv ska ge en mörkare bild. Genom att mäta exponeringen mot en 18% gråskiva slipper vi alltså problemet med felexponering av svarta pantrar och vita marmorstatyer som diskuterades ovan. Men i praktiken är det ju inte alltid man kan placera ut gråkort i motivet vid fotografering.

11.3. Histogram

Många digitalkameror kan presentera ett histogram över hur bildens pixlar fördelar sig på olika ljusvärden, Fig. 11.2. På horisontella axeln i histogrammet avsätts pixelvärde (ofta 0-255 eftersom bilder med 8 bitar är vanligast). På vertikala axeln avsätts hur många pixlar som har ett visst värde. Histogrammet kan utnyttjas som en utmärkt exponeringsmätare i det fall att man har ett så kallat ”live-view” histogram, dvs att man innan exponeringen kan se hur bildens histogram kommer att se ut. Vissa kameror kan också presentera separata histogram för de tre färgkanalerna för bilder som sparats i kamerans minne. Detta möjliggör en bedömning både av bildens färgbalans och exponering, se kap. 13.6

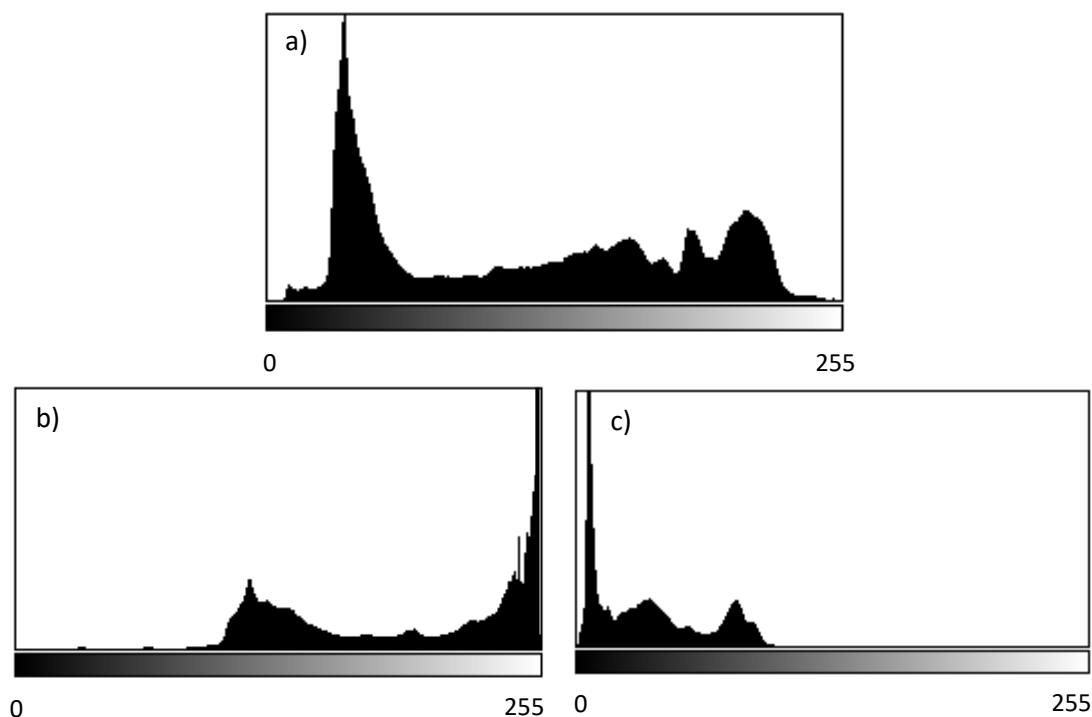


Fig. 11.2. Många digitalkameror kan på displayen visa ett histogram för att underlätta inställning av korrekt exponering. a) Histogram för en korrekt exponerad bild. b) Överexponerad bild. c) underexponerad bild. Se texten för detaljer.

I Fig. 11.2 a) visas histogrammet för en korrekt exponerad bild. De ljusaste pixlarna ligger då nära, men inte ända uppe vid, värdet 255. När pixelvärdet blir 255 har vi bottnat, dvs överexponerat, bilden och information går förlorad. De mörkaste pixlarna har låga värden, men inte ända nere vid noll eftersom detta skulle innebära att vi förlorar information i mörka bildpartier. Delfigurerna b) och c) visar hur histogrammen ser ut för över- respektive underexponerade bilder. I överexponerade bilder kommer vissa bildpixlar att ha värdet 255, vilket betyder att vi har förlorat information i ljusa bildpartier. I underexponerade bilder hamnar vissa bildpixlar på alltför låga värden, vilket gör att information i mörka partier förloras. Naturligtvis är det ofta acceptabelt att enstaka bildpixlar har värdena 255 eller 0 om det rör sig om ointressanta delar av bilden (t.ex. ett blänk i en fönsterruta eller en sotsvart skugga). Med ett histogram har man emellertid fullständig kontroll över vilka pixelvärden man har i bilden. Detta gör att man ofta kan få en bättre exponerad bild om man arbetar i manuell mod, och använder histogrammet, än om man låter kameran själv automatiskt ställa in exponeringen.

Om motivet har ett förhållandevis lågt kontrastomfång kommer pixelvärdena inte att spänna över hela intervallet från nästan 0 till nästan 255. I sådana fall är det i allmänhet bäst att exponera så att de ljusaste pixlarna hamnar ganska nära 255, vilket innebär att inga pixlar hamnar nära 0. På så sätt kommer vi att få förhållandevis höga signalnivåer i alla pixlar, vilket är fördelaktigt ur brussynpunkt.

12. Färglära

12.1. Inledning, färgseendet

Det mänskliga ögat har tre olika sorters färgkänsliga receptorer, så kallade tappar, i näthinnan. Dessutom finns icke färgkänsliga receptorer som kallas stavar, vilka bara är aktiva vid låga ljusnivåer (mörkerseende). De olika typerna av tappar är (grovt förenklat) känsliga för rött, grönt respektive blått ljus, Fig. 12.1. Känslighetskurvorna överlappar dock, vilket betyder att såväl monokromatiskt ljus som mer utbredda emissionsspektra ofta stimulerar flera typer av receptorer samtidigt. Hur färgintryck uppstår i hjärnan är ett mycket komplicerat skeende (hjärnan är en kraftfull bildprocessor). Färgintrycket beror inte bara på de ljusvåglängder och styrkor som ett föremål sänder ut mot ögat, utan också på sådana faktorer som färg och ljushet hos föremålets omgivning. I detta kompendium ska vi bara använda en mycket förenklad modell av färgseendet, vilken dock räcker för att förstå hur vi kan fotografera i färg.

Enligt denna enkla modell beror intrycket av färg och ljushet på kombinationen av signalstyrkor från de tre olika typerna av tappar (a, b, c i Fig. 12.1). Exempelvis kommer monokromatiskt (enfärgat) ljus av våglängden 450 nm att i huvudsak stimulera blåkänsliga tappar. Hjärnan uppfattar då ljuset som blått. Ljus med våglängden 530 nm stimulerar inte de blåkänsliga tapparna alls, medan de grönkänsliga stimuleras kraftigt och de rödkänsliga något mindre. Detta uppfattar hjärnan som grön färg. Om vi ökar våglängden till 560 nm stimuleras grön- och rödkänsliga tappar ungefär lika starkt, vilket vi uppfattar som gult ljus. Vid våglängder över 630 nm stimuleras praktiskt taget bara rödkänsliga tappar, och ljuset upplevs som samma röda färg oberoende av våglängden.

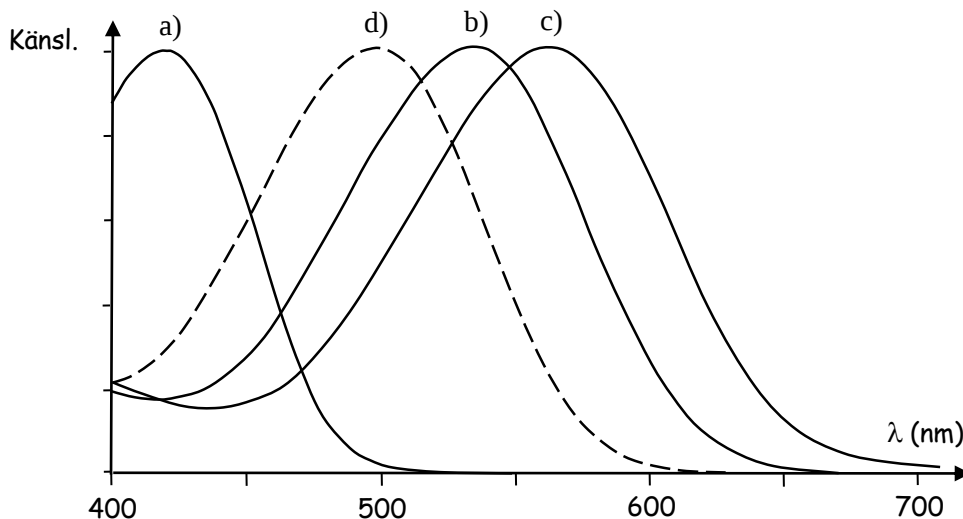


Fig. 12.1. Spektrala känslighetskurvor för de tre typerna av tappar som finns i det mänskliga ögat. För enkelhets skull benämner vi dessa blåkänsliga a), grönkänsliga b) och rödkänsliga c) tappar. Detta är något oegentligt eftersom känslighetskurvorna överlappar varandra kraftigt. Kurvan d) gäller för stavarna som används vid mörkerseende. Kurva a) antyder att det finns en inte obetydlig känslighet för våglängder under 400 nm, men dessa våglängder absorberas i ögat innan de når näthinnan.

Verkliga föremål sänder praktiskt taget aldrig ut monokromatiskt ljus, utan oftast ett brett, kontinuerligt spektrum. Men även i detta fall gäller att hjärnans färgintryck beror på kombinationen av signalstyrkor från de tre typerna av tappar. Sålunda är inte citroner gula därför att de sänder ut monokromatiskt ljus med en våglängd runt 560 nm (författaren har kontrollerat detta med en spektrometer!). Istället sänder citronen ut ett kontinuerligt spektrum med högre intensitet i området 500 till 700 nm, och lägre intensitet i området 400 till 500 nm. Detta ger en kraftig, och ungefär lika stor, utsignal från grön- och rödkänsliga tappar, och en betydligt lägre från de blåkänsliga. Detta kommer enligt vad som sagts ovan att ge ett gult färgintryck, liknande det som erhålls från monokromatiskt ljus med 560 nm våglängd. Vi kan därför dra slutsatsen att ögat inte är något bra instrument för att bedöma spektralfördelningen för ljus. Många helt skilda spektralfördelningar kan ge samma färgintryck i hjärnan. För fotografins del är detta dock en fördel, eftersom vi kan återskapa korrekta färgintryck utan att vi behöver exakt återskapa motivets spektralfördelning.

12.2. Additiv färgsyntes

En slutsats av vad som sagts ovan, är att vi kan skapa en mängd olika färgintryck (och ljushetsnivåer) genom att belysa ögat med ljus bestående av rött, grönt och blått i olika blandningsförhållanden, exempelvis med den uppställning som visas i Fig. 12.2. Genom att variera styrkorna på dessa ljuskällor kan vi variera styrkan på utsignalerna för de tre typerna av tappar.

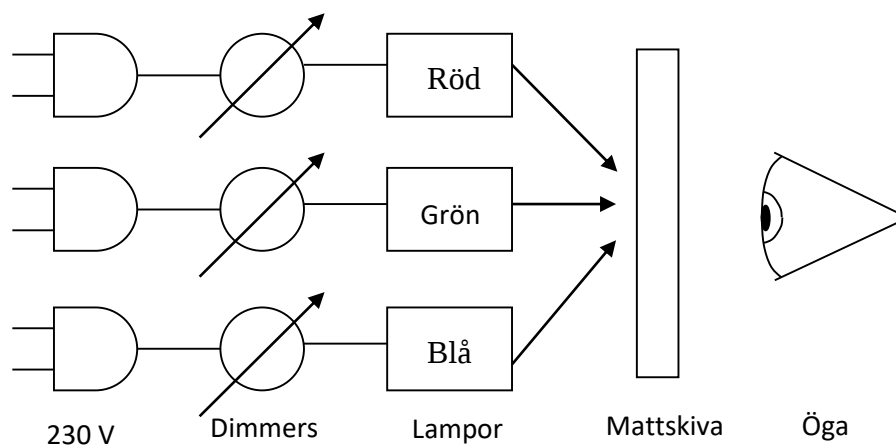


Fig. 12.2. Experimentuppställning för att åstadkomma olika färgintryck hos en betraktare. Genom att variera styrkan på de tre lamporna kan såväl färgton som ljusstyrka varieras.

I Fig. 12.3 ges några exempel på färgintryck som erhålls vid olika intensiteter på de tre lamporna. För enkelhets skull antar vi att ljuset från lamporna vardera upptar ca. 1/3 av synliga spektrat, och inte överlappar spektralt. Dessa tre spektraldelar kallar vi för enkelhets skull blått (B), grönt (G) och rött (R) ljus.

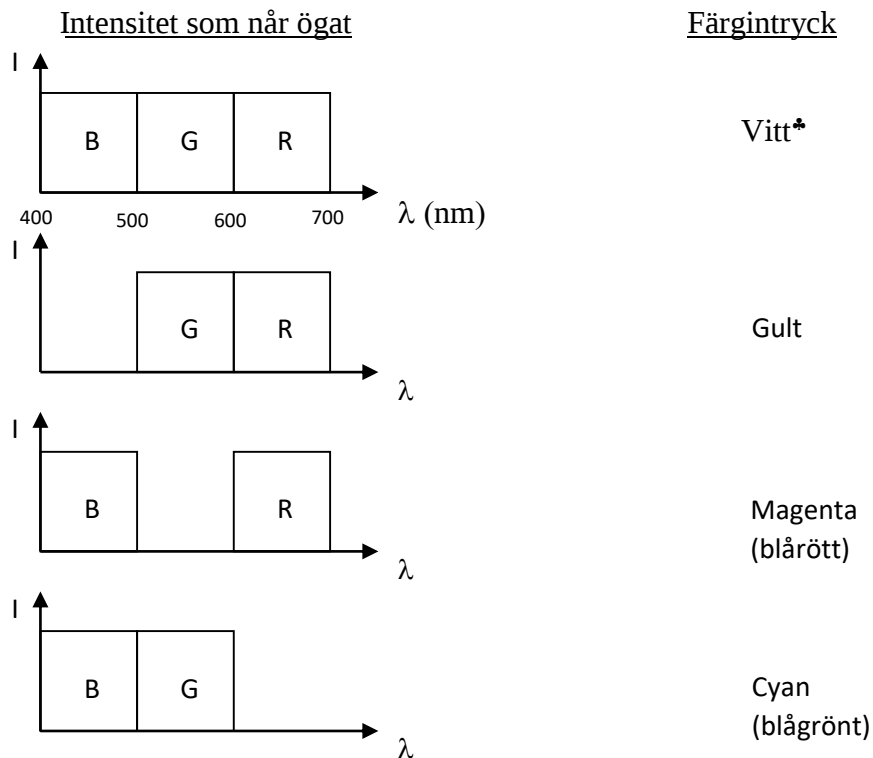


Fig. 12.3. Exempel på ljus med olika spektralsammansättning, samt de färgintryck som åstadkoms.

I exemplen i Fig. 12.3 är lamporna antingen av eller lyser med full styrka. Det innebär att man upplever de erhållna färgerna gult, magenta och cyan som mättade (= kraftiga). Vad skulle hända om man låter den blå lampan lysa med halv styrka, och de gröna och röda med full styrka? Jo, detta skulle vi uppleva som gult (eftersom grönt och rött dominerar), men färgmättningen skulle vara lägre (blekgul färg) eftersom vi har blandat in en del blått. Vi kan också se detta fall som att vi blandar mättat gult ljus med vitt ljus, vilket naturligt nog ger en blekare gul färg. Samma resonemang gäller för färgmättningen för de övriga färgerna.

Metoden att skapa färgintryck som beskrivits ovan kallas **additiv färgsyntes**, eftersom vi sänder ut och blandar ljus i tre olika våglängdsband. Beroende på blandningsförhållandet erhålls olika färgintryck (inte bara de som visas i figuren ovan). Exempel på områden där additiv färgsyntes används är TV- och datorskärmar. Dessa är uppbyggda av ett stort antal tätt liggande R-, G- och B-emitterande punkter vilkas intensiteter kan styras. Eftersom vi betraktar skärmarna på stort avstånd kan ögat inte uppfatta individuella punkter, och därför behövs inte någon mattskiva som i Fig. 12.2 för att blanda ljuset. I fotografins barndom skapades också färgbilder enligt den additiva principen, men senare färgfilmer fungerar istället enligt den subtraktiva principen som beskrivs i avsnitt 12.3.

* Spektrat som visas har samma intensitet för alla våglängder. Detta är inte riktigt fallet med "vitt" ljus typ solljus, vars spektrala emission följer lagarna för svartkroppsstrålare, se kap. 13. Inom området 400 – 700 nm är dock skillnaderna mellan dessa två fall mycket liten.

För att kunna återge en färgbild krävs att man i varje bildpunkt vet hur kraftig styrkan ska vara på R-, G- och B-komponenterna. Ett krav vid fotograferingen är då att man separat registrerar dessa tre komponenter. I digitalkameror uppnås detta vanligen genom att vissa pixlar i sensorn bara är känsliga för blåa våglängder, medan andra bara är känsliga för gröna respektive röda våglängder, se kap. 14. I färgfilm erhålls samma effekt genom att olika emulsionslager är känsliga för olika våglängder.

12.3. Subtraktiv färgsyntes

Subtraktiv färgsyntes baseras på att vitt ljus får absorberas i ett eller flera på varandra följande filter. Denna typ av färgsyntes används när man trycker färgbilder i t.ex. böcker och tidningar. Subtraktiv färgsyntes används också i färgfilm och färgfotopapper.

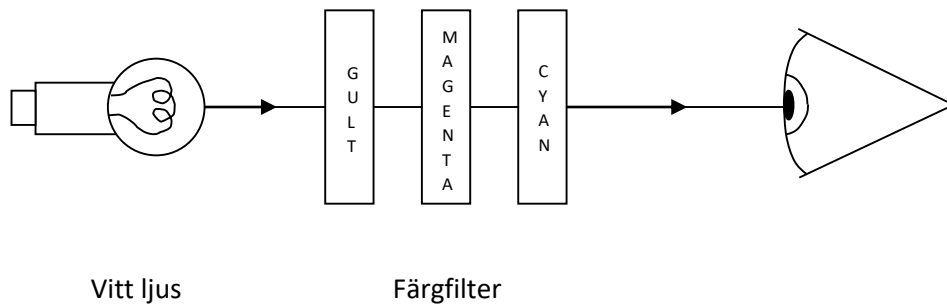


Fig. 12.4. Principen för subtraktiv färgsyntes.

Genom att variera färgmättnaden på de tre filtren kan olika färgintryck skapas. De använda färgfiltren har färgerna gult, magenta och cyan, och brukar betecknas med bokstäverna Y, M och C (efter engelskans Yellow, Magenta och Cyan). YMC-filtrens spektrala transmissionskurvor ser (i det ideala fallet) ut som i Fig. 12.5.

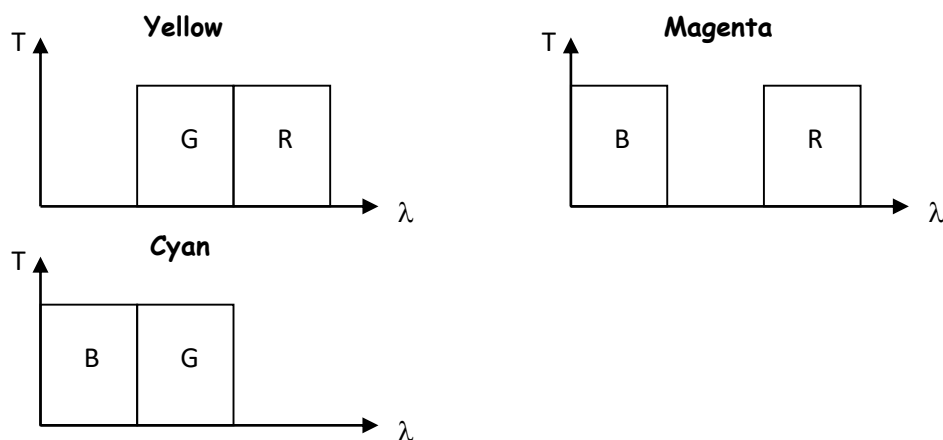


Fig. 12.5. Spektrala transmissionskurvor för mättade gul-, magenta- och cyanfilter.

För att få fram spektrala transmissionen vid kombination av flera filter i en ”sandwich” gäller att man skall *multiplisera* spektrala transmissionskurvorna för de enskilda filterna. Exempel på resultat vid kombination av ovanstående filter blir då:

$$Y + M \Rightarrow \text{Rött} \quad Y + C \Rightarrow \text{Grönt} \quad M + C \Rightarrow \text{Blått} \quad Y + M + C \Rightarrow \text{Svart}$$

De YMC-filter som visas i Fig. 12.5 har en maximal färgmättnad i och med att t.ex. gulfiltret absorberar 100% av infallande blått ljus. I analogi med vad som sades i fallet additiv syntes, kan man tänka sig ett gulfilter som transmitterar en del blått ljus (men inte så mycket som grönt och rött). Ett sådant gulfilter kommer då att ha en lägre färgmättnad. Det som karakteriserar ett gulfilter är alltså att det transmitterar bra (ofta nära 100 %) i grönt och rött, och mindre i blått (hur mycket beror på färgmättnaden). Analogi gäller för magenta- och cyanfilter. Eftersom ett gulfilter alltså absorberar i blått säger man att gult och blått är **komplementfärger**. På samma sätt är magenta och grönt komplementfärger, och likaså cyan och rött. Förståelsen av dessa begrepp är av grundläggande betydelse för förståelsen av hur subtraktiv färgsyntes fungerar.

En förenkling som gjorts i ovanstående figurer är att det antagits att filterkurvorna har ett idealiserat ”fyrkantigt” utseende. I verkligheten är inte detta fallet, utan kurvorna har (mer eller mindre) avrundade hörn samt överlappar varandra delvis. Detta ändrar dock inte det principiella resonemanget. Något man bör vara observant på är att många filterkurvor visar spektrala *densiteten* istället för spektrala *transmissionen*. Spektrala densiteten definieras som

$$D(\lambda) = -\log_{10} \left(\frac{1}{T(\lambda)} \right), \quad (12.1)$$

där $T(\lambda)$ betecknar filtrets transmission som funktion av våglängden. Hög densitet motsvarar alltså låg transmission och vice versa.

12.4. Färganalys

Ofta finns ett behov av att kvantitativt kunna beskriva olika färger, t.ex. genom de R-, G- och B-komponenter som beskrevs i avsnitt 12.2. Ett stort problem med detta är att färg är något vi upplever i våra hjärnor, och att det därför inte går att mäta direkt (åtminstone inte med dagens vetenskapliga metoder). Man tvingas därför använda indirekta mätmetoder, och att införa begreppet ”normalobservatör”. Men mer om detta i avsnitt 12.5, först ska vi titta på ett enklare fall av färgbeskrivning utan anspråk på att värdena ska kunna hänföras till någon mätnormal.

Låt oss betrakta en digital färgbild. Denna består av ett stort antal pixlar, var och en beskriven med tre mätetal R, G och B. Dessa tre tal anger hur kraftigt datorskärmen röda, gröna och blåa färgpunkter ska lysa. Talen är ofta i intervallet 0-255 (8 bitar), och kan visas genom att föra cursorn över en bild i Photoshop eller liknande program. Antag att en pixel har värdena RGB = (250, 50, 50). Enligt vad som tidigare sagts i detta kapitel, inser vi att detta motsvarar en omättad röd färg med

hög intensitet. En annan pixel i bilden har kanske värdena $RGB = (50, 10, 10)$. Även detta motsvarar en omättad röd färg, men med lägre intensitet. I bägge fallen har vi att grön- och blåvärdena är en femtedel så stora som rödvärdet. Förhållandet mellan RGB -värdena är alltså lika stort, vilket innebär att vi har samma *färgton* och *färgmättnad*. Däremot är intensiteten (ljusheten) olika i de bägge fallen. Vid färganalys brukar man nöja sig med att betrakta färgtonen och färgmättnaden, och inte bry sig om ljusheten. Detta åstadkoms genom att räkna fram normerade rgb -värden (färgkoordinater) genom

$$r = \frac{R}{R + G + B}, \quad g = \frac{G}{R + G + B}, \quad b = \frac{B}{R + G + B} \quad (12.2)$$

r , g och b , ligger i intervallet 0 till 1 och spänner upp ett plan enligt Fig. 12.6a. Eftersom vi vet att $r + g + b = 1$, så finns all information i en projektion i rg -planet (vi vet ju att $b = 1 - r - g$). En sådan projektion visas i Fig. 12.6b, och motsvarar att vi tittar uppifrån och ner längs b -axeln i Fig. 12.6a. Varje pixel i en digital bild kan på detta sätt tilldelas en position i färgtriangeln beroende på vilken färgton och färgmättnad den har.

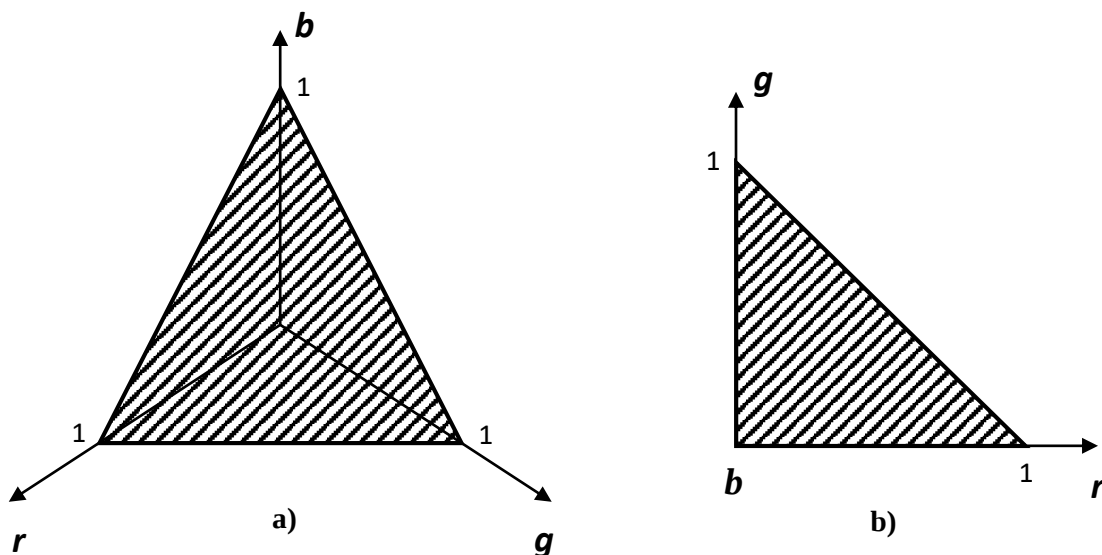


Fig. 12.6. Färgtonen och färgmättnaden för t.ex. en bildpixel kan anges genom dess rgb -komponenter. Eftersom de normerade rgb -värdena alltid ligger på det streckade planet i a), så räcker det att betrakta den tvådimensionella projektionen i delfigur b) för att entydigt ange färgtonen och färgmättnaden.

De rgb -värden som erhålls på detta sätt ansluter inte till någon absolut mätskala, och kan därför inte jämföras mellan olika system. Exempelvis kan samma digitala bild återges med något olika färgtoner på olika datorskrämar, trots att rgb -värdena är identiska. För att råda bot mot detta kan man använda det så kallade CIE-systemet som beskrivs i nästa avsnitt.

12.5. CIE-systemet

CIE-systemet* bygger på begreppet normalobservatör, vars egenskaper har fastställts som medelvärde av mätningar på ett stort antal individer. Till skillnad från de färgkoordinater, *rgb*, som erhöles i föregående avsnitt, är CIE-värdena knutna till välbestämda referenser. Detta innebär att om färgkoordinaterna är lika enligt CIE-systemet, så kommer också en (normal)observatör att uppleva färgerna som lika i verkligheten. CIE-värden anges ofta för datorskärmar, ljuskällor mm, och tillåter då en rättvis jämförelse av deras färgegenskaper.

Som utgångspunkt för en beskrivning av CIE-systemet kan vi ta den experimentuppställning för additiv färgsyntes som visas i Fig. 12.2. Men istället för bredbandiga emissionsspektra, antar vi istället att ljuskällorna sänder ut monokromatiskt ljus av våglängderna 435.8, 546.1 och 700 nm (valet av våglängder följer CIE-standarden). På samma sätt som beskrevs i avsnitt 12.2, kan man med dessa ljuskällor åstadkomma en mängd färg- och ljushetsintryck hos en observatör. Låt oss nu införa begreppet ”tristimulus values” (vi använder engelska beteckningar eftersom nästan all litteratur i ämnet är på detta språk). Tristimulus values (givna som R, G, B-värden) talar om hur höga intensiteter vi behöver ha på de tre lamporna i experimentuppställningen för att få samma färgton och ljushet som vårt försöksobjekt, t.ex. en viss färg som visas på en datorskärm. ”Chromaticity coordinates”, *rgb*, är tristimulus values normerade enligt ekv. 12.2.

Som ett första experiment med denna uppställning kan vi tänka oss att vi undersöker vilka tristimulus values vi får för spektralrena färger. Uppgiften blir då att undersöka t.ex. hur höga intensiteter av våglängderna 435.8, 546.1 och 700 nm vi ska blanda för att det färgmässigt för vårt öga ska se ut som monokromatiskt ljus med våglängden 570 nm. Sedan upprepar man experimentet för alla våglängder i synliga området (med samma ljusintensitet för alla våglängder). När man plottar tristimulus values som funktion av våglängden på det ljus som ska matchas, får man tre funktioner som kallas color-matching functions och betecknas $\bar{r}_\lambda$, $\bar{g}_\lambda$, $\bar{b}_\lambda$, Fig. 12.7.

* CIE = Commission Internationale de l’Eclairage, dvs internationella belysningskommissionen, som antog dessa normer 1931.

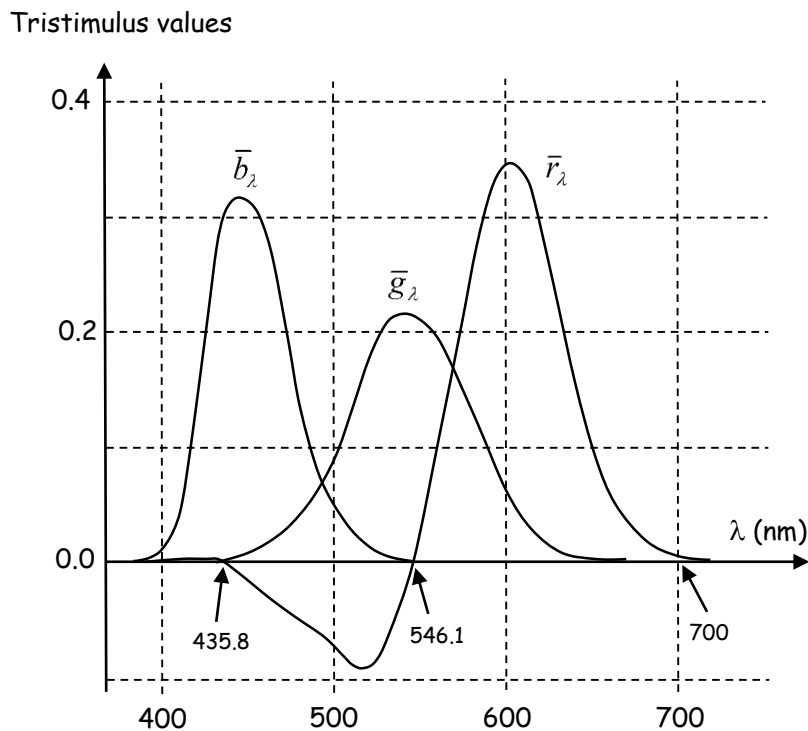


Fig. 12.7. "Color-matching functions", $\bar{r}_\lambda$, $\bar{g}_\lambda$, $\bar{b}_\lambda$, (tristimulus values för rena spektralfärger av samma intensitet). Vid de valda primärfärgerna, 435.8, 546.1 och 700 nm, är av naturliga skäl två av funktionerna lika med noll.

Av Fig. 12.7 framgår att funktionen $\bar{r}_\lambda$ kan anta negativa värden, t.ex. i området runt 500 nm. Vad betyder detta, man kan väl inte ha negativa ljusintensiteter? Jo, det betyder helt enkelt att monokromatiskt ljus med en våglängd av t.ex. 500 nm *inte* kan matchas färgmässigt genom att blanda ljus av våglängderna 435.8, 546.1 och 700 nm. I vilka proportioner man än blandar dessa färger, kommer det för ögat inte att färgmässigt likna ljus med våglängden 500 nm. Om man däremot blandar 500 nm ljuset med en lagom mängd (given av funktionen $\bar{r}_\lambda$'s negativa värde) av 700 nm ljus, så kan den blandningen färgmässigt matchas mot en lagom blandning (given av funktionsvärdena $\bar{g}_\lambda$ och $\bar{b}_\lambda$) av våglängderna 435.8 och 546.1 nm. Som vi senare ska se har man i CIE-systemet fixat till det hela så att man inte behöver räkna med negativa värden, men rent matematiskt/fysikaliskt utgör det inget hinder att tristimulus values kan bli negativa. Kom ihåg att syftet med CIE-systemet är att *mäta* färger, inte att *skapa* dem (det är ett annat problem, och naturligtvis kan man då inte jobba med negativa intensiteter).

Nu är vi mogna att titta på något mer avancerat än bara monokromatiskt ljus. Låt oss anta att vi vill bestämma tristimulus values för ljus av en godtycklig spektralfördelning, P_λ t.ex. enligt Fig. 12.8.

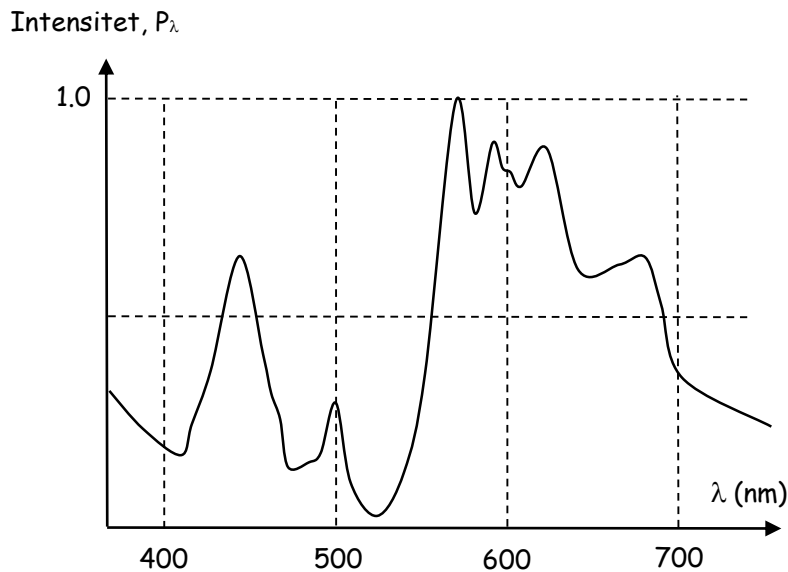


Fig. 12.8. Exempel på spektralfördelning för ljus vars tristimulus values ska bestämmas.

Vi kan då börja med att titta på ett infinitesimalt spektralområde $d\lambda$ vid våglängden λ . Detta lilla spektralområde kan vi behandla som monokromatiskt ljus, och tristimulus values ges av

$$\begin{aligned} R &= P_\lambda \bar{r}_\lambda d\lambda \\ G &= P_\lambda \bar{g}_\lambda d\lambda \\ B &= P_\lambda \bar{b}_\lambda d\lambda \end{aligned} \quad (12.3)$$

Alla tre värden blir infinitesimala genom att $d\lambda$ ingår i ekvationerna, men det är naturligt eftersom ljusintensiteten inom spektralområdet $d\lambda$ också är infinitesimal.

För att få tristimulus values för hela spektralfördelningen får summerar vi ihop bidragen för alla små delintervall $d\lambda$, dvs vi får integraluttrycken

$$\begin{aligned} R &= \int P_\lambda \bar{r}_\lambda d\lambda \\ G &= \int P_\lambda \bar{g}_\lambda d\lambda \\ B &= \int P_\lambda \bar{b}_\lambda d\lambda \end{aligned} \quad (12.4)$$

Dessa uttryck förutsätter att vi har linearitet, dvs att när man blandar ljus av olika färger så kan man addera de ingående ljusens tristimulus values. Experiment som utförts visar att detta stämmer med god noggrannhet. Genom att normera ovanstående RGB -värden enligt ekv. 12.2 erhålls "chromaticity coordinates" rgb . Dessa värden ger endast information om färgtonen och färgmättnaden, eftersom informationen om ljusheten försvinner i normeringen. Som vi tidigare sett, räcker det med två koordinatvärden, t.ex. r och g för att karakterisera färgton och

färgmättnad eftersom $r + g + b = 1$. En viss färg kommer då att representeras av en punkt i färgtriangeln i Fig. 12.6b. Det visar sig finnas många spektralfördelningar, med helt olika utseende, som resulterar i samma rgb -värden. Detta bekräftar det som sades i inledningen av kapitlet, nämligen att ögat inte är något bra instrument för att bedöma spektrala innehåll i ljus.

Chromaticity coordinates som räknas ut på ovanstående sätt kan emellertid resultera i negativa värden, vilket betyder att man hamnar utanför den skuggade triangeln i Fig. 12.6b. Detta beror på att color-matching function $\bar{r}_\lambda$ är negativ i området runt 500 nm, Fig. 12.7. Man har ansett att detta är en nackdel, och för att undvika negativa värden inför man en matematisk transform som räknar om rgb -värden till xyz -värden. Vi ska här inte gå in på hur denna transformation sker, men man använder helt enkelt tre formler av typen $x = f(r, g, b)$. Efter denna transformation har man fått tre färgkoordinater, xyz , som alltid antar värden mellan 0 och 1. Effekten av transformationen kan beskrivas som att man ändrar de color-matching functions som visas i Fig. 12.7 till att se ut enligt Fig. 12.9.

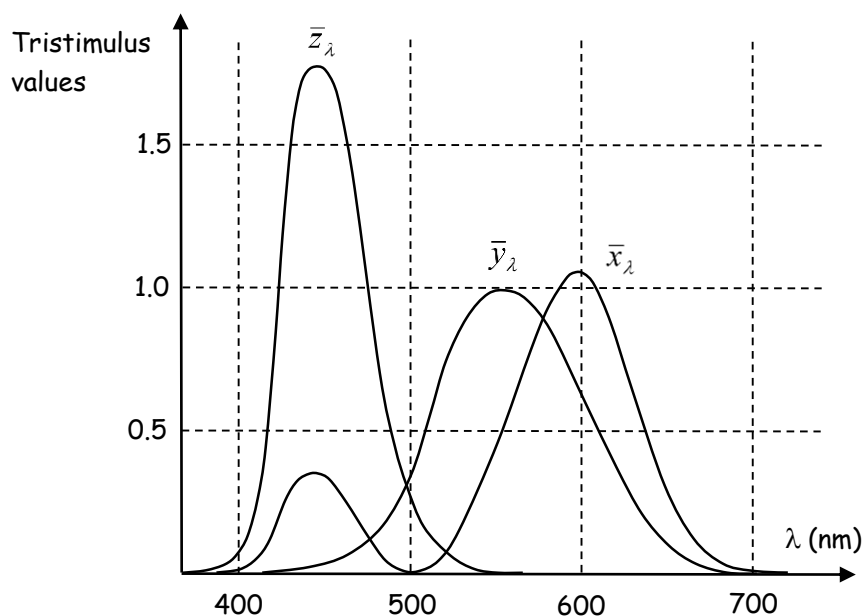


Fig. 12.9. CIE color-matching functions $\bar{x}_\lambda$, $\bar{y}_\lambda$, $\bar{z}_\lambda$. Kurvan $\bar{y}_\lambda$ representerar ögats totala spektrala känslighet (ljushetsintrycket som erhålls för olika λ).

I analogi med ekv. 12.4, kan vi nu definiera CIE tristimulus values, XYZ , utgående från de nya color-matching functions som visas i Fig. 12.9.

$$\begin{aligned} X &= \int P_\lambda \bar{x}_\lambda d\lambda \\ Y &= \int P_\lambda \bar{y}_\lambda d\lambda \\ Z &= \int P_\lambda \bar{z}_\lambda d\lambda \end{aligned} \tag{12.5}$$

Valet av color-matching functions, Fig. 12.9, gör att Y -värdet svarar mot ljushetsintrycket i ögat medan X och Z ger färginformation. Genom normering enligt

$$\begin{aligned}
 x &= \frac{X}{X+Y+Z} \\
 y &= \frac{Y}{X+Y+Z} \\
 z &= \frac{Z}{X+Y+Z}
 \end{aligned}
 \tag{12.6}$$

erhålls CIE chromaticity coordinates, xyz . Av dessa brukar x och y ritas upp i form av ett CIE(x, y)-chromaticity diagram. I Fig. 12.10 visas ett sådant diagram där de spektralrena färgerna mellan 400 och 700 nm lagts in.

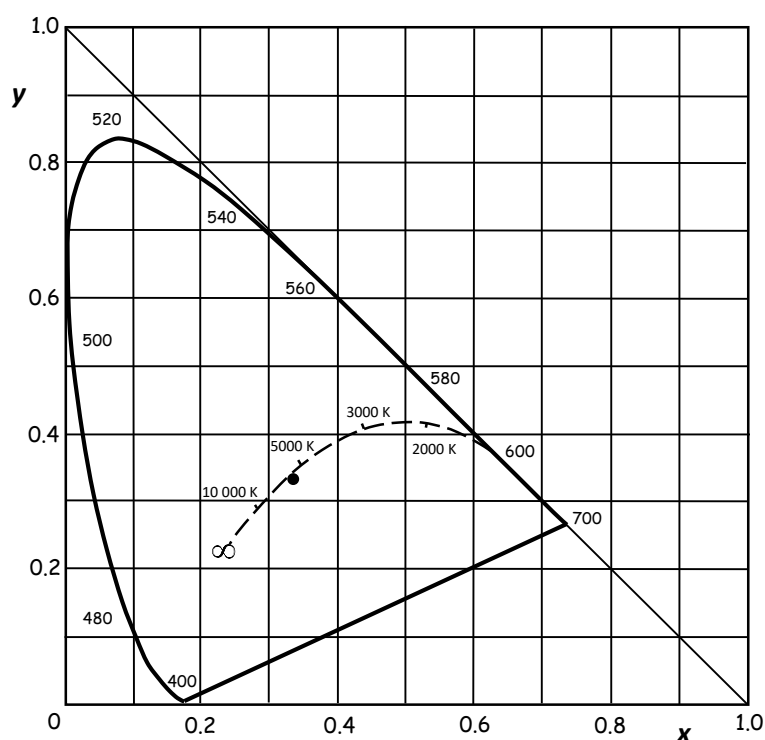


Fig. 12.10. CIE(x, y)-chromaticity diagram. Den hästskoformade kurvan visar läget för monokromatiskt ljus med våglängder mellan 400 och 700 nm. Den räta linjen mellan 400 och 700 nm motsvarar blandljus med olika proportioner av 400 och 700 nm ljus. Den svarta punkten motsvarar en ljuskälla med jämn spektralfördelning mellan 400 och 700 nm (vitt ljus i Fig. 12.3). Streckade linjen motsvarar svartkroppsstrålare med olika färgtemperaturer (se kap. 13) från 2000 K till ∞ .

Figur 12.10 innehåller ingen ljushetsinformation; den försvinner vid normeringen enligt ekv. 12.6. Olika färgtoner och färgmättnader kommer att representeras av olika positioner i den triangel som ligger till vänster om diagonalen. Alla områden i triangeln representerar emellertid inte möjliga färgtoner. I själva verket kommer alla färger att hamna inom ett område som utgörs av den hästskoformade kurvan, svarande mot spektralfärgerna, och den räta linjen mellan 400 och 700 nm punkterna på hästskokurvan.

Ett användningsområde för CIE(x, y)-chromaticity diagram är att uppskatta färgåtergivningen för "output-medium", t.ex. en datorskärm eller bläckstråleskrivare. Om vi väljer exemplet datorskärm, så byggs bilden upp av små

lysande röda, gröna och blå prickar. Antag att ljuset från dessa prickar svarar mot de färgkoordinater som visas i Fig. 12.11. Då gäller att endast färger som ligger inom den triangel som spänns upp av RGB-prickarnas färgkoordinater kan återges på skärmen. För att få en stor dynamik i färgåtergivningen, t.ex. att kunna återge mättade färger, är det därför viktigt att RGB-prickarnas färgkoordinater ligger utspridda så mycket som möjligt innanför hästskokurvan.

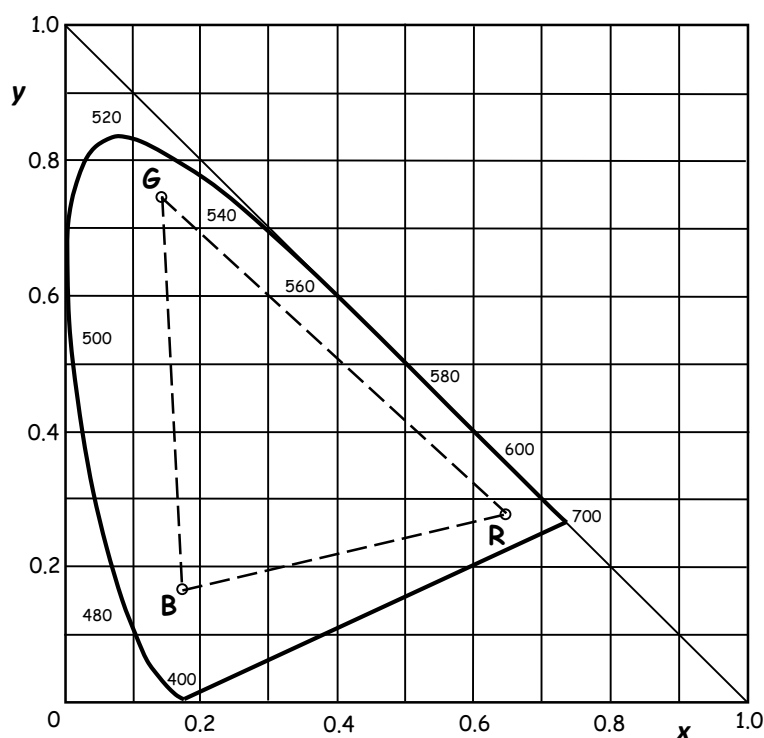


Fig. 12.11. RGB-punkterna representerar färgkoordinater för de tre typerna av färgpunkter i en tänkt datorskärm. Datorskärmen kommer då att kunna återge färger som ligger innanför den streckade triangeln.

12.6. Färgrymder: sRGB och AdobeRGB

I digitalkameror kan man ofta välja om jpeg-bilder ska sparas som sRGB eller AdobeRGB. Det som påverkas av denna inställning är vilka pixelvärden som kommer att representera olika färgtoner och mättnadsgrader. Det har skrivits mycket, inte minst på internet, om olika färgrymder och deras för- och nackdelar (det finns fler rymder än sRGB och AdobeRGB). I detta kompendium ska vi bara helt ytligt beröra ämnet för att förklara vad detta med färgrymder handlar om.

Låt oss ta ett enkelt exempel. Antag att vi fotograferar med en digitalkamera, och sedan betraktar bildfilerna på en datorskärm. Som visas i Fig. 12.11 beror skärmens ”spännvidd” vad gäller att återge olika färger (det brukar kallas ”gamut”) på var i CIE-diagrammet skärmens R-, G- och B-punkter ligger. Läget för dessa tre primärfärger (”primaries” på engelska) är naturligtvis lite olika för olika skärmar. I Fig. 12.11 ligger t.ex. röd primärfärg ungefär vid $x = 0.64$, $y = 0.28$. Primärfärgerna spelar en stor roll för färgrymder, och vi ska börja med att titta på sRGB.

Tanken med sRGB är att man ska få samma färgåtergivning av bilder på olika bildskärmar, och att man ska utnyttja ”typiska” bildskärmars kapacitet på ett bra sätt. Därför definierade man R-, G- och B-primärfärgerna baserade på typiska skärmprestanda (CRT-skärmar, dvs. elektronstrålerör) vid 1990-talets mitt*. Man fick då följande koordinater (x,y) för primärfärgerna:

$$R = (0.64, 0.33)$$

$$G = (0.30, 0.60)$$

$$B = (0.15, 0.06).$$

Dessa punkter finns inlagda i diagrammet i Fig. 12.12. Där finns också markerat en triangel som visar den gamut (färgspännvidd) som erhålls, samt vitpunkten W med koordinaterna $(0.31, 0.33)$ †.

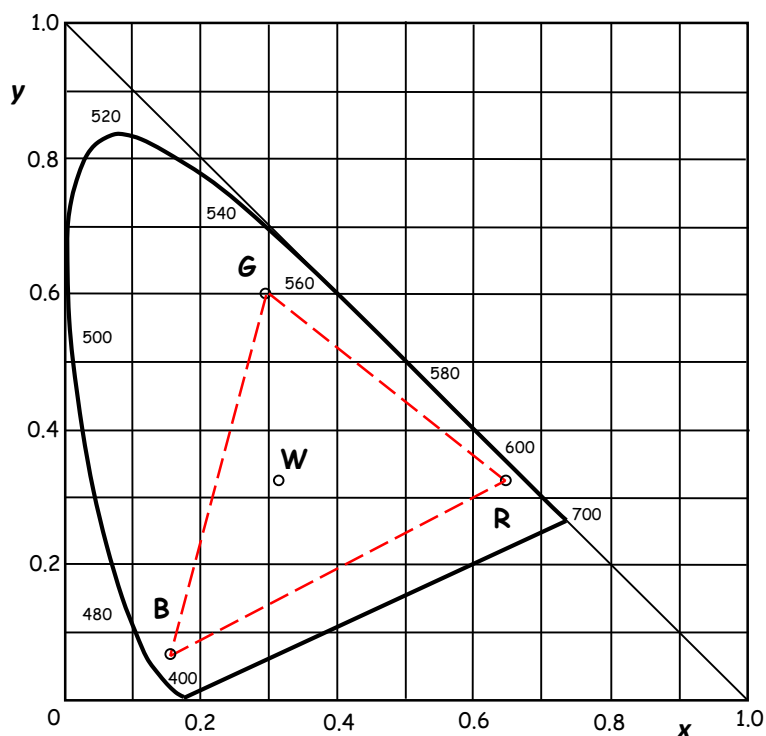


Fig. 12.12. **RGB**-punkterna representerar färgkoordinater för de tre primärfärgerna i färgrymden sRGB, och vitpunkten är markerad med **W**. Datorskärmens gamut (färgspännvidd) representeras av den streckade röda triangeln.

sRGB är den utan jämförelse vanligaste färgrymden. Så gott som alla bildskärmar använder den, och det är den färgrymd som används på internet och i de flesta andra sammanhang (sRGB står för ”standard RGB”). Som synes i Fig. 12.12 är färgåtergivningen något begränsad, särskilt inom det blågröna området. Men detta återspeglar, som sagt, typiska verkliga skärmprestanda. En jpeg-bild kan återge

* Det stämmer fortfarande rätt bra för de flesta skärmar.

† Denna vitpunkt har beteckningen D65 och motsvarar färgen på dagsljus med en korrelerad färgtemperatur av 6500 K (en blandning av direkt solljus och blått himmelsljus).

$2^8 \times 2^8 \times 2^8 = 1.68 \times 10^7$ olika färger (se kap. 14.7), och med sRGB kommer alla dessa färger att ligga inom röda streckade triangeln i Fig. 12.12. Man kommer då att optimalt utnyttja en skärm med ”typiska” verkliga prestanda*. Detta innebär att om skärmen skulle ha bättre gamut än den som ges av sRGB, så utnyttjas inte skärmens möjligheter till färgåtergivning fullt ut. Men, kanske någon undrar, om nu skärmen har en större gamut än vad som är typiskt, varför då inte låta t.ex. en rent grön pixel hamna färgmässigt högre upp än i Fig. 12.12 om motivet verkligen var klargrönt? Jo, därför att då kommer bildens färger att återges olika på olika datorskärmar, och det gäller inte bara mättade färger utan alla färgtoner. Det är just det problemet man vill undvika.

AdobeRGB utvecklades av Adobe Systems i slutet av 1990-talet. Likheter med sRGB är stora. En jpeg-bild med AdobeRGB innehåller, precis som sRGB, 8 bitar i varje färgkanal. Både sRGB och AdobeRGB kan därför återge lika många färger, nämligen $2^8 \times 2^8 \times 2^8 = 1.68 \times 10^7$. En jämförelse mellan färgrymderna för sRGB och AdobeRGB visas i Fig. 12.13. Skillnaden är att den gröna primärfärgen i AdobeRGB ligger lite högre upp (markerad **G'**), vilket ger en större gamut.

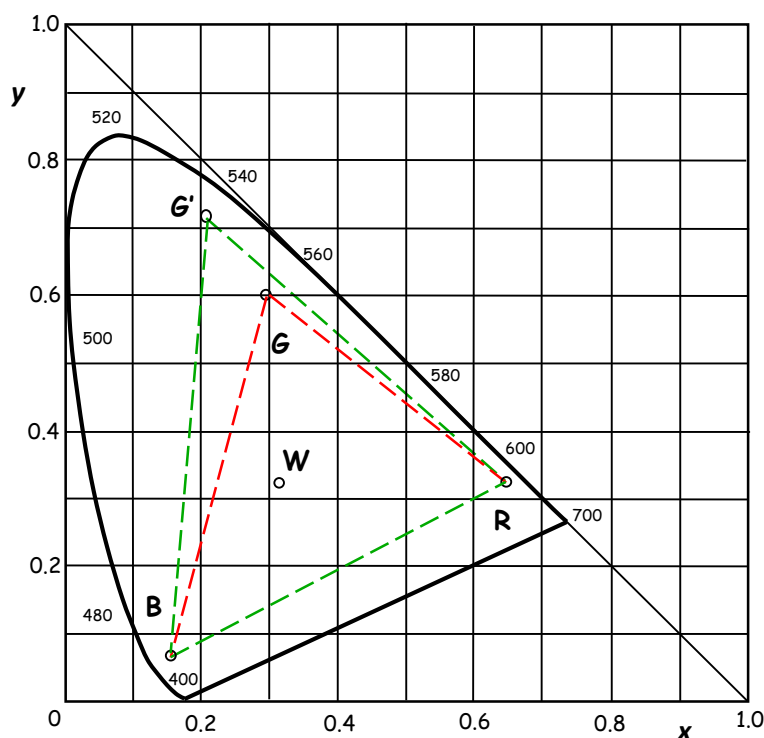


Fig. 12.13. Jämförelse mellan sRGB (rödmarkerad) och AdobeRGB (grönmarkerad). Det är endast grön primärfärg som skiljer sig åt. Den ligger för AdobeRGB på $x = 0.21$ och $y = 0.71$ (för sRGB gäller $x = 0.30$ och $y = 0.60$).

Att AdobeRGB har en större gamut än sRGB är en potentiell fördel. Det medför att man skulle kunna återge högre färgmättnad i gröna färgtoner. Men det krävs en

* Eftersom olika skärmar har lite olika egenskaper ska man för bästa resultat göra en kalibrering med utrustning som mäter ljuset från skärmen. Då kan man skapa en så kallad ICC profil, som skärmen sedan kan använda för att t.ex. pixelvärden (255,255,255) ska ge en färg som motsvarar vitpunkten och (255,0,0) motsvarar röd primärfärg etc. (Anm. 255 kan bytas mot vilket som helst annat 8-bitars tal som är skilt från noll.)

mycket bra datorskärm för att den ska leva upp till den gamut som motsvarar AdobeRGB. Och faran är stor att en bild med färgrymden AdobeRGB i slutändan kommer att visas som sRGB på t.ex. internet eller en bildskärm av normal kvalitet. Resultatet kommer då att bli en förlust i färgmättnad, inte bara ner till sRGB-nivå, utan klart sämre än så. Alltså, en AdobeRGB-bild som återges via sRGB kommer att se sämre ut än en sRGB-bild. Detta beror på att den gröna triangeln ”krymps” ner till den röda. Därvid kommer även omättade färger som redan ligger inom röda triangeln att krympa ner mot vitpunkten, dvs återges alltför omättade. Och dessa omättade färger utgör ju den absolut största andelen av pixlarna i en normal bild.

Men om man omvandlar AdobeRGB-bilden till sRGB innan den visas på en sRGB-skärm kan man undvika problemet med bleka omättade färger. Eftersom AdobeRGB täcker en större färgrymd, så är det möjligt att låta alla färger som ligger utanför röda linjerna i Fig. 12.13 projiceras ner på röda linjerna. Däremot förblir alla färger innanför röda linjerna oförändrade. Det finns också andra metoder för att omvandla AdobeRGB till sRGB, men vi går inte in på det här. Eftersom information går förlorad när man omvandlar AdobeRGB till sRGB, så kan man inte efteråt komma tillbaka till AdobeRGB.

När kan då AdobeRGB vara av praktisk nytta? Jo, det är framförallt om man vill printa ut sina bilder på en riktigt högkvalitativ printer. Dessa kan ofta återge mer mättade färger än vad en bildskärm kan göra. Men det gäller då att se till att printern verkligen är inställd för AdobeRGB, eftersom det i de flesta sammanhang antas att bildmaterial använder sRGB.

En slutsats angående val av färgrymd skulle kunna vara att sRGB är ett säkert kort. Den ger bra färgåtergivning i alla sammanhang, men kanske inte den allra yppersta i avancerade tillämpningar (speciellt printning). AdobeRGB kan vara ett bättre val om man ska framställa utskrifter av yppersta kvalitet. Men för sådana mer avancerade tillämpningar är det antagligen bättre att fotografera i RAW-format, eftersom man då kan framställa ett godtyckligt antal jpeg-bilder med olika färgrymder ur samma RAW-bild*. Att arbeta med bilder i RAW-format möjliggör också mycket mer avancerad bildbehandling än om man låter kameran själv tillverka sina jpeg-bilder.

Förutom att styra hur bildens färger återges, så innehåller färgrymderna sRGB och AdobeRGB även en specifikation av hur bildens gammakorrektion ska utföras. Gammakorrektion utgör en icke-linjär skalning av bildens pixelvärden, och beskrivs mer utförligt i kap. 14.6. Syftet med gammakorrektion är att på ett bra sätt kunna återge gråtoner och färger trots att man i jpeg-bilder bara har 256 nivåer i varje delfärg.

Förutom sRGB och AdobeRGB finns ytterligare ett antal färgrymder. Men dessa används mest i speciella, professionella tillämpningar.

* En RAW-bild innehåller typiskt mer färginformation än vad som ryms i både AdobeRGB och sRGB.

13. Ljuskällor

13.1. Inledning

Vi ska i detta kapitel titta lite på olika ljuskällor som används inom fotografi, samt deras spektrala egenskaper. Ljuskällorna kan indelas i tre huvudkategorier:

1. Temperaturstrålare (svartkropsstrålare)
2. Urladdningslampor
3. LED-lampor

13.2. Begreppen svartkropsstrålare och färgtemperatur

Glödlampor, solen och spisplattor är strålningskällor som sänder ut ett spektrum som approximativt överensstämmer med det från en ideal svartkropsstrålare. För sådana strålkällor gäller att den totalt utstrålade effekten ökar med temperaturen. Vid ökande temperatur kommer också emissionsmaximum att förskjutas mot kortare våglängder. Denna spektralförskjutning ger en färgförändring hos ljuset, ett faktum som är speciellt viktigt vid färgfotografi. Vi ser i Fig. 13.1 att vid lägre temperaturer så är utstrålningen inom det synliga spektralområdet (ca. $0.4 - 0.7 \mu\text{m}$) låg. Det är först vid temperaturer på 5000-6000 K som vi får ett emissionsmaximum ungefär mitt i det synliga området.

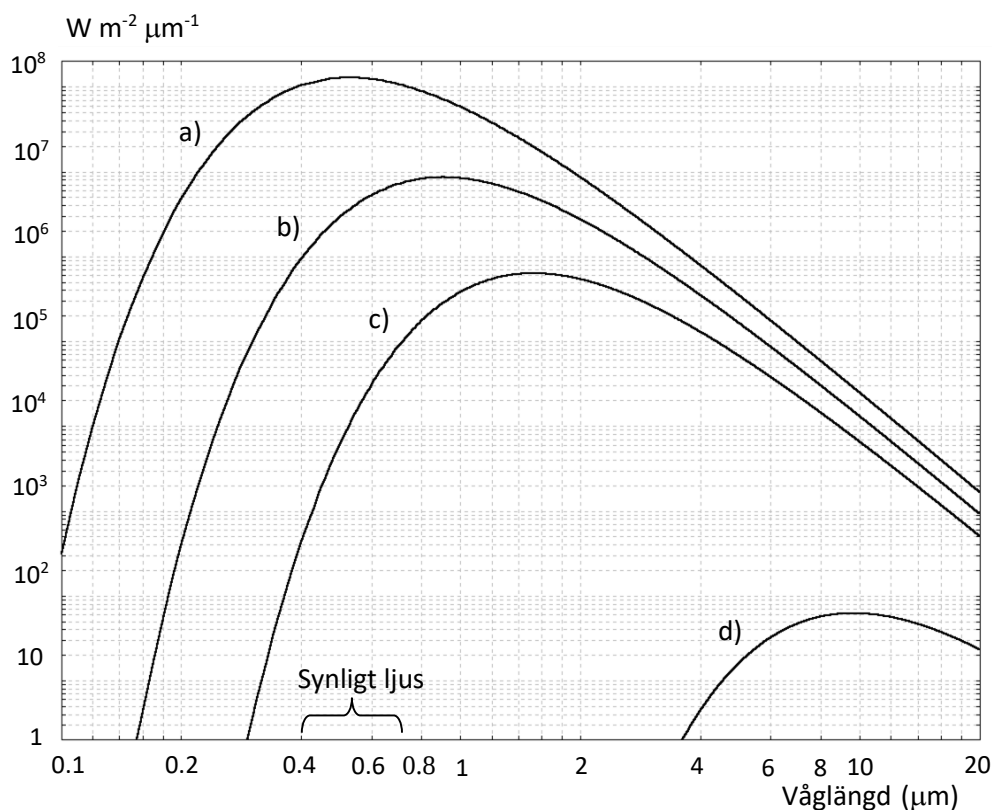


Fig. 13.1. Spektrala emissionskurvor för svartkropsstrålare med olika temperaturer. a) 5500 K (dagsljus), b) 3200 K (fotolampor), c) 1900 K (stearinljus), d) 300K (rumstemperatur)

För att karakterisera en ljuskällas spektralfördelning används begreppet **färgtemperatur**. Om färgtemperaturen anges till t.ex. 5500 K (dagsljus), betyder det att ljuskällans spektralfördelning överensstämmer (åtminstone approximativt) med spektralfördelningen hos en svartkroppsstrålare med den temperaturen. Ljuskällan behöver dock inte ha just denna temperatur, eftersom spektralfördelningen kan påverkas med färgfilter. I digitalkameror ställs färgtemperaturen (även kallad vitbalansen) in på lämpligt värde vid fotograferingen, och även finjusteringar kan ofta göras. Se vidare kap. 14.8. Man ska dock ha klart för sig att färgsammansättningen hos dagsljus kan variera en hel del med tiden på dagen, årstid och atmosfäriska förhållanden (t.ex. dis). Ljusabsorption i jordatmosfären kommer också att påverka spektralfördelningen för dagsljus, se Fig. 13.2 som visar ett uppmätt dagsljusspektrum en solig sommardag.

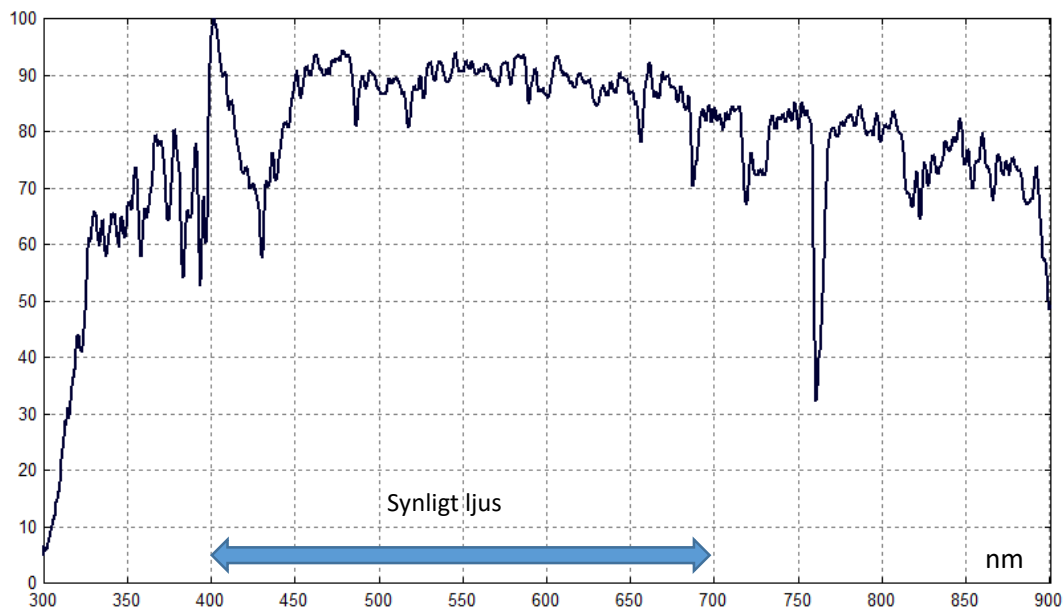


Fig. 13.2. Dagsljusspektrum en solig sommardag i Stockholm. Jämfört med kurva a) i Fig. 13.1 är denna kurva ojämnare, men i grova drag är intensitetsfördelningen mellan de blå, gröna och röda våglängdsområdena densamma. Ojämnheterna i kurvan beror huvudsakligen på ljusabsorption i jordens atmosfär (även solens atmosfär ger en liten påverkan). Den kraftiga dippen vid 760 nm beror sålunda på jordatmosfärens syre (O_2), och de låga värdena när man närmar sig 300 nm beror på ozon (O_3) i jordatmosfären.

En ljuskälla med en färgtemperatur på ungefär 5500 K kan i fotosammanhang anses som ganska ideal. Den kommer att ge en mycket jämn spektralfördelning inom det synliga området (Fig. 13.2, om man bortser från lite ”brus”), och levererar därför ett nästan perfekt vitt ljus*. Detta gör att motivets samtliga färger registreras på ett bra sätt. En ljuskälla med låg färgtemperatur, t.ex. 3200 K (Fig. 13.3), kommer att ge ett gulaktigt ljus som gör att blåa föremål avbildas alltför mörka. Genom att ställa in kamerans vitbalans på 3200 K kommer de blå färgerna att förstärkas i bilden, men den förstärkningen medför också ett ökat brus i bildens blå partier. Motsvarande problematik uppstår vid höga färgtemperaturer. Då erhålls ett

* För en svartkroppsstrålare med temperaturen 5500 K varierar intensiteten mindre än $\pm 10\%$ inom våglängdsintervallet 400-700 nm.

blåaktigt ljus som gör att röda färger behöver förstärkas i bilden, vilket återigen ger ökat brus.

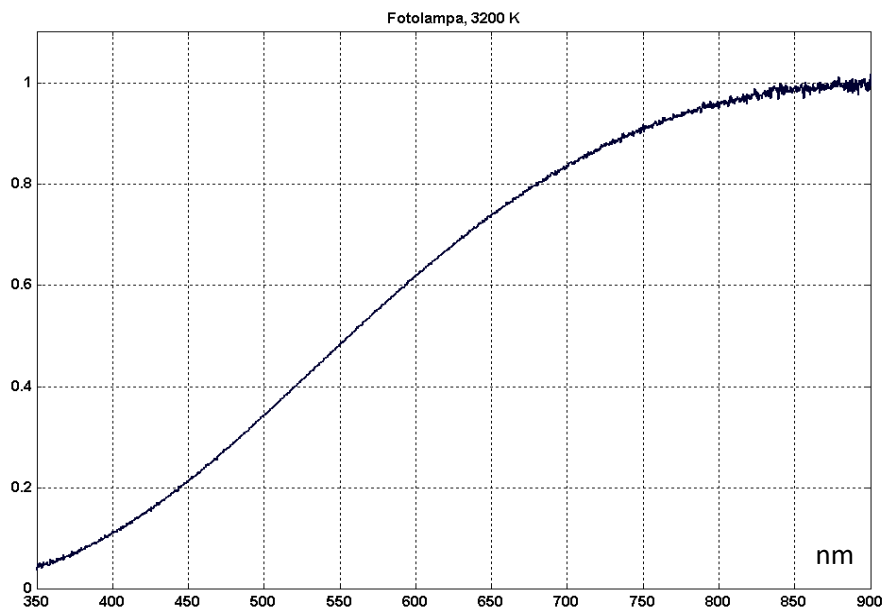


Fig. 13.3. Uppmätt spektralkurva för en 3200 K fotolampa (glödlampa). Notera den ojämna spektralfördelningen med lite blått och mycket rött ljus. Detta måste kompenseras i kameran, eller i datorn vid efterbehandlingen, för att inte bilderna ska få en rödgul färgton.

För att mäta upp färgtemperaturen hos en ljuskälla kan man använda en så kallad färgtemperaturmätare. Denna fungerar vanligen så att den mäter upp hur hög belysning man har inom två eller tre olika spektralband (t.ex. blått, grönt och rött). Sedan jämförs storleken på dessa värden, och utgående från denna jämförelse beräknas en färgtemperatur. Detta fungerar bra för ljuskällor som är approximativt svartkroppsstrålare, men för andra ljuskällor blir värdena osäkra eller meningslösa.

Under senare år har energisparlampor och LED-lampor (se avsnitt 13.3 och 13.4) i stor utsträckning ersatt glödlampor. Det innebär att inomhusbelysningen idag har en spektralfördelning som ofta skiljer sig markant från en svartkroppsstrålare (Fig. 13.5 och 13.6). Detta innebär att man strängt taget inte kan tala om färgtemperatur för dessa nyare ljuskällor. Trots det kan man på förpackningarna för lysrör, energisparlampor och LED-lampor ofta se en färgtemperatur angiven (ofta till ca. 2700 K). Det man anger i dessa fall är egentligen inte färgtemperaturen, utan något som kallas "korrelerade färgtemperaturen". Detta är den temperatur som en svartkroppsstrålare ska ha för att dess färg rent visuellt ska matcha ljuskällans färg så bra som möjligt. Ofta blir dock färgmatchningen långt ifrån perfekt. Det bör också betonas att även en god visuell färgmatchning inte alltid innebär att de fotografiska resultaten blir desamma för den aktuella ljuskällan och svartkroppsstrålaren. Fig. 13.4 visar hur man i praktiken bestämmer korrelerade färgtemperaturen. Det man gör är att bestämma temperaturen för den svartkroppsstrålare som i CIE-färgdiagrammet (kap. 12.5) ligger närmast den aktuella ljuskällan.

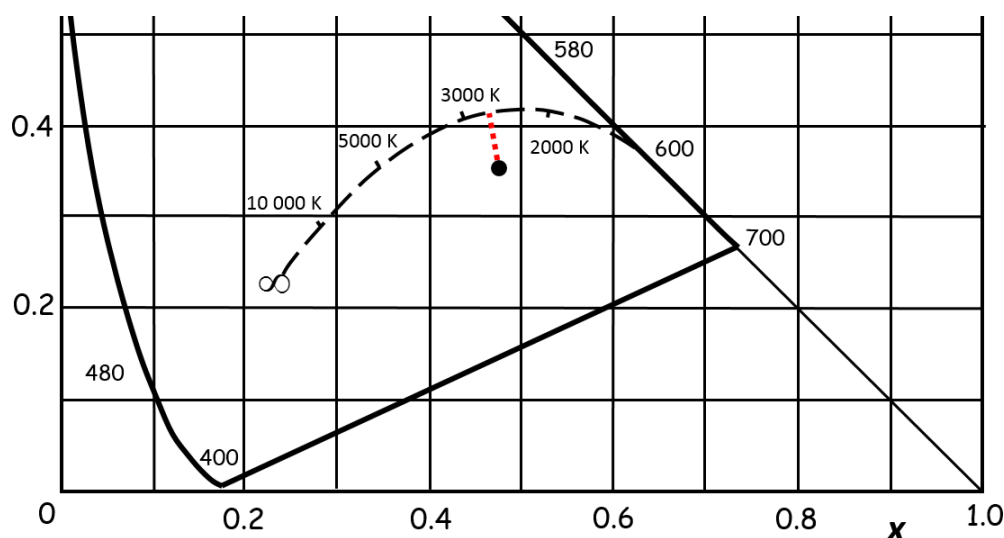


Fig. 13.4. Principen för bestämning av korrelerad färgtemperatur. Figuren visar ett delområde av Fig. 12.10. CIE-koordinaterna för svartkroppsstrålare med olika temperaturer anges av den streckade kurvan. Svarta prickerna är en tänkt ljuskälla som inte är en svartkroppsstrålare. För att hitta dess korrelerade färgtemperatur drar man en linje kortaste vägen till den streckade svartkroppskurvan (prickad röd linje). Den temperatur man hamnar på då utgör den korrelerade färgtemperaturen, i detta fall ca. 2700 K. Om prickade linjen är längre än 0.05 enheter i x - eller y -led, brukar man anse att det inte är meningsfullt att tala om en korrelerad färgtemperatur.

Oavsett om man kan tilldela en ljuskälla en korrelerad färgtemperatur eller inte, så kan man tilldela den ett så kallat "color rendering index", CRI, som mått på färgåtergivning. Detta index baseras på ljuskällans spektralfördelning, vilken jämförs med spektralfördelningen hos en svartkroppsstrålare. Svartkroppsstrålare får sålunda högsta möjliga värde, 100, medan andra ljuskällor får värden mindre än 100. CRI är inget pålitligt mått på en ljuskällas lämplighet för färgfotografering, men ger trots detta en viss vägledning. CRI-värden i intervallet 80-100 är vanliga för lampor i hemmiljö, och resulterar ofta i bilder av god kvalitet. Lampor gjorda speciellt för fotoändamål har ofta CRI i intervallet 95-100, medan gatubelysning ofta ligger på 0-50 (natrium lågtryckslampor kan till och med ge negativa värden!).

13.3. Urladdningslampor

De vanligaste urladdningslamporna är olika typer av lysrör och energisparlampor. Energisparlampor kallas även lågenergilampor, och är i princip små lysrör. I dessa lampor har man en elektrisk urladdning i förtunnad kvicksilverånga. Vissa lamptyper (t.ex. gatubelysning) använder också elektriska urladdningar i natriumånga. I Fig. 13.5 visas exempel på emissionsspektrum från ett lysrör.

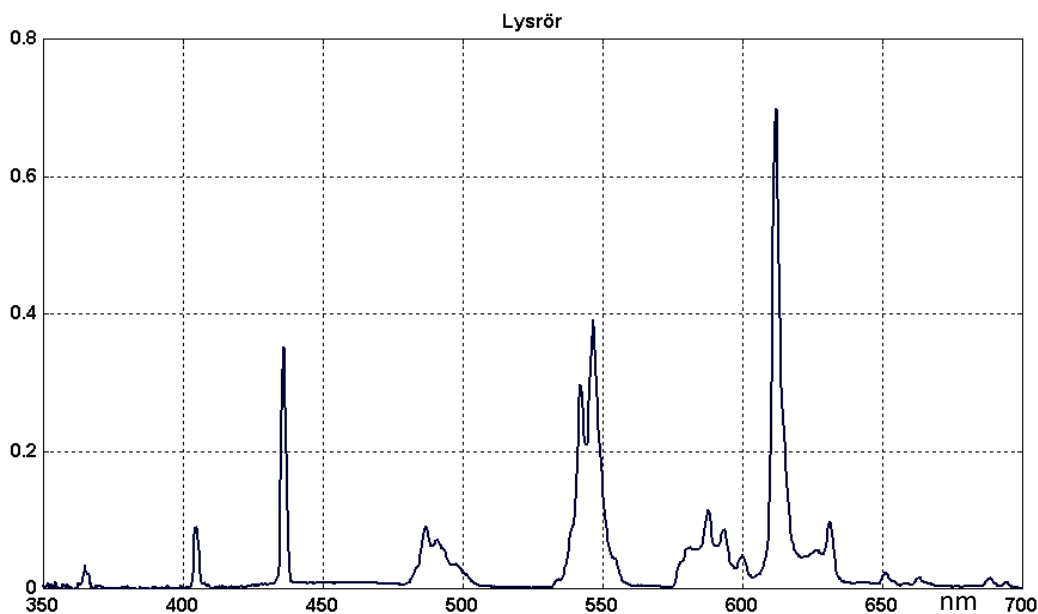


Fig. 13.5. Emissionskurva för lysrör som används till inomhusbelysning i kontor och liknande. Ett antal spektrallinjer från kvicksilver syns, plus det spektrum som ges av det fluorescerande lyspulvret som finns på glaskolvens insida. Lyspulvret omvandlar ultraviolett strålning till synligt ljus, plus att det bidrar till en behagligare färg på ljuset.

Urladdningslampor bör användas med viss försiktighet vid färgfotografi. Gatubelysning med kvicksilver högtryckslampor ger t.ex. ofta en sjukligt grön färgton beroende på kvicksilvers starka emission vid 546 nm. (Det är denna spektrallinje som gör att grönskan ser så fräsch ut i gatlampans kvicksilversken.) Högtrycksnatriumlampor, som också ofta används till gatubelysning, saknar nästan helt våglängder under 550 nm, vilket gör att blåa och gröna färger försvinner i bilderna. Vanliga lysrör och energisparlampor ger emellertid ofta ett bra resultat trots att spektralfördelningen kan se ut som i Fig. 13.5. Lysrör finns dock i ett flertal varianter med olika (korrelerad) färgtemperatur. I hemmiljö används ofta en låg färgtemperatur (2700 K), medan man på arbetsplatser och i offentliga miljöer ofta har en högre färgtemperatur (t.ex. 4000 K). Därför kan man behöva experimentera lite med olika vitbalansinställningar vid fotografering i lysrörljus.

13.4. LED-lampor (Light Emitting Diode)

LED-lampor har, tillsammans med urladdningslampor, ersatt glödlampor i många sammanhang. Vinsten man gör är att dessa bägge lamptyper förbrukar mycket mindre energi än glödlampor. För att ange en ljuskällas effektivitet brukar man ange det så kallade ljusutbytet, dvs hur många lumen man får per watt inmatad elektrisk effekt. En glödlampa kan typiskt ge ca. 15 lm/w, lysrör eller energisparlampor ca. 60 lm/w och LED-lampor ca. 80 lm/w. Gatubelysning med urladdningslampor kan ge ett ljusutbyte över 100 lm/w. Det medför alltså stora energivinster att ersätta glödlampor med andra lamptyper. Som tidigare omtalats så kan urladdningslampor i t.ex. gatubelysning vara olämpliga (eller i varje fall oberäkneliga) vid färgfotografi på grund av ljusets spektralfördelning. Hur är det då med LED-lampor?

En LED-lampa består av ett antal lysdioder som utsänder ljus inom ett ganska smalt spektralområde runt 450 nm, dvs blått ljus. Detta blå ljus absorberas av ett fluorescerande ämne som omger lysdioderna, och som därvid utsänder ett brett kontinuerligt spektrum med ett maximum som ofta ligger runt 600 nm. I Fig. 13.6 visas spektralfördelningarna för ett lysrör och en LED-lampa. Energisparlampor har ett spektrum som liknar lysrör. I Fig. 13.7 illustreras hur lite av glödlampans utstrålning som sker i det synliga området.

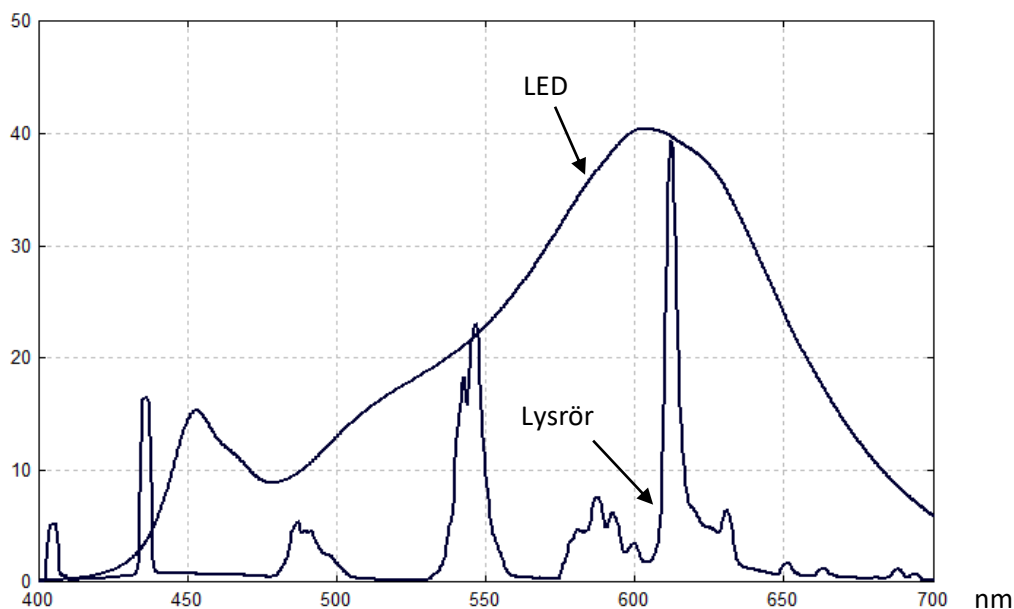


Fig. 13.6. Emissionskurvor för LED-lampa och lysrör. På grund av LED-lampans mer kontinuerliga spektralfördelning är den lämpligare för färgfotografering än lysröret (inga färger saknas i spektrum).

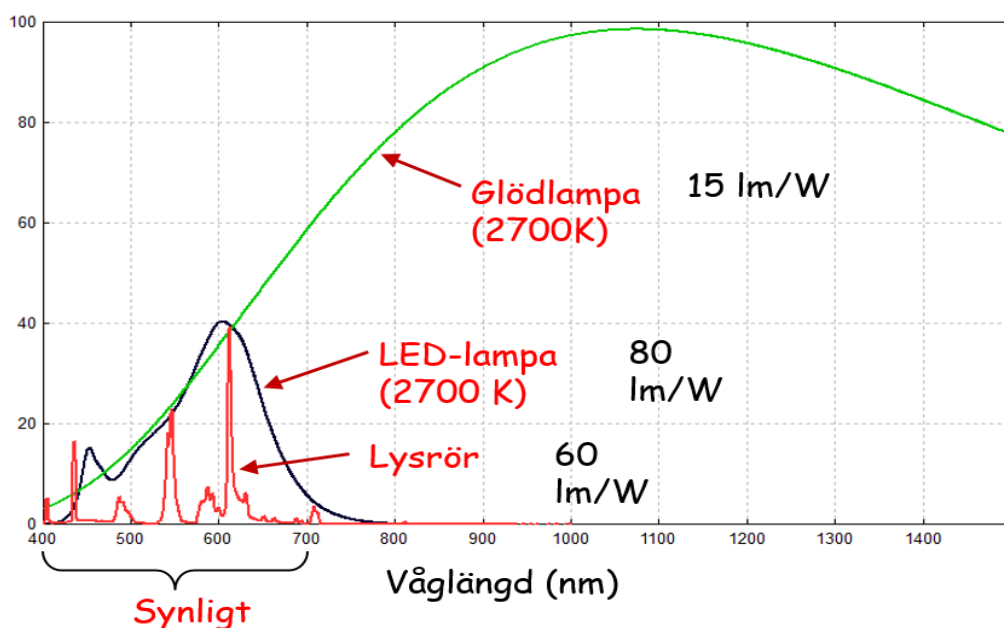


Fig. 13.7. Emissionskurva för glödlampa jämfört med LED-lampa och lysrör. Huvuddelen av glödlampans utstrålning sker utanför det synliga området i form av värmestrålning, vilket ger ett lågt ljusutbyte.

LED-lampans spektrum är mycket mer likt glödlampsljus (Fig. 13.3) än vad fallet är för lysrör. Detta innebär mindre risk för konstiga färger vid användande av LED-lampor. Men som framgår av Fig. 13.6 har LED-lampor en puckel i blåa delen av spektrumet, vilket beror på att lysdiodernas verkliga våglängd inte helt absorberas av det fluorescerande skiktet utan lyser igenom lite. Dessutom innehåller LED-ljuset lite mindre rött än vad fallet är med glödlampsljus. Olika lampfabrikat kan säkerligen skilja sig åt en del vad gäller dessa detaljer, även om den angivna färgtemperaturen är densamma (oftast ca. 2700 K). Man kan därför behöva göra små justeringar av vitbalansen i kameran för att få en optimal färgåtergivning.

Det förtjänar också att påpekas att det finns LED-lampor och lysrör/energiparlampor som är speciellt framtagna för att användas i en fotostudio. I dessa lampor har man i regel en bättre färgkvalitet på ljuset (t.ex. genom förbättrade fluorescensämnen) vilket gör det lättare att få färgbilder med bra kvalitet. Färgtemperaturen på denna typ av fotolampor ligger i allmänhet på ca. 5500 K (dvs de har dagsljuskaraktär).

13.5. Fotoblixt

Urladdning i xenon-gas under högt tryck och hög temperatur ger nästan svartkroppsstrålning, Fig. 13.8. Detta utnyttjas i blixtaggregat, och fungerar väl för både svartvit och färgfotografering. Färgtemperaturen i en fotoblixt varierar lite mellan olika blixtaggregat. Den varierar också för ett och samma blixtaggregat beroende på hur hög energi man tar ut ur blixten när bilden tas. Vanligen ligger färgtemperaturen på ca. 5500 K, dvs den motsvarar dagsljus. Vid ateljéfotografering används ofta en eller flera blixhtar som ljuskälla. Kamerainställningarna innan exponeringen sker vanligen i glödlampsljus. Många kameror har inbyggd blixt, men av utrymmesskal blir det små blixtaggregat med ganska begränsad räckvidd.

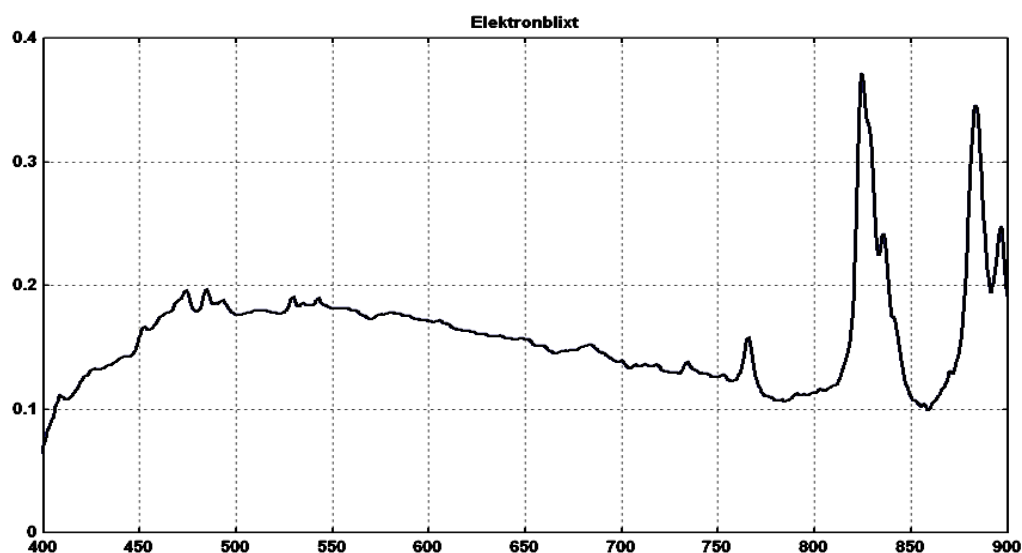


Fig. 13.8. Emissionskurva för en fotoblixt. Den ger mer svartkroppsliknande strålning än övriga urladdningslampor. De kraftiga spektrallinjerna över 800 nm är inte störande, eftersom vanliga kameror inte är känsliga i detta våglängdsområde.

Innan blixten är klar att avfyras måste en kondensator laddas upp till hög spänning. Detta kan ta flera sekunder med hjälp av de relativt små batterier som oftast används. Detta begränsar i hur snabb följd blyxtbilder kan tas. Utlösningen av blixten synkroniseras till kamerans slutare med hjälp av en kabel som ansluts till kamerahuset (eller via en kontakt i blyxtsockeln). Urladdningen i blyxtröret startas med en högspänningspuls som gör gasen ledande, varefter den i kondensatorn upplagrade energin urladdas genom gasen. Urladdningstiden är ofta ca. 1 ms. För att blyxtsynkroniseringen ska fungera bra tillsammans med ridåslutare får emellertid inte slutartiden vara för kort (se kap. 4.3). Den kortaste tillåtna tiden ligger ofta runt 1/250 s. Om en för kort exponeringstid används hinner inte ridåslutaren med, och den kommer därför att skymma delar av sensorn vid exponeringen. Tanken vid blyxtexponeringen är att ridåerna ska vara helt öppna innan blixten tänds. Exponeringstiden bestäms därmed helt och hållet av blyxtens brinntid. Många digitalkameror har en spärr som gör att man inte kan ställa in för kort tid när blixten är aktiverad. Det finns också ”intelligenta” blyxtar som kan användas även vid de allra kortaste slutartiderna för en ridåslutare. Istället för en enda blyxt kommer då ett helt tåg av blyxtar under hela den tid som den smala slutarspringan rör sig fram över sensorn.

13.6. Kontroll av vitbalansen i en bild

Som tidigare nämnts i detta kapitel kan vitbalansinställningen på kameran behöva justeras för att få god färgkvalitet i bilderna. Hur ska man då ”ute i fält” avgöra om färgerna återges bra eller inte, och hur ska man veta åt vilket håll en eventuell justering ska göras? En grov vägledning får man naturligtvis genom att titta på hur färgerna återges på kamerans display. Detta är användbart för en första grovjustering. Men för att få mera noggrann information är det bra att studera RGB-histogrammen som kan visas i många kameror (se kap. 11.3 angående histogram).

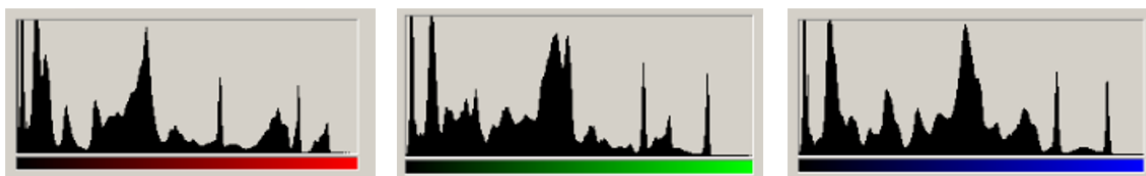


Fig. 13.9. Exempel på histogram för en bilds röda, gröna och blå färgkanaler (i den ordningen från vänster till höger). I detta fall är vitbalansen korrekt inställd, vilket ger en korrekt exponering i samtliga kanaler (jämför Fig. 11.2)

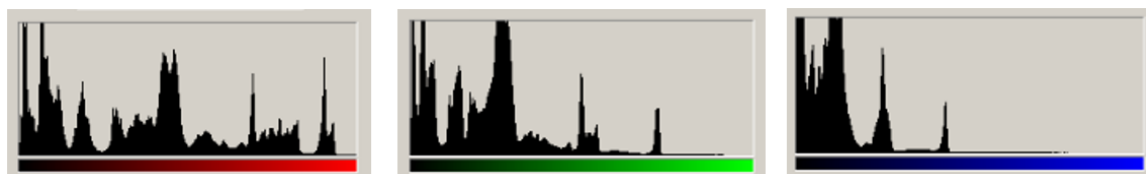


Fig. 13.10. RGB-histogram för bild av samma motiv som i Fig. 13.9, men i detta fall är vitbalansen felaktigt inställd. Rödkanalen är korrekt exponerad, medan grönkanalen är lite underexponerad och blåkanalen är mycket underexponerad.

Bilderna vars histogram visas i Fig. 13.9 och 13.10 togs med en glödlampa som ljuskälla (färgtemperatur ca. 2700 K). I ena fallet, illustrerat i Fig. 13.9, ställdes kamerans vitbalans på 2700 K, medan i andra fallet, Fig. 13.10, vitbalansen var ställd på 5500 K. Korrekt vitbalans är inställd när RGB-histogrammen har ungefär samma utsträckning i horisontell riktning. Men i detta sammanhang kan en liten varning vara på plats. Histogrammens utseende och horisontella utsträckning beror naturligtvis på motivets färginnehåll. Om i stort sett hela motivet har en orange-röd färgton så skulle histogrammet för en korrekt bild se ut som i Fig. 13.10. I normalfallet så finns dock oftast partier i motivet som är ljust gråa eller vita, och då kommer en inställning som ger histogram enligt Fig. 13.9 att korrekt återge motivets färger.

Även om vitbalansen justeras såsom beskrivits, kan det hända att mindre justeringar behöver göras i efterhand då bilderna behandlas i en dator. Det är då till stor hjälp om man någonstans i bilden har ytor som man vet är helt neutralgrå i verkligheten. Eftersom sådana ytor inte alltid finns naturligt i ett motiv, så lägger professionella fotografer ofta in en kalibrerad grå- och färgskala någonstans i motivet (t.ex. i bildkanten så att den lätt kan beskäras bort i den färdigbehandlade bilden). Vid efterbehandlingen av bilden kan man då justera RGB-värdena så att grå- och färgskalan återges så naturtroget som möjligt. Detta brukar gå mycket bra om det är fråga om mindre justeringar av vitbalansen. Men om det fordras stora justeringar blir bildkvalitén lidande. Man kan råka ut för ett ökat brus i bilderna, samt att antalet nivåer som kan utnyttjas i någon av färgkanalerna är för litet. Det sistnämnda problemet begränsar antalet färgtoner som kan återges. Ett annat problem man kan råka ut för vid felaktig vitbalansinställning, är att någon färgkanal kan vara bottnad (överexponerad) trots att histogrammet före exponeringen ser helt optimalt ut. Detta beror på att histogrammet som visas före exponeringen vanligen är luminanshistogrammet, vilket är ett medelvärde av de tre färgkanalerna. Om en kanal är överexponerad, och en annan underexponerad, så kan medelvärdet se bra ut och därmed vara vilseledande. Som vi ser är det därför viktigt att försöka få vitbalansen så bra som möjligt redan vid fotograferingen.

Det som sagts ovan gäller vid fotografering i de vanligt förekommande filformaten jpeg och tiff. Vid fotografering i RAW-format sker all behandling av bilddata efteråt på en dator, och med betydligt större bit-djup än i en jpeg-bild (se kap. 14.7). Vitbalansinställningen vid fotograferingen spelar därmed ingen roll. Men eftersom vitbalansinställningen påverkar bilden som visas på LCD-skärmen och i en elektronisk sökare, så kan det vara bra att ha en ungefär korrekt vitbalansinställning för att kunna bedöma hur slutbilden kommer att se ut. Vid fotografering i RAW-format är det också mindre risk för felexponering. Om man exponerar efter luminanshistogrammet som visas på LCD-skärmen före exponering brukar det bli rätt, oberoende av inställd vitbalans. Vid ändring av vitbalans förskjuts visserligen blå- och rödhistogrammen, men åt motsatta håll så att medelvärdet ligger ganska konstant. Och grönhistogrammet påverkas bara marginellt av vitbalansinställningen. Resultatet blir att medelvärdet av de tre histogrammen påverkas mycket lite av vitbalansinställningen. Anledningen till att man kan få problem vid jpeg-fotografering, är att beroende på vitbalansinställning lägger man på olika digital förstärkning eller dämpning av pixelvärdena i kameran efter exponeringen. Då kan det hända att någon färgkanal bottnar, trots att det ursprungliga pixelvärdet inte var överexponerat.

14. Matrissensorer för digital fotografi

14.1 Digitalkameran

I digitala kameror används en elektronisk sensor istället för som tidigare fotografisk film. Detta innebär många fördelar.

- Man har inga kostnader för film och framkallning.
- Misslyckade bilder kan raderas och minnesutrymmet återanvändas.
- Man ser resultatet omedelbart.
- Det är lätt att bearbeta och förbättra sina bilder i en dator. Utskrifter i hög kvalitet kan framställas på bläckstråle- eller laserskrivare.
- Bilderna är i digital form och kan publiceras direkt på internet, eller användas i datorframställda dokument (som detta kompendium)

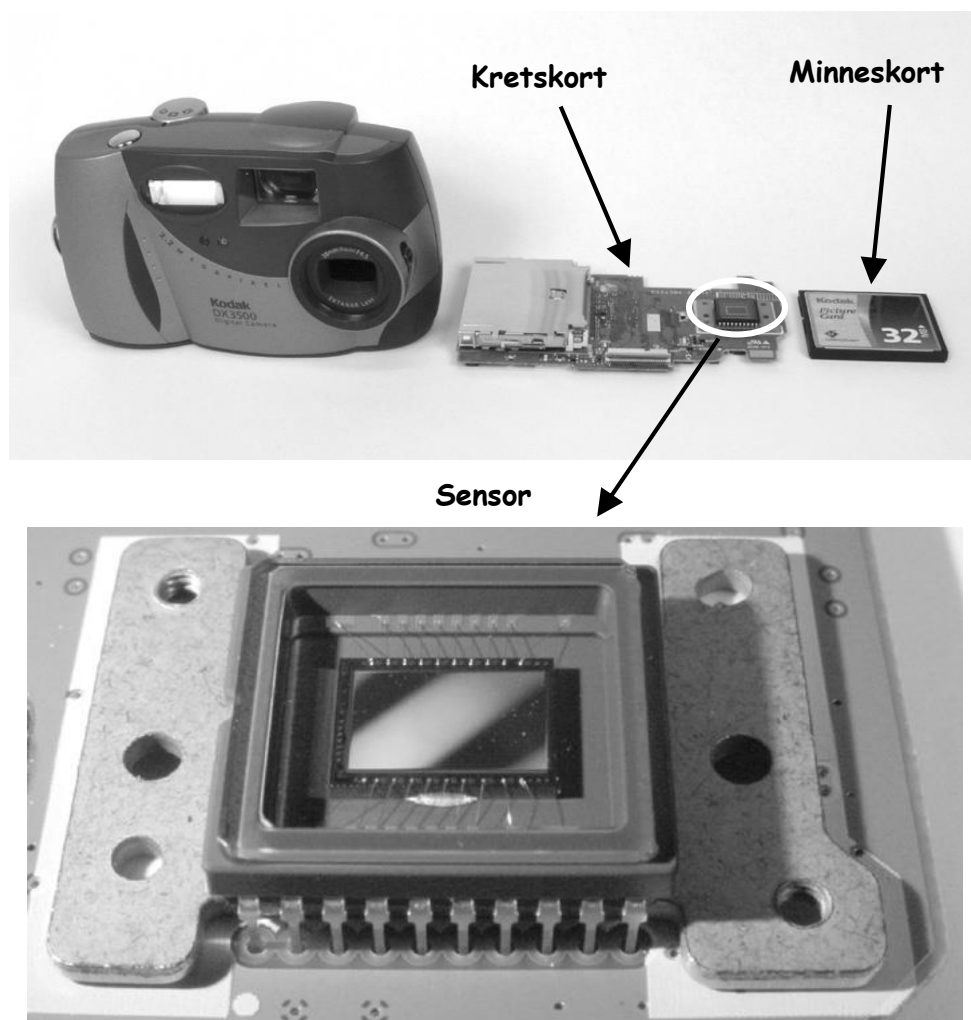


Fig. 14.1. Exempel på uppbyggnaden av en enkel digitalkamera. Sensorn är i detta fall en CCD med en ljuskänslig yta av 4 mm x 6 mm. I mer avancerade kameror är sensorn större. I professionella sammanhang används ofta sensorer som är 24 mm x 36 mm eller ännu större. Antalet pixlar ligger typiskt på 10-20 miljoner i enklare kameror, medan proffskameror kan ha över 50 miljoner pixlar.

Sensorer i digitalkameror varierar kraftigt i storlek. Från ca. 4 x 5 mm upp till ca. 40 x 50 mm. Antalet pixlar (detektorelement) varierar mellan ca. 10 miljoner och 60 miljoner. (Detta är läget i oktober 2019. Situationen förändras hela tiden på grund av den snabba utvecklingen.) I nedanstående tabell ges en översikt över några vanliga sensorstorlekar.

Mobilkameror	ca. 4 x 5 mm	ca. 10 mpixlar
1/3.1"	3.5 x 4.7 mm	10-15 mpixlar
1/2.3"	4.55 x 6.17 mm	ca. 20 mpixlar
Four Thirds System	13.0 x 17.3 mm	typ. 16 mpixlar
APS-C	15.6 x 23.5 mm	upp till 24 mpixlar
Fullformat (FX)	24 x 36 mm	upp till ca. 60 mpixlar
Mellanformat (Hasselblad)	36.7 x 49.1 mm	50 mpixlar

Det förtjänar att påpekas redan nu att bildkvalitén som en sensor kan leverera beror i hög utsträckning på storleken, och inte bara på antalet pixlar. En sensor med stor area kommer att ge bilder som är skarpare och innehåller mindre brus. Det lägre bruset innebär att kameran kan användas under sämre ljusförhållanden (genom att öka ISO-talet). Dessutom kommer en sensor med stor area att klara av att fotografera motiv med stort kontrastomfång, utan att man förlorar detaljer i de ljusaste och mörkaste partierna. Dessa saker kommer att beröras närmare i detta kapitel.

Sensorer i digitalkameror finns av två olika typer, CCD och CMOS. Från början dominerade CCD-sensorer på grund av att de gav den bästa bildkvalitén. Utvecklingen har emellertid gjort att CMOS-sensorens kvalitet har ökat kraftigt, samt att man kan tillverka dessa sensorer i stora storlekar till ett förhållandevis lågt pris. Detta har gjort att CMOS under årens lopp har tagit över allt mer av marknaden för sensorer, i alla fall vad gäller konsumentprodukter. En annan fördel med CMOS är att kretsarna är snabba och att man kan läsa ut bara en delmängd av pixelantalet. Detta gör att digitalkameror med CMOS också kan användas för att spela in video i HD-format (1920 x 1080 pixlar, eller ännu högre). CMOS-kretsar har vidare en lägre strömförbrukning, vilket naturligtvis är en fördel i batteridriven utrustning. Det enda som egentligen talar för CCD-sensorn i dagsläget är att den för en given sensorstorlek ger en något bättre bildkvalitet, vilket kan vara en fördel om kamerans dimensioner måste hållas nere. I vetenskapliga tillämpningar och vid professionell fotografering, där man har mycket högt ställda krav på bildkvalitet, kan CCD också vara det bästa valet. Uppbyggnaden och funktionen hos CCD- och CMOS-sensorer kommer att beskrivas senare i detta kapitel.

Det är lätt att bli överväldigad av alla möjligheter som datorn ger till "bildförbättringar", och man faller lätt för reklamens budskap att "endast fantasin sätter begränsningar". Detta är naturligtvis nonsens. Precis som i klassisk fotografi med film så är det viktiga vilka grundprestanda vad gäller upplösning, brus och dynamik som kamerasystemet har. Brister i dessa avseenden kan aldrig skylas över av aldrig så mycket datorbearbetningar. Därför ska vi i detta och kommande kapitel tala en hel del om olika mått på bildkvalitet, och hur den kan mätas.

Optiken i en digitalkamera skiljer sig inte på något väsentligt sätt från optiken i en kamera för film. Största skillnaden är att man i regel arbetar med kortare

brännvidder. Detta beror på att matrissensorerna i många fall är mindre till storleken än en filmruta*. En kortare brännvidd påverkar skärpedjupet (som blir större), samt hur höga bländartal som kan användas innan diffraktionen blir störande (gränsen flyttas nedåt). Dessa saker behandlas i kapitlen om skärpedjup och upplösningsförmåga. I detta kapitel ska vi koncentrera oss på sensorn i en digitalkamera. Denna består av en integrerad krets tillverkad med kiselteknologi. Som redan nämnts finns två varianter av dessa sensorer, vilka går under benämningarna CCD och CMOS. CCD står för "charge-coupled device" och CMOS för "complementary metal oxide semiconductor". Största skillnaden mellan dessa två typer handlar om hur pixeldata läses ut. Vi ska börja med att beskriva CCD-sensorn.

14.2 CCD-sensorn

Om man tittar på en CCD-sensor i hög förstoring ser man att den består av en regelbunden matris av små ljuskänsliga element (pixlar). Centrum-till-centrum avståndet mellan pixlarna varierar mellan olika CCD-sensorer, men typiskt ligger det i intervallet 1 – 5 μm . Idealt skulle pixlarna ligga kant i kant så att hela sensorytan vore ljuskänslig. Detta är inte möjligt av praktiska skäl, mer om detta senare. Man brukar därför tala om "fill factor", vilken anger hur stor procentuell andel av sensorytan som är ljuskänslig. I värsta fall kan fill factor vara så låg som 20%, vilket innebär att ljuskänsligheten blir låg. Ett knep som ofta används för att öka på fill factor, är att lägga en liten lins strax ovanför varje pixel, Fig. 14.2. På så sätt kommer ljuset att koncentreras på det ljuskänsliga elementet, och de okänsliga mellanrummen träffas bara av ett fåtal fotoner. Detta gör att fill factor typiskt ökar från ca. 30% till ca. 70%.

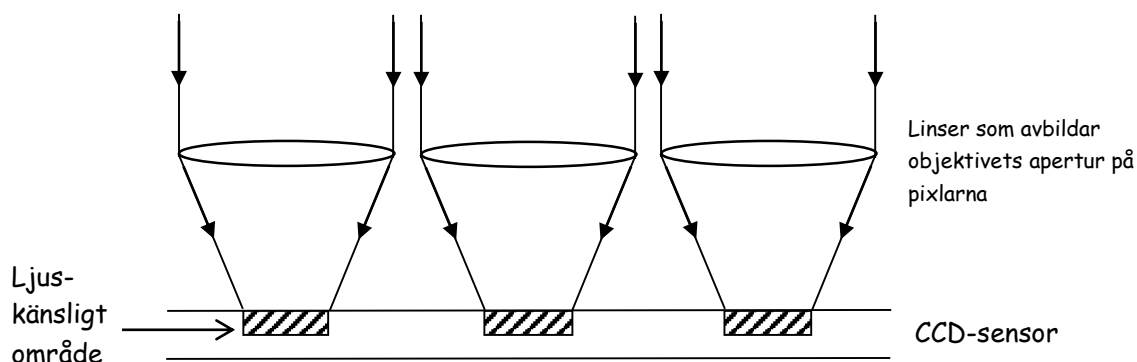


Fig. 14.2. Genom att montera mikrolinser ovanför detektorelementen kan "fill factor" ökas kraftigt, vilket medför att ljuset utnyttjas bättre.

En pixel kan bestå av en ljuskänslig diod eller en så kallad "photogate". Den grundläggande principen för ljusdetektion är densamma i de bägge fallen. Vi ska illustrera funktionen genom att titta på hur en photogate detekterar fotoner, samt

* I teknisk litteratur anges sensorstorleken ofta med ett mått i tum, t.ex. $\frac{1}{2}$ " för en "halvtumssensor". Detta mått anger inte, såsom man skulle kunna tro, diagonalmåttet för den ljuskänsliga ytan. Av dumma historiska skäl är verkliga diagonalmåttet för sensorn istället ca. 2/3 av tumangivelsen!

hur signalen läses ut. I Fig. 14.3 visas en förenklad tvärsnittsbild av en photogate. Uppifrån och ner består den av en tunn elektrod (ledande material), oxidlager (isolator) och p-dopat kisel (halvledare).

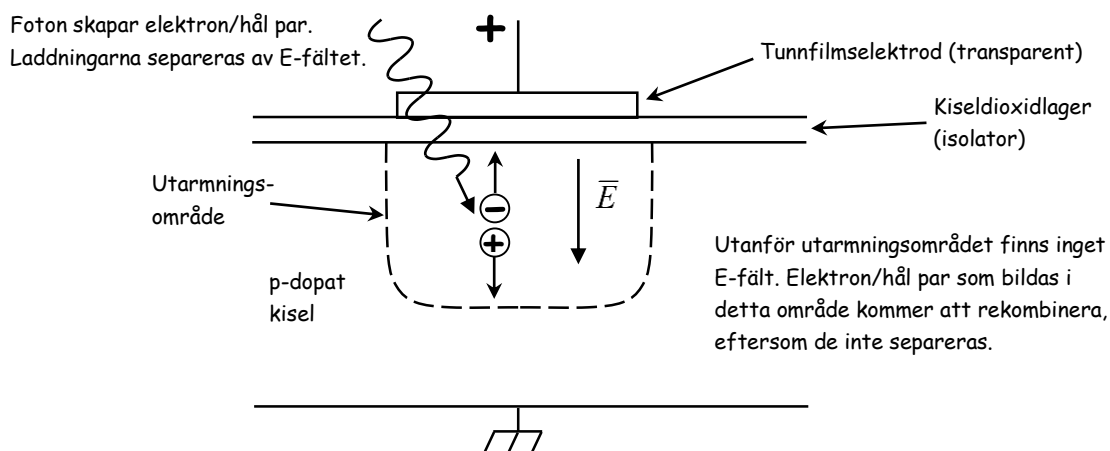


Fig. 14.3. Tvärsnittsbild genom en "photogate." Funktionen beskrivs i texten.

När man lägger en positiv spänning på elektroden kommer de positiva laddningsbärarna, dvs. hålen, att repelleras. Detta skapar ett så kallat utarmningsområde (eng. "depletion layer") under elektroden där inga rörliga laddningsbärare finns kvar. I detta område finns ett elektriskt fält med en fältstyrkevektor riktad från elektroden. Ju högre spänning som läggs på elektroden, desto större blir utarmningsområdet. En infallande foton med tillräckligt hög energi kan slå loss en bunden elektron från kristallgittret. Därvid skapas ett elektron/hål par. Om detta sker i utarmningsområdet kommer elektronen och hålet att separeras av det elektriska fältet. Elektronen vandrar uppåt mot elektroden, där den lägger sig uppe under oxidskiktet (som är en isolator). Hålet däremot vandrar i motsatt riktning och kommer att lämna utarmningsområdet, se Fig. 14.4. Den lägsta fotonenergi som behövs för att skapa ett elektron/hål par i kisel är ca. 1.2 eV, svarande mot en våglängd av ca. en mikrometer. Detta innebär att en photogate av kiselmaterial har god känslighet för synligt ljus och nära infraröd strålning.

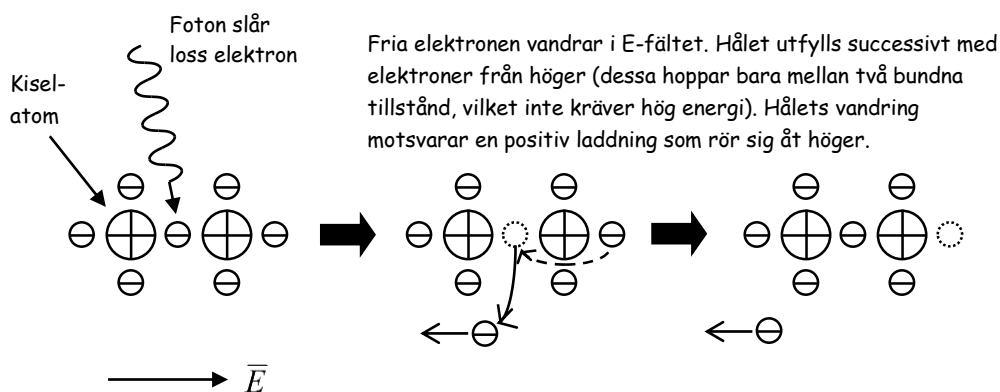
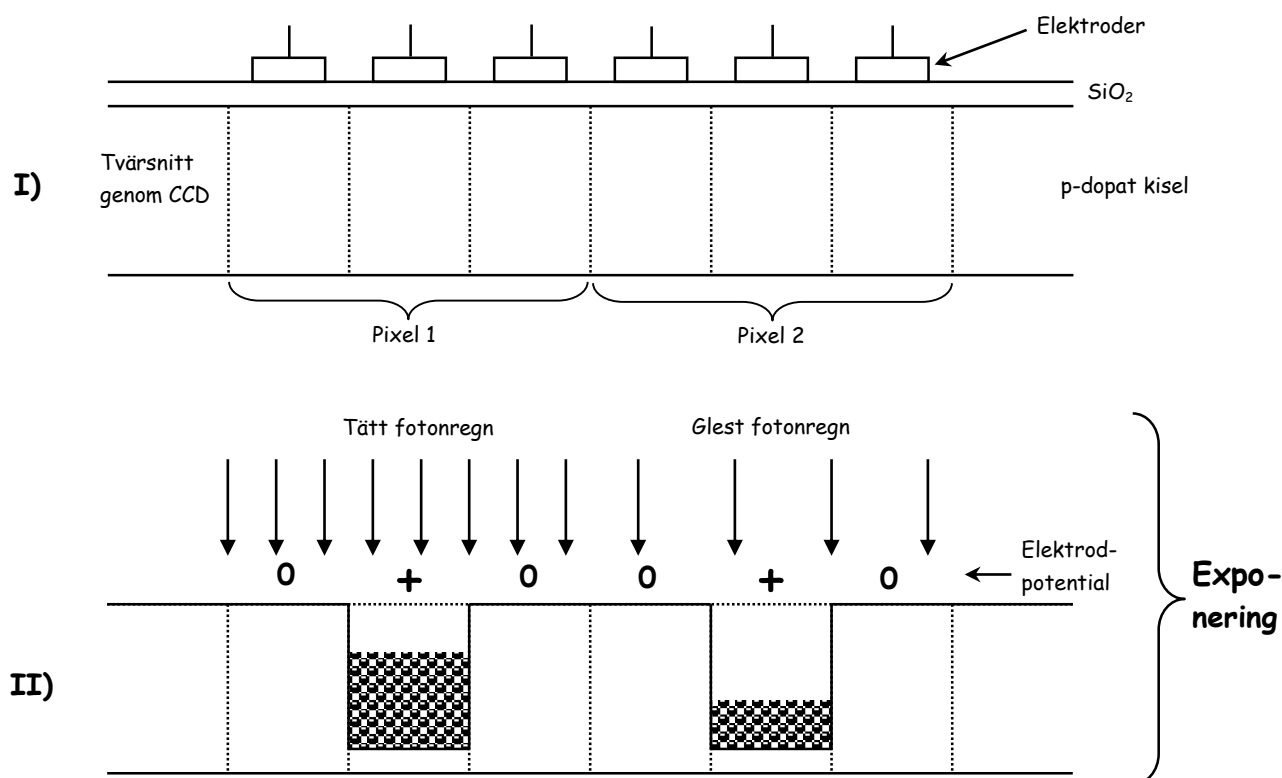
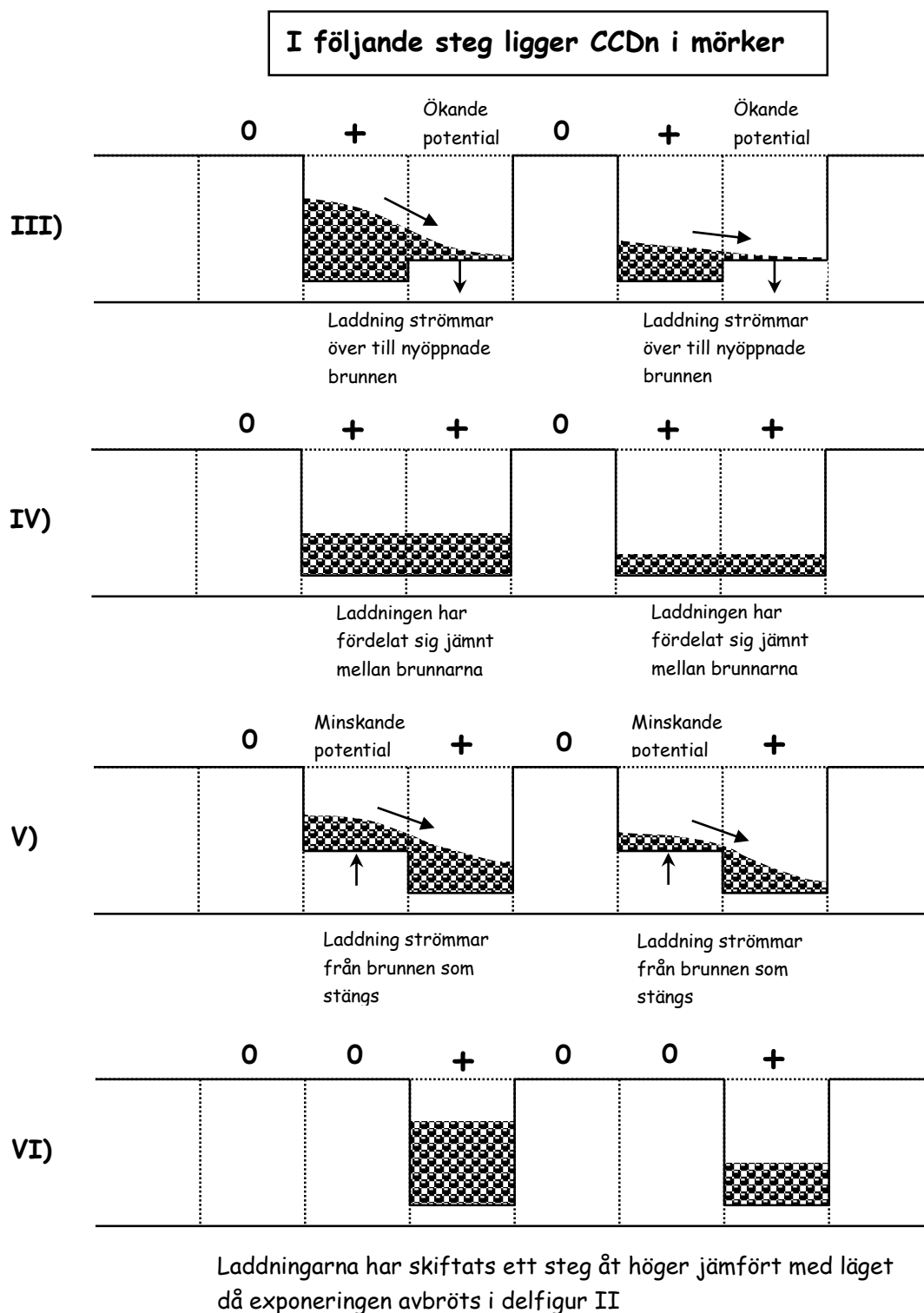


Fig. 14.4. En infallande foton kan skapa ett elektron/hål par.

Ju fler fotoner som absorberas i utarmningsområdet, desto fler elektroner bildas. Dessa elektroner kommer att ansamlas alldeles under den positiva elektroden. Det finns dock en gräns för hur många dessa elektroner kan bli. De kommer att repellera varandra inbördes, och de kommer att med sin negativa laddning skärma, dvs. minska, den positiva elektrodens inverkan på djupare liggande skikt i kiselmaterialiet. Härigenom minskas utarmningsområdet för att slutligen helt försvinna. Om denna situation uppstår kan elektroner börja avlägsna sig och driva bort till intilliggande pixlar. Denna icke önskvärda situation kallas "blooming", och yttrar sig så att en lokal kraftig överexponering kan sprida sig över en stor yta och förstöra hela bilden. Detta fenomen var mer besvärande på äldre sensorer. Numera kan man genom diverse konstruktionsknep begränsa blooming-effekten. Begränsningen att varje pixel bara kan lagra upp ett visst antal elektroner kvarstår dock fortfarande. Detta max-antal brukar benämnas "well capacity" beroende på att man brukar likna elektronuppsamlingen i pixlarna med att man fyller på vatten i brunnar. Tvärsnittsarean på varje brunn (dvs. pixel) bestäms av elektrodens storlek, och dess djup bestäms av pålagd spänning (inom vissa gränser). Alltså, ju större pixelarea och ju högre pålagd spänning, desto fler fotoner kan pixeln detektera under exponeringstiden. Antalet kan skilja sig en hel del mellan olika CCD-sensorer, men 4×10^4 är en rimlig siffra. Låt oss nu titta på en mycket förenklad bild av hur en endimensionell CCD kan vara uppbyggd, Fig. 14.5. Vi använder här brunnsmodellen, och tänker oss att de infallande fotonerna likt regndroppar fyller på i brunnarna. För enkelhets skull utesluter vi oxidskikt och elektroder ur de undre delfigurerna. Som synes består varje pixel av flera elektroder svarande mot olika brunnar vars djup styrs av spänningen på respektive elektrod. Antalet elektroder per pixel kan variera mellan olika sensorer.

Fig. 14.5. Nedanstående sex delfigurer (I-VI) illustrerar med hjälp av "brunnsmodellen" vad som händer vid exponering och utläsning av en CCD-sensor.





I Fig. 14.5 visar delfigur I ett tvärsnitt genom CCD:n. Exponeringen påbörjas genom att lägga en lämplig positiv spänning (det rör sig om någon volt) på elektroderna märkta med + i delfigur II. Infallande fotoner skapar då fria elektroner, vilka likt vattendroppar fyller på i brunnarna. Fotoner som träffar mellan brunnarna ger inga elektroner. Under exponeringstiden är endast en elektrod per pixel positiv, och bara under denna detekteras fotoner. + respektive 0 står för potential på respektive elektrod. Efter exponeringstiden stängs slutaren och inga fler fotoner faller in. Nu ska laddningarna i brunnarna läsas ut från CCD:n. I figurerna III och

IV visar första fasen i denna utläsning. Genom att höja potentialen på elektroderna omedelbart till höger om de redan befintliga brunnarna så breddas dessa brunnar. Laddning strömmar över till de nyöppnade brunnarna, och i sluttillståndet fördelas de ursprungliga laddningarna över dubbelt så stora brunnareor. Nivåerna blir därför bara hälften så höga (figur IV). Andra fasen i utläsningen visas i figur V och VI. Där sänks potentialen för de ursprungliga brunnarna till noll, och därför kommer all laddning att ansamlas i de nyöppnade brunnarna. När denna fas är avslutad har laddningsmängderna i de två ursprungliga brunnarna skiftats ett steg åt höger. Man kan sedan fortsätta på samma sätt och stegvis förflytta laddningarna åt höger. Det är denna laddningsöverföring mellan cellerna som gett kretsen namnet charge-coupled-device (CCD). Låt oss nu titta på hur datautläsningen sker i det tvådimensionella fallet. För enkelhets skull betraktar vi en 2×2 matris som efter exponering har laddningsmängderna $q_1 - q_4$ i de olika pixlarna.

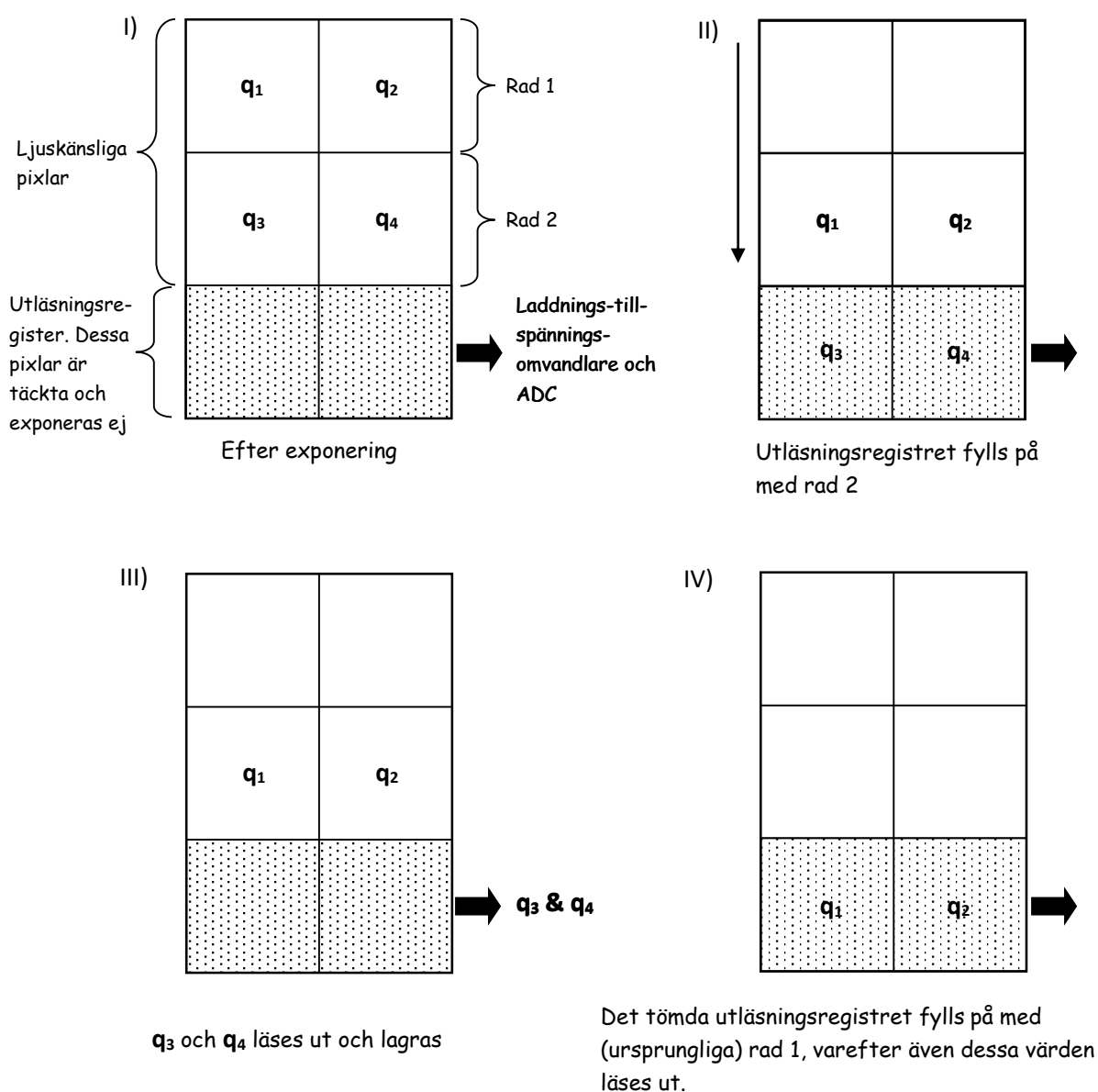


Fig. 14.6. Illustration av tvådimensionell datautläsning i en förenklad 2×2 matrissensor.

I verkligheten finns tusentals rader och kolumner, vilket innebär att en laddningsmängd kan komma att skiftas mellan åtskilliga tusen pixlar innan den slutligen läses ut och analog/digital-omvandlas. Förflyttning från en pixel till nästa kräver dessutom flera skiftsteg såsom framgick av Fig. 14.5. För att detta ska fungera krävs att man har oerhört små laddningsförluster i varje skiftsteg. I praktiken kan det krävas 10 000 eller fler skiftsteg. Ett mått på hur små förlusterna är i ett sådant skiftsteg ges av "charge transfer efficiency", som brukar ligga på typiskt 0.999995. Detta innebär att efter 10 000 steg återstår $0.999995^{10000} = 0.95$ av ursprungliga laddningsmängden. Mycket större förluster än så kan man inte tolerera i högkvalitativa sammanhang. Utläsningen av pixeldata sker ofta med en hastighet av storleksordningen 20 Mpixel/s.

Den hittills beskrivna CCD:n är en så kallad "full frame transfer CCD", en typ som används mest i vetenskapliga sammanhang (t.ex. astronomi). Den innehåller en stor matris med ljuskänsliga pixlar samt ett utläsningsregister längst ned. Utläsningstiden för pixeldata är förhållandevis lång (ofta ca. en halv sekund), vilket gör att denna typ av CCD inte är lämplig om man vill ha "live-view" visning i reducerad upplösning på en LCD-skärm på en kamera. Vidare behövs, om inte exponeringstiden är mycket lång, en mekanisk slutare eftersom sensorn är ljuskänslig under datautläsningen. Visserligen kan man öka utläsningshastigheten genom att använda mer än ett utläsningsregister, men det tar fortfarande ganska lång tid att tömma pixlarna på laddning. En "frame transfer" CCD tillåter att man mycket snabbare (< 1 millisekund) kan tömma de ljuskänsliga pixlarna på laddning, och därefter påbörja en ny exponering om så önskas. Även denna typ används mest i vetenskapliga sammanhang. I frame transfer utförande består CCD:n av två pixelmatriser, en aktiv som detekterar ljus och en passiv som är täckt med ett skikt som absorberar alla fotoner. I Fig. 14.7 visas ett exempel på en frame transfer CCD med 2000×3000 aktiva pixlar (ljuskänsliga) och lika många lagringspixlar (okänsliga för ljus).

Exponeringen börjar med att en positiv spänning läggs på elektroderna i de aktiva pixlarna (exponeringssteget i Fig. 14.5). Efter avslutad exponering så skiftas laddningsmängderna i pixlarna kolumnvis nedåt, så att de överförs till lagringsmatrisen. I det valda exemplet, och med en skiftfrekvens av 20 Mpixel/s,

så går hela processen på $2000 \times \frac{1}{20 \times 10^6}$ sekunder = 100 μ s. Detta är så snabbt att

man i vissa vetenskapliga tillämpningar inte behöver någon mekanisk slutare som avbryter exponeringen. Sedan pixelvärdena överförts till lagringsmatrisen är de okänsliga för ljus, och kan i lugn och ro (i det valda exemplet ca. 0.3 s) läsas ut. Under tiden kan en ny exponering påbörjas. Till skillnad från en full frame transfer CCD, så kan man alltså både exponera och läsa ut data samtidigt med en frame transfer CCD. En nackdel är dock att en mindre procentuell del av ytan är ljuskänslig.

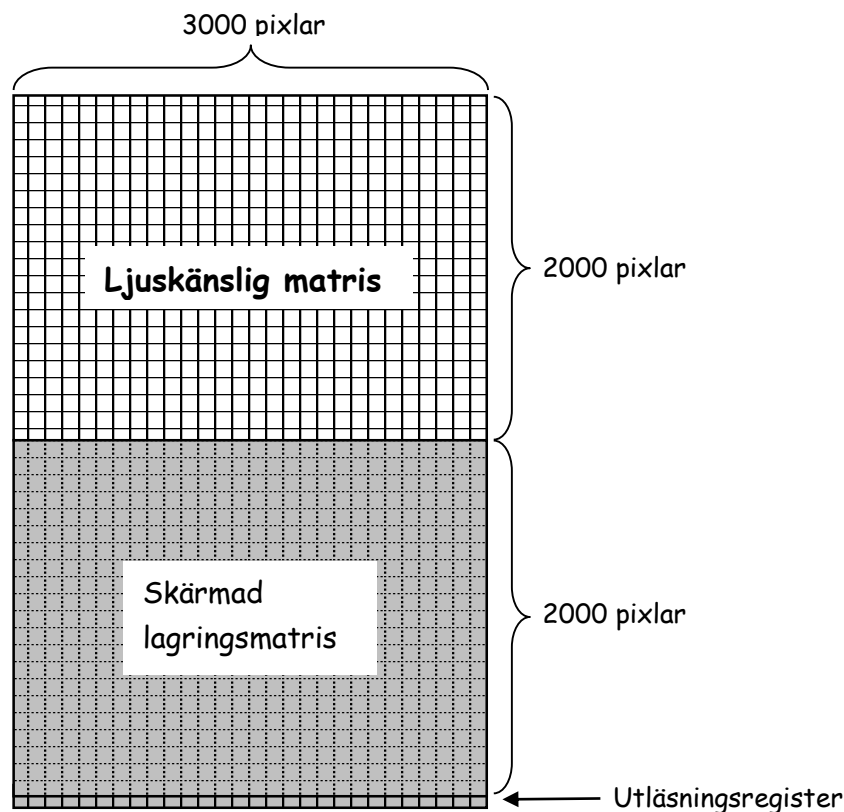


Fig. 14.7. Uppbyggnaden av en "frame transfer" CCD.

Den typ av CCD som sitter i konsumentprodukter (digitalkameror och videokameror), samt även i många professionella kameror, kallas "interline transfer" CCD. Den har fördelen att laddningsöverföringen till lagringspixlarna går ännu snabbare än med en frame transfer CCD. I många fall kan man därför helt slopa mekanisk slutare. Den har också fördelen att man lätt och snabbt kan läsa ut en grov bild med få pixlar som kan användas till "live-view" på en LCD-skärm på kameran. Detta möjliggörs genom att man kan arbeta i flera "utläsningsmoder", ju färre utlästa pixlar desto högre bildfrekvens. I Fig. 14.8 visas den principiella uppbyggnaden av en interline transfer CCD. I denna typ av sensor används vanligen fotodioder och inte photogates. Fotodioder fungerar i stort sett på samma vis som photogates, men har vanligen högre kvantverkningsgrad (färre fotoner går till spillo) eftersom ingen elektrod ligger över utarmningsområdet och absorberar ljus.

Varannan kolumn i en interline transfer CCD består av aktiva, ljuskänsliga pixlar. När exponeringen är färdig skiftas alla pixelladdningar **ett** steg åt höger in i pixlarna i lagringskolumnerna. Detta kan teoretiskt ske mycket snabbt (på ca. 1 μ s), men i praktiken kan problem med strökapacitanser, restladdningsmängder etc. göra att man inte fullt ut kan utnyttja den potentiella hastigheten. Jämfört med en frame transfer matris är dock hastighetsvinsten avsevärd. Efter det att laddningarna skiftats över till lagringskolumnerna så sker utläsningen av data på samma sätt som tidigare beskrivits med hjälp av ett (eller flera) utläsningsregister (se Fig. 14.8). Fill factor är ofta ganska låg på denna typ av CCD (ner till 20%), och därför brukar den

användas i kombination med mikrolinser, se Fig. 14.2. Utläsningshastigheten av data ligger ofta i storleksordningen 50 Mpixel/s för en interline transfer CCD.

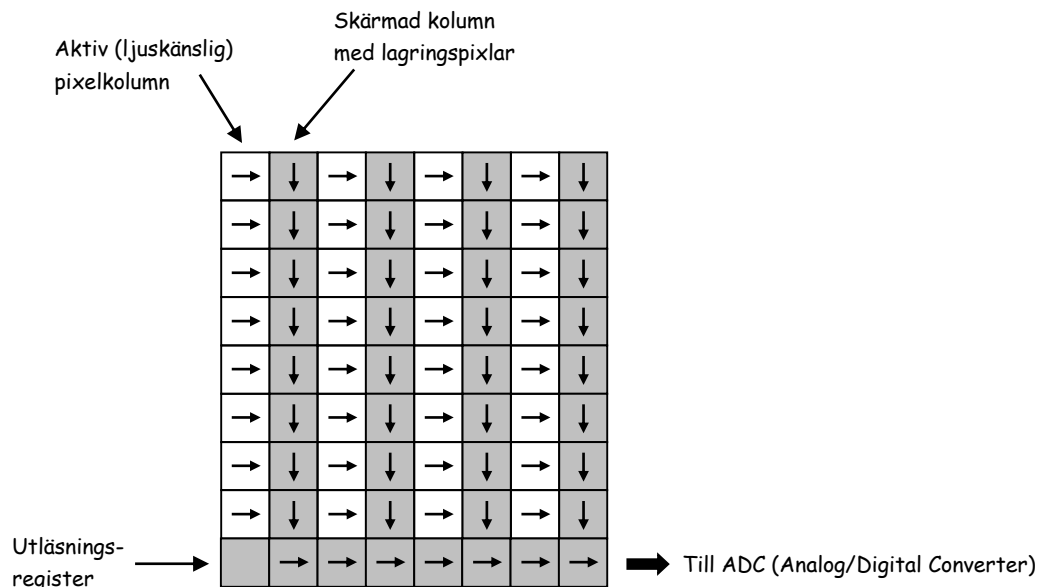


Fig. 14.8. Principiell uppbyggnad och funktion hos en "interline transfer" CCD. Pilarna visar hur laddningarna förflyttas vid utläsning av data.

14.3 CMOS-sensorn

Vi har nu sett principen för hur CCD-sensorer fungerar. Den andra typen av matrissensor som nämndes i kapitlets inledning betecknas med CMOS (complementary metal oxide semiconductor). En stor fördel med CMOS-tekniken är att man på samma krets (chip) kan integrera inte bara de ljuskänsliga pixlarna utan även annan elektronik som klockkretsar, förstärkare och även ADC. En annan fördel är att effektförbrukningen är lägre, vilket gör att kamerans batterier räcker längre. Dessutom är pixlarna individuellt adresserbara, så att man snabbt kan läsa ut bara de pixelvärden man för tillfället behöver. I CMOS-sensorer används vanligen fotodioder för att detektera ljus. Figur 14.9 visar fotodiodens princip, samt hur den är implementerad i en CMOS-sensor i en mycket förenklad version.

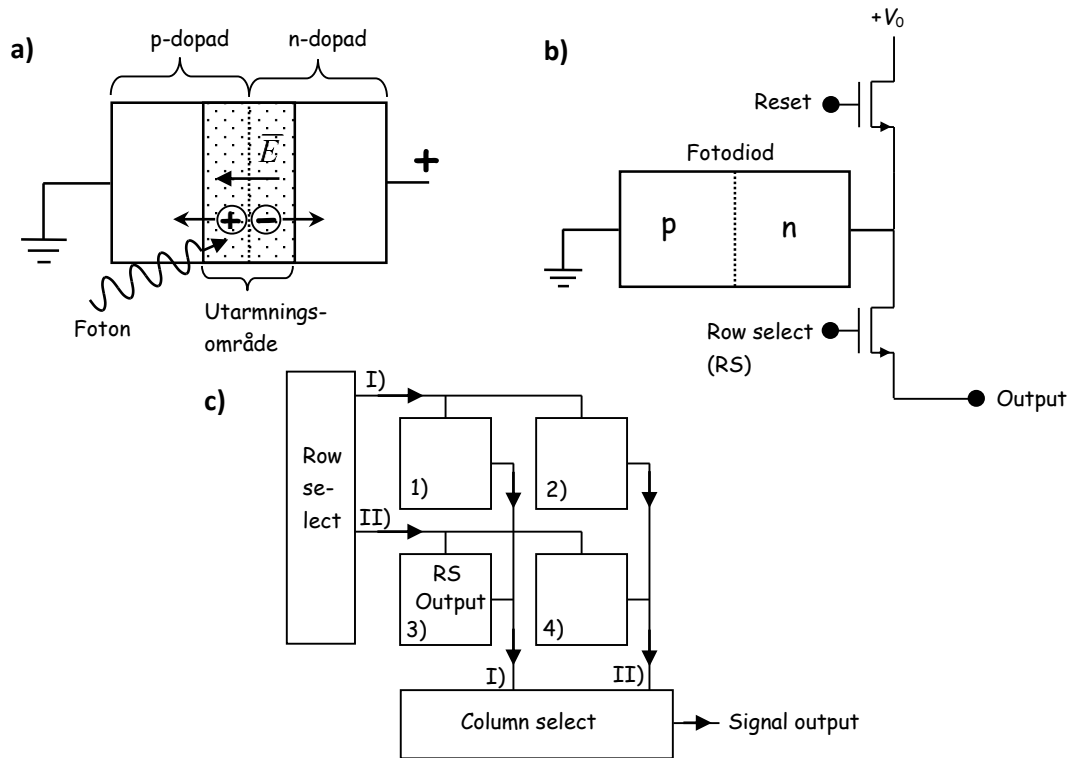


Fig. 14.9. Uppbyggnaden av en CMOS-sensor i kraftigt förenklad version. Funktionen beskrivs mer utförligt i texten. a) Ljusdetektionen i varje pixel sker i en fotodiod. b) Varje pixel innehåller, förutom en fotodiod, också andra elektroniska komponenter för signalutläsning mm. c) Principiella uppbyggnaden av en förenklad matrissensor med 4 pixlar. Genom row och column select styrs vilket värde som läses ut.

Fotodioder detekterar ljus på i princip samma sätt som de photogates vi tidigare tittat på, men det finns vissa praktiska skillnader. Figur 14.9a visar en backspänd diod där det bildas ett utarmningsområde utan rörliga laddningsbärare. I utarmningsområdet finns ett elektriskt fält, vilket gör att de elektron/hål par som bildas av infallande fotoner separeras och färdas åt var sitt håll. Denna process är helt ekvivalent med den som sker i en photogate (Fig. 14.3). Utarmningsområdet skapas vid exponeringstidens början, genom att man under ett kort ögonblick lägger på en positiv spänning på diodens n-dopade sida. Detta sker genom att aktivera "Reset"-ingången i Fig. 14.9b. Dioden fungerar då som en kondensator som laddas upp till spänningen V_0 , och ett utarmningsområde uppstår. Därefter kopplas spänningskällan bort, och kondensatorn (dvs dioden) urladdas successivt av de laddningar som skapas av infallande fotoner. Aktuell spänning över dioden, V , kommer att läggas ut på "output" om "Row select" aktiveras (transistorn behövs för att elektriskt isolera dioden från omgivningen). När utsignalen är avläst, kan man göra "reset" på pixeln (dvs ladda upp den på nytt) och börja exponera nästa bild. Sambandet mellan utspänningen, V , och antalet fotoner som detekterats, N , ges av

$$N = \frac{C(V_0 - V)}{e} \quad (14.1)$$

där e står för elementarladdningen och C är diodens kapacitans. Typiska storleksordningar för en verklig fotodiod är $C \approx 10$ fF (femtofarad) och $(V_0 - V)_{\max} \approx 1$ volt, vilket ger $N_{\max} \approx 50\,000$. Detta ger ungefär samma "well capacity" som för en typisk photogate i en CCD. Man kan öka well capacity genom att installera en liten kondensator parallellt med fotodioden i Fig. 14.9. Detta görs på vissa CMOS-sensorer (den extra kondensatorn är naturligtvis integrerad på samma krets). Om man väljer en kondensator med samma kapacitans som fotodiodens, i det valda exemplet 10 fF, så fördubblar man well capacity. Men samtidigt minskar man känsligheten till hälften, dvs det krävs dubbelt så hög exponering för att få lika hög utspänning $V_0 - V$. Om tillgången på ljus är god är en hög well capacity att föredra, medan en hög känslighet kan vara att föredra om det är ont om ljus. Detta kommer att diskuteras närmare i kap. 14.5. För att kunna erbjuda bägge möjligheterna, har vissa kameror möjlighet att koppla in och ur den extra parallella kondensatorn (detta kallas "dual gain" på engelska). Det brukar fungera så att vid låga ISO-tal (se kap. 14.10) är kondensatorn inkopplad, vilket ger hög well capacity och låg känslighet. När man ställer in ett högre ISO-tal (gränsen ligger ofta vid ISO 500 – 1000) så kopplas kondensatorn ur, vilket minskar well capacity men ökar känsligheten.

I Fig. 14.9c visas schematiskt hur signalutläsningen går till i en enkel 2 x 2 pixel CMOS-sensor. Antag att vi vill läsa ut värdet för pixel 3. Då lägger vi först ut en positiv spänning på "Row select II", vilket gör att utsignalerna från pixlarna 3 och 4 hamnar på utgångarna (dvs på kolumnerna I resp. II). Genom "Column select" registret kan vi välja att signalen på kolumn I ska läggas på utgången (Signal output), och att signalen från kolumn II kommer att blockeras. På detta sätt kan man, även för en godtyckligt stor sensormatris, välja precis vilket pixelvärde man vill ha på "Signal output" i varje ögonblick. Man får härigenom stor flexibilitet att bara läsa ut exakt de pixlar som önskas, vilket möjliggör en mycket hög bildfrekvens om man nöjer sig med få pixlar per bild. Detta utnyttjas i höghastighetsvideo. Bra för höghastighetsvideo (och vanlig fotografi) är också att CMOS generellt sett tillåter en högre utläsningshastighet av data än en CCD. För en CMOS-krets kan utläsningshastigheten ligga runt 500 Mpixel/s, alltså en hel storleksordning högre än för en CCD. Detta beror på att en hel del signalbehandling (laddning- till spänningsomvandling) sker lokalt i varje pixel, och kan sålunda utföras parallellt i många pixlar samtidigt. CMOS-kretsar har också flera tusen signalförstärkare och ADC:er (en för varje pixelrad eller kolumn), vilket också gör kretsen snabb. I en CCD sker denna signalbehandling seriellt för varje pixelvärde vid utläsningen, vilket tar längre tid.

Med tanke på de många fördelarna, kan man undra varför inte alla digitalkameror använder CMOS-sensorer snarare än CCD. Ursprungligen var det så att CMOS gav en klart sämre bildkvalitet. Högre mörkersignal och brus har, åtminstone tidigare, varit ett problem. Mörkersignalen uppstår genom att inte bara fotoner, utan även termisk rörelse i materialet, kan skapa laddningsbärare. Detta ger samma effekt som ströljus och försämrar bildkvaliteten. Det högre bruset associerat med CMOS beror huvudsakligen på variationer mellan de förstärkare som är integrerade i varje pixel. Detta ger ett så kallat "fixed pattern noise" som innebär att olika pixlar har lite olika ljuskänslighet. Egentligen bör inte detta benämnas brus, eftersom det innebär en systematisk och inte slumpvis avvikelse från korrekt värde. Fixed pattern noise är mycket störande, och ursprungligen användes CMOS-sensorer mest i enkla

applikationer där liten storlek och låg strömförbrukning var viktigare än god bildkvalitet. På senare år har emellertid bilden förändrats. Fixed pattern noise går att kompensera eftersom det är en systematisk avvikelse, och nyare CMOS-kretsar har sådan kompensation inbyggd i hårdvaran. Detta, plus att man kan tillverka stora (24 mm x 36 mm eller mer) CMOS-sensorer till överkomligt pris, gör att CMOS idag används i de flesta digitalkameror. I kompaktkameror med små sensorer (mindre än ca. 5 mm kantlängd) används både CCD och CMOS. CCD verkar ha en viss nisch här på grund av sin något högre bildkvalitet, vilket är speciellt viktigt vid liten sensorstorlek. I riktiga proffskameror typ Hasselblad kan man hitta både CCD- och CMOS-sensorer, ofta i storlekar uppåt 35 mm x 50 mm. Här spelar pris och strömförbrukning ofta mindre roll, vilket gör att CCD ofta har använts. Men CMOS har nått en sådan kvalitet att den numera ofta används även i denna kameratyp.

Sammanfattningsvis kan man alltså säga att för en given sensorstorlek så tenderar CCD att ge något bättre bildkvalitet. Men genom att man kan tillverka CMOS i stora storlekar till rimligt pris så har CMOS tagit över mycket av marknaden. Att sensorstorlek, snarare än antalet ”megapixlar”, betonats i ovanstående resonemang beror på att pixlarna inte bör vara för små till ytan, eftersom detta kommer att begränsa bildkvaliteten. Mer om detta i kapitlet om bildkvalitet.

14.4 Elektroniska slutare i CMOS-sensorer

Elektroniska slutare fungerar lite olika beroende på sensortyp, men vi begränsar oss till att titta på den vanligaste sensortypen, CMOS. I en CMOS-sensor används vanligen en metod som kallas ”rolling shutter” (rullande slutare), och det är i princip en elektronisk version av ridåslutare (se kap. 4.3). I Fig. 14.10 visas principen för en rolling shutter.

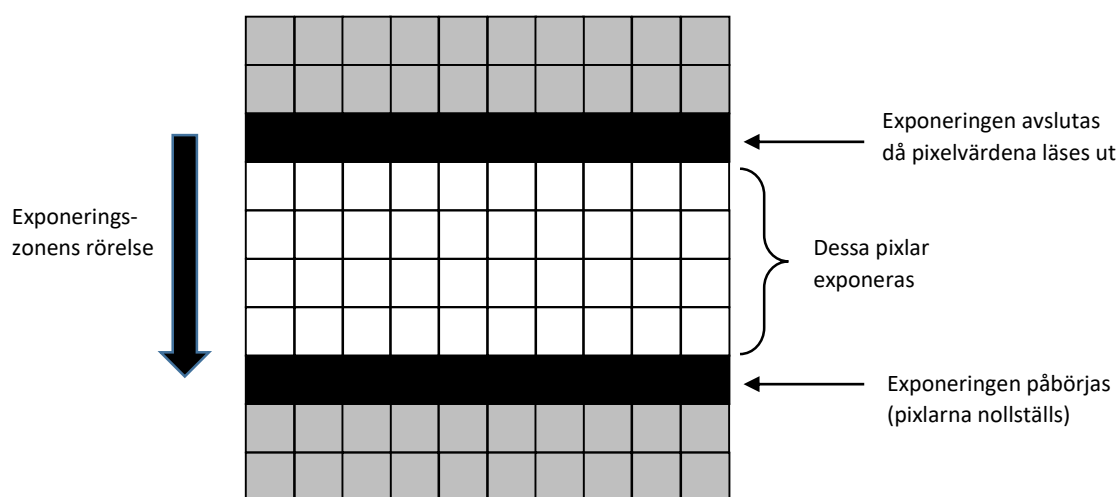


Fig. 14.10. ”Rolling shutter” i en CMOS-sensor fungerar som en elektronisk ridåslutare. Genom att variera fördröjningen mellan exponeringens början och utläsningen av pixelvärden kan exponeringstiden varieras.

Med en rolling shutter kan man nå mycket korta exponeringstider, kortare än med en mekanisk slutare eftersom inga rörliga delar finns. Men precis som för en mekanisk ridåslutare så exponeras inte alla pixlar i sensorn samtidigt, vilket kan ge kraftigt distorderade former om motivet (eller fotografen) rör sig. Och här har vi akilleshälen för en rolling shutter. Även om den klarar mycket korta exponeringstider, så är totala tiden det tar att exponera en bild förhållandevis lång, ibland uppåt 100 millisekunder. Detta är klart långsammare än en mekanisk ridåslutare som klarar av samma sak på typiskt 4-5 millisekunder. Därför är det mycket större risk att få rörelsedistorsion i bilderna med en elektronisk ridåslutare. Användningen av blixtnöjs också, eftersom man då måste använda mycket långa exponeringstider som kan innebära att icke önskat omgivningsljus påverkar exponeringen[†].

Varför tar då hela exponeringen så lång tid med rolling shutter trots att den inte har några rörliga delar? Svaret ligger i den elektronik som sköter utläsningen av pixelvärden. Det finns en gräns för hur snabbt detta kan ske, och i dagsläget (2016) ligger den gränsen vid några hundra megahertz. För en kamera med många pixlar (36-50 megapixel) och rolling shutter, innebär detta att hela exponeringsförloppet skulle kunna ta uppåt en tiondels sekund. Då är risken mycket stor att få bilddistorsion på grund av att motivet eller fotografen rör sig. Därför brukar rolling shutter bara användas på mindre sensorer där totala utläsningstiden blir kortare[†]. Det är inte omöjligt att utläsningshastigheten kan ökas i framtiden. Man får emellertid komma ihåg att högre utläsningshastighet medför att större elektrisk bandbredd behöver utnyttjas, något som i sin tur ökar bruset (kap. 14.5). En annan möjlighet är att utnyttja en sensorarkitektur som tillåter så kallad ”global shutter”. Då exponeras alla pixlar samtidigt, och värdena lagras därefter lokalt i varje pixel efter exponeringens slut. Detta liknar funktionen hos en interline transfer CCD, se Fig. 14.8. Resultatet blir att med en global shutter så kan exponeringen påbörjas och avslutas samtidigt för alla pixlar. Efter exponeringstidens slut kan sedan pixelvärdena läsas ut i lugn takt, all exponering är ju redan avslutad. Global shutters används redan i videosammanhang på mindre sensorer, men har ofta inneburit en viss kompromiss vad gäller brus och mörkersignal. Det är möjligt att kretsteknologin kommer att utvecklas, och att global shutters kommer att införas även på stillbildskameror med stora sensorer. I så fall kan mekaniska slutare kanske helt elimineras på alla kameror.

Slutligen kan nämnas att det finns en hybridvariant mellan elektronisk rolling shutter och mekanisk ridåslutare. Den brukar kallas ”electronic first curtain” (elektronisk första ridå). På systemkameror kan man ofta ställa om mellan helt mekanisk ridåslutare och electronic first curtain. När man använder electronic first curtain så är mekaniska slutaren helt öppen vid start, och exponeringen påbörjas precis som vid rolling shutter, se Fig. 14.10. Denna start av exponeringen kan emellertid svepa ner över sensorn mycket snabbare än vid rolling shutter, eftersom pixlarna bara behöver nollställas och inga värden läses ut. Exponeringen avbryts emellertid av en mekanisk ridå, som efter en viss fördröjning sveper ner över

[†] En ytterligare anledning att ha en mekanisk slutare i en systemkamera, är att då denna är stängd så minskar risken att få damm på sensorn vid objektivbyte.

[†] En kamera som erbjuder ”live view” på LCD-skärm eller i sökaren använder i detta läge alltid elektronisk slutare (en eventuell mekanisk slutare är då öppen hela tiden). Att detta fungerar även på systemkameror med stora sensorer beror på att man i live view läget bara läser ut en mindre delmängd av det totala antalet pixlar.

sensorn med samma hastighet som när exponeringen startades. Sedan kan pixelvärdena läsas ut i mycket långsammare takt, eftersom pixlarna då ligger i mörker och inte påverkas av en fördröjning.

En fördel med electronic first curtain jämfört med en helt mekanisk slutare, är att man undviker vibrationer som orsakas av första ridåns rörelse. Detta har betydelse framförallt när man använder långa brännvidder. Electronic first curtain ger också en minskad fördröjning mellan att man trycker ner avtryckaren och att bilden tas, samt reducerar också ljudet från slutaren något. En nackdel med electronic first curtain är att det finns en fara att man får en ojämn exponering över bildytan vid mycket korta exponeringstider (t.ex. 1/1000 sekund och kortare). Orsaken till detta, är att den mekaniska ridån som avslutar exponeringen inte ligger i sensorplanet utan en bit framför, se Fig. 14.11. Detta medför att mekaniska ridåns skugga i sensorplanet ligger olika långt efter den elektroniska ridån beroende på om man befinner sig i över- eller underkant av bildfältet. Detta i sin tur beror på att ljusets infallsriktning mot bildplanet är olika i över- och underkant såsom framgår av Fig. 14.11. Graden av ojämnhet i exponeringen beror emellertid på det använda objektivet brännvidd och konstruktion*, samt avståndet mellan ridå och sensor. Effekten är som nämnts försumbar vid längre exponeringstider. Kamerafabrikanterna brukar rekommendera att man inte använder electronic first curtain vid mycket korta exponeringstider.

* Det finns en typ av objektiv som kallas ”telecentrisk”, och där ljusets infallsvinkel mot bildplanet är densamma över hela bildytan. För sådana objektiv kommer man att få en helt jämn exponering även när electronic first curtain används vid mycket korta exponeringstider. Telecentrisk objektiv är dock mycket ovanliga i fotografiska sammanhang.

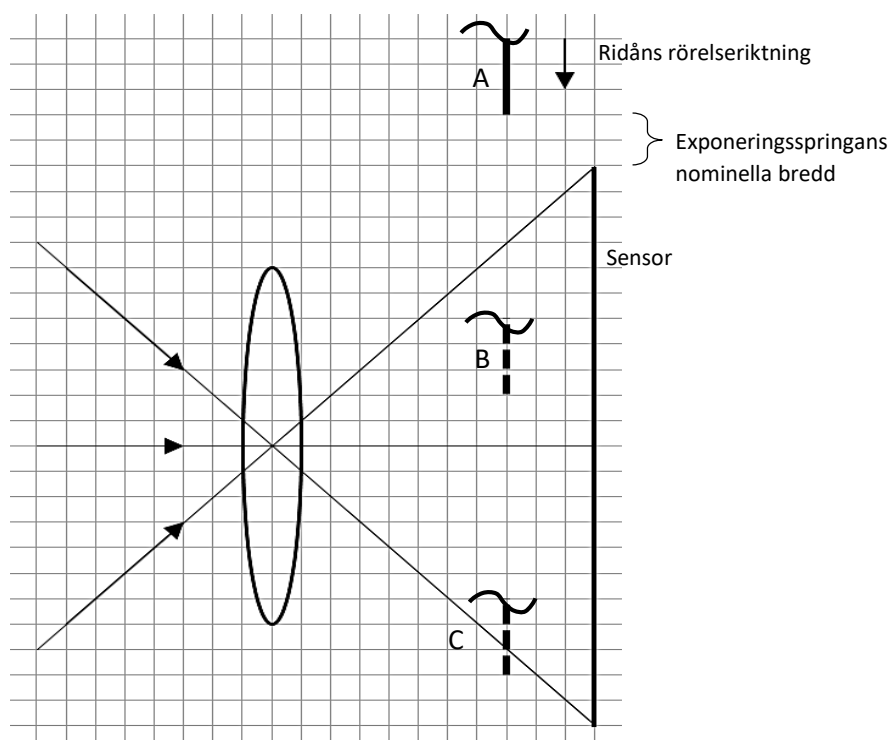


Fig. 14.11. Illustration av skuggningseffekt som kan förekomma vid användning av ridåslutare med "electronic first curtain". För enkelhets skull visas endast centrala strålar genom linsen. När ridåns främre kant befinner sig i läge A påbörjas exponeringen av översta pixelraden i sensorn. Exponeringen av översta raden avbryts först när ridån rört sig 5 rutor nedåt i figuren, dvs effektiva springbredden är 5 rutor istället för 2 (nominella värdet). När ridåns framkant befinner sig i läge B påbörjas exponeringen av pixlarna i bildcentrum, och den avbryts när ridån rört sig 2 rutor nedåt. Effektiv springbredd = nominellt värde, 2 rutor. När ridåns framkant befinner sig i läge C påbörjas exponeringen av pixlarna i sensorns understa rad, men då blockeras ljuset redan av ridån och exponeringen blir noll. Effektiv springbredden blir alltså noll i detta fall. Slutresultatet blir en bild som är överexponerad i övre delen, korrekt exponerad i centrum, och underexponerad i nedre delen. Effekten har överdrivits i figuren för tydlighets skull.

14.5. Fotometrisk kvalitet

Vad gäller den fotometriska kvalitén på de pixeldata som kommer från matrissensorer så finns några mått som ofta citeras, och som kan vara värda att känna till*. De storheter som kommer att beröras här är "Well capacity" (osäkert vad det heter på svenska), "Brus", "Mörkersignal", "Dynamisk vidd", "Signal/brus förhållande" och "Kvantverkningsgrad".

Well capacity har berörts tidigare i detta kapitel. Det talar om hur många elektroner som maximalt kan lagras upp i varje pixel under exponeringen. Allmänt gäller att ju fler desto bättre. Många elektroner betyder ett säkert pixelvärde (mindre

* Intressanta data kan hittas på <http://www.photonstopphotos.net/> och www.dxomark.com.

brus i bilderna) och större kontrastomfång (se Dynamisk vidd). Sensorer med stora pixlar har högre well capacity än sensorer med små pixlar, eftersom well capacity ökar ungefär linjärt med pixelarean. En fullformatsensor (24 mm x 36 mm) med 24 Mpixlar (6 μm pixelstorlek) kan typiskt ha en well capacity på 50 000 – 80 000. En APS-C sensor (23.6 mm x 15.8 mm) med 24 Mpixlar (4 μm pixelstorlek) kan typiskt ligga på 30 000 – 40 000. Mobilkameror har mycket små sensorer, typiskt ca. 4 mm x 6 mm, och ofta uppåt 20 Mpixlar. Detta innebär mycket små pixlar, bara drygt 1 μm , och en well capacity runt 2000 – 5000.

Brus innebär att man har slumpvisa variationer överlagrade på signalen från sensorn. Ju mera brus, desto grynigare bilder kommer man att få. Brus brukar uttryckas som så kallat RMS-värde (efter engelskans root mean square), vilket innebär att man mäter standardavvikelsen i signalen. Brus orsakas av flera olika faktorer som här kommer att beskrivas.

”Read-out noise”, alltså utläsningsbrus på svenska, härstammar från de elektroniska komponenter som används i sensorns laddning-till-spänningsomvandling, analoga signalförstärkning samt i ADC:n (analog-till-digital omvandlaren, se kap. 14.7). Signalens väg i kameran visas schematiskt i Fig. 14.12.

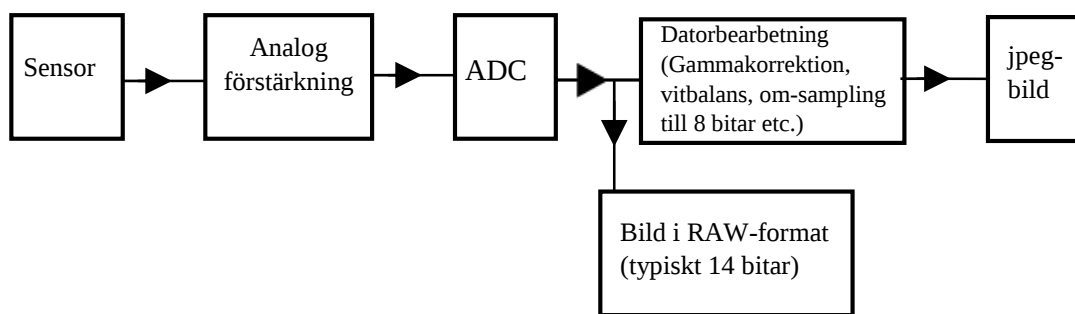


Fig. 14.12. Förenklad bild av signalens väg i en digitalkamera.

När sensorn ligger i mörker, och exponeringstiden är mycket kort, är det i stort sett bara read-out noise som förekommer i signalen. Detta brus brukar ofta uttryckas som ett antal elektroner, typiskt ca. 2-5. Detta betyder att om antalet upplagrade elektroner i sensorn uppvisar denna standardavvikelse, så skulle det ge den observerade brusnivån i utsignalen från ADC:n. Eftersom analoga förstärkningen ökar linjärt med inställda ISO-talet (som beskrivs i kap. 14.10) så ökar också read-out noise linjärt med ISO-talet. Vid mycket låga ISO-tal, dvs låg analog förstärkning, får man dock en avvikelse från linearitet. Detta beror på ett konstant bakgrundsbrus från ADC:n, vilket är oberoende av ISO-talet.

Read-out-noise kan beskrivas matematiskt genom uttrycket

$$\sigma_{read-out} = \sqrt{\sigma_{ADC}^2 + (\sigma_{sensor} \cdot g)^2} \quad (14.2)$$

där $\sigma_{read-out}$ är totalt read-out noise mätt på ADC-utgången, σ_{ADC} är brusbidraget från ADC:n, σ_{sensor} är bruset från sensor och analog förstärkare, och g är analoga förstärkningen (eng. gain). Samtliga σ -värden avser standardavvikelse. Bruset från ADC:n ligger typiskt på $\lesssim 1$ digital enhet (teoretiskt minimum är $\frac{1}{\sqrt{12}}$ enheter). För

låga ISO-tal, säg runt 100, har en 24 Mpixel fullformatssensor typiskt ett värde på $\sigma_{\text{sensor}} \cdot g$ som också ligger på $\lesssim 1$ digital enhet (med 14 bitars ADC som är typisk i digitalkameror).

Som exempel på verkliga prestanda kan vi utgå från publicerade data för Sony A7*. För denna kamera har följande värden erhållits: $\sigma_{\text{ADC}} = 0.94$ och $\sigma_{\text{sensor}} \cdot g = 0.755$ (vid ISO 100). Det innebär att totalt read-out noise ligger på $\sqrt{0.94^2 + 0.755^2} = 1.21$ digitala enheter, vilket innebär att bruset är $\frac{1.21}{2^{14}-1} = 7.4 \times 10^{-5} = 0.007\%$ av maximal signalnivå. Detta representerar den lägsta brusnivå vi någonsin kan få i en bild. Så låg brusnivå kan man alltså förvänta sig i riktigt mörka bildpartier när exponeringstiden är mycket kort. Men observera att detta gäller bara för en bild registrerad i RAW-format. I t.ex. en jpeg-bild representeras varje pixelvärde av endast 8 bitar (trots att ADC:n ger typiskt 14 bitar). Detta kommer att innebära en högre brusnivå jämfört med maximal signalnivå.

När man ökar ISO-talet ökar bruset påtagligt, eftersom analoga förstärkningen, g , ökar linjärt med ISO-talet. För Sony A7 så ökar alltså $\sigma_{\text{sensor}} \cdot g$ till 7.55 vid ISO-tal 1000, vilket ger ett read-out noise på 7.6 digitala enheter (i 14-bitars representation). Vid ISO-tal 10 000 ökar read-out noise till 76 enheter. Olika kameror har naturligtvis lite olika prestanda, men de siffror som angivits Sony A7 torde vara ganska typiska för en 24 Mpixel fullformatskamera.

Vissa kameror har så kallad ”dual gain” design (kap. 14.3). Det innebär att man vid ISO-tal runt 500 – 1000 ökar sensorns känslighet (på bekostnad av well capacity). Detta innebär att analoga förstärkningen kan reduceras, och därmed reduceras även read-out noise. Som exempel på detta kan vi ta värden för kameran Sony A7 III*. När ISO ökar från 500 till 640 så minskar read-out noise från 4.1 digitala enheter till 2.5. Vid fortsatt ökning av ISO ökar bruset ungefär linjärt med ISO-talet, men hela tiden på en lägre nivå än utan dual gain design.

Mörkerbrus orsakas av statistiska variationer i mörkersignalen (som kommer att beskrivas nedan). Mörkerbrusets RMS-värde växer med kvadratroten ur exponeringstiden, och kan därför begränsa bildkvaliteten vid långa exponeringstider. Kameror med stora pixlar tenderar att ha högre mörkersignal och därmed högre mörkerbrus. Men detta kompenseras mer än väl av den högre signalnivån som stora pixlar ger. Ett typiskt värde för mörkerbruset vid rumstemperatur (bruset beror på temperaturen) kan vara 1-2 elektroner RMS för en exponeringstid av 1 sekund. Eftersom exponeringstiden vid normal fotografering är betydligt kortare än så, kan mörkerbruset ofta försummas.

Fotonbrus är oftast det helt övervägande bruset vid lite högre ljusnivåer, dvs överallt utom i bildernas skuggpartier. Fotonbrusets RMS-värde är kvadratroten ur antalet detekterade fotoner. När vi närmar oss maximal exponering, dvs ”well capacity”, kan vi typiskt ha ca. 60 000 elektroner i en fullformatsensor, vilket ger ett RMS-brus på $\sqrt{60000} \approx 240$ elektroner, vilket är betydligt högre än de brusbidrag som beskrivits tidigare. Fotonbruset beror på den slumpvisa ankomsttiden för enskilda fotoner, och är alltså en inneboende egenskap hos ljuset självt. Detta sätter en teoretisk gräns för hur lågt bruset kan bli i en bild.

* Data hämtade från <http://www.photonstophotos.net/>

Mörkersignalen anges ofta som (medel)antalet termiskt genererade elektroner i en pixel per sekund. Typiskt kan detta värde ligga runt 2-3 för en sensor med 6 μm pixlar, men variationen mellan olika pixlar kan vara stor. Ofta finns ett antal pixlar i sensorn som har mycket högre mörkersignal än resten. Mörkersignalen växer linjärt med exponeringstiden och kan bli besvärande vid riktigt långa tider. Mörkersignalen kan emellertid minskas kraftigt genom att man kyler sensorn. Detta används i t.ex. astronomiska tillämpningar. Många matrissensorer har några extra pixlar som alltid ligger i mörker och registrerar enbart mörkersignalen. Denna (medel)mörkersignal kan sedan subtraheras från de utlästa pixelvärdena. Eftersom mörkersignalen varierar kraftigt mellan olika pixlar gör en del kameror kompensationen så att man efter bildexponeringen registrerar en mörkerbild med samma exponeringstid men med stängd slutare. Denna mörkerbild subtraheras sedan pixel för pixel från den första bilden. Härigenom kan den fotometriska "offset" som mörkersignalen ger kompenseras för varje pixel. Däremot kan man inte komma åt det brus (dvs. slumpvisa variationer i tiden) som mörkersignalen ger upphov till. Detta brus beror på kvadratroten ur mörkersignalen. Om man minskar mörkersignalen (t.ex. genom kylning), så kommer man alltså att också minska bruset i bilderna. Mörkersignalen är större för stora pixlar eftersom den är ungefär proportionell mot pixelarean.

Dynamisk vidd för en sensor definieras som $\frac{\text{Maximal utsignal}}{\text{Brusnivå i mörker}}$. Den ger ett mått på förhållandet mellan den högsta och lägsta exponering (kontrastomfånget) som sensorn klarar av att registrera i en och samma bild. Man sätter då upp som krav att de ljusaste partierna inte får bli kritvita och urfrätta, dvs vi får inte överskrida "well capacity". Likaså får exponeringen i skuggpartierna inte bli så låg att de försvinner i bakgrundsbruset. Man kräver därför att utsignalen från sensorn i ett skuggparti måste ligga en "viss bit" över mörkersignalen. Som "viss bit" har man valt brusnivån (mätt som standardavvikelse) när sensorn befinner sig i mörker.

Signal- och brusnivåerna brukar ofta räknas om så att de uttrycks i hur många elektroner de motsvarar. Om well capacity = 60 000 (typiskt värde för fullformatsensor) så svarar detta tal alltså mot "Maximal utsignal". Men vad är det som bestämmer "Brusnivån i mörker"? Jo, denna sammansätts av read-out noise och mörkerbrus (se ovan). Om exponeringstiden är så kort att vi kan försumma mörkerbruset, så innebär det att brusnivån i mörker utgörs av enbart read-out noise. Som vi såg tidigare kan detta brus typiskt vara 2-5 elektroner. Om vi som exempel väljer 4 elektroner, så innebär det att dynamiska vidden blir 15 000. Detta gäller vid lägsta möjliga ISO-inställning (vanligen runt 100). När man utnyttjar högre ISO-tal kommer dynamiska vidden att sjunka, eftersom man inte utnyttjar sensorns hela well capacity (se kap. 14.10). Ju mindre pixlar en sensor har, desto lägre blir well capacity och därmed även dynamiska vidden.

Som ett konkret exempel på dynamisk vidd kan vi återigen betrakta Sony A7. Med 14 bitars ADC blir maximal utsignal $2^{14} - 1 = 16383$. Som vi tidigare såg var read-out noise 1.21 vid ISO 100, vilket ger en dynamisk vidd av $\frac{16383}{1.21} \approx 13\,500$. Vid ISO 1000 respektive 10 000 var read-out noise 7.6 respektive 76, vilket ger dynamiska viddar av ca. 2200 och 220. Ofta uttrycks dynamiska vidden i bländarsteg, vilket för en dynamisk vidd av 13 500 blir 13.7 ($2^{13.72} \approx 13\,500$). Ett

annat sätt är att uttrycka den i decibel (dB), vilket blir $20 \times \log_{10}(13\,500) = 83$ dB. Precis som för read-out noise, gäller att prestanda försämras i fallet att man har en 8-bitars bild, typ jpeg, istället för en RAW-bild. Detta kommer att undersökas närmare i kap. 14.7. Som jämförelse kan också nämnas att det mänskliga ögat har en dynamisk vidd av ca. 10 000.

En kamera kan, åtminstone teoretiskt, återge ett motiv med större kontrastomfång än vad som framgår av dynamiska vidden. Detta beror på att ströljus i optiken ger ett lägre kontrastomfång i optiska bilden än vad motivet har (se kap. 9.15). Men ströljuset medför att kontrasten blir låg i mörka bildpartier, vilket ofta begränsar dynamiska vidden för kameran som helhet. Det finns emellertid gott om motiv med extremt stort kontrastomfång, t.ex. motljusbilder, och där räcker inte dynamiska vidden för någon digitalkamera till. Då kan man använda HDR (High Dynamic Range) fotografi. Detta innebär att flera olika kraftiga exponeringar kombineras i datorn till en enda bild.

Signal/brus förhållande, ofta förkortat SNR (efter engelskans signal-to-noise ratio), är en parameter som ibland förväxlas med "Dynamisk vidd". Som framgår av namnet definieras SNR som $\frac{\text{Utsignal}}{\text{Brusnivå i utsignalen}}$, där brusnivån anges som standardavvikelse. Till skillnad från dynamisk vidd, så gäller för SNR att både utsignal och brus ska mätas upp samtidigt, alltså vid samma exponeringsnivå (med "Utsignal" avses nivån över mörkersignalvärdet). Detta innebär att när detektorn ligger i mörker är $\text{SNR} = 0$. Störst SNR får man alltid vid maximal utsignal. I detta fall brukar oftast gälla att brusnivån helt domineras av fotonbrus. Då gäller att bruset (standardavvikelsen i antalet elektroner) är lika med kvadratroten ur antalet upplagrade elektroner i en pixel. Om vi återvänder till vårt tidigare exempel med en maximal signal motsvarande 60 000 elektroner, så blir bruset $= \sqrt{6.0 \times 10^4} = 245$ elektroner (medan read-out noise var ca. 2-5 elektroner, och mörkerbruset oftast är ännu mindre). Detta ger oss $\text{SNR} = \frac{6.0 \times 10^4}{\sqrt{6.0 \times 10^4}} = 245$, eller 48 dB (vi har då försummat andra bruskällor än fotonbrus). Detta är maximal SNR, vilken erhålls vid lägsta ISO-tal, och i ett mycket ljusst parti i bilden. När man går mot mörkare partier går SNR slutligen ner mot noll.

En ökning av ISO-talet innebär att man inte utnyttjar sensorns hela well capacity, vilket kommer att ge ett lägre SNR. En ökning av ISO från 100 till 1000 i ovanstående exempel, skulle innebära att SNR sjunker från ca. 240 till knappt 80; och vid ISO 10 000 skulle SNR sjunka till lite drygt 20. För att få maximalt SNR ska pixlarna alltså kunna lagra så många elektroner som möjligt. Detta betyder att de ska ha så stor area som möjligt. Av denna anledning kan man inte krympa matrissensorerna allt för mycket, och det finns därför en gräns för hur små kamerorna kan göras.

Vad har man då för praktisk nytta av ett högt SNR vid fotografering? Jo, ett lågt SNR kommer att ge bilder som blir väldigt gryniga. Mest märkbart är detta om man tittar på lite större partier med samma gråton, t.ex. ett himmelsparti eller liknande. För att få lite känsla för hur bilder med olika SNR ser ut i praktiken, återges i Fig. 14.13 ett antal bilder som är identiska förutom att SNR varierar mellan 50 och 1. Det kan dock vara tekniskt svårt att i detta kompendium återge bilder med högt

SNR på ett rättvist sätt. Som jämförelse kan därför nämnas att ett SNR på 50 – 100 uppfattas som en bra bild. Vid SNR runt 30 börjar man tydligt uppfatta bruset i bilden, och vid SNR 10 upplever man bilden som klart grynig. I färgbilder har brusprickarna ofta olika färger (Fig. 14.27 b), vilket upplevs som mer störande än svartvitt brus. Det förtjänar dock att påpekas att man störs mindre av grynigheten i en bild som innehåller många pixlar jämfört med en bild som innehåller få pixlar, om bägge har samma SNR. Detta beror troligen på att ögat har en förmåga att medelvärdesbilda vid betraktning av en tätt pixellerad bild.

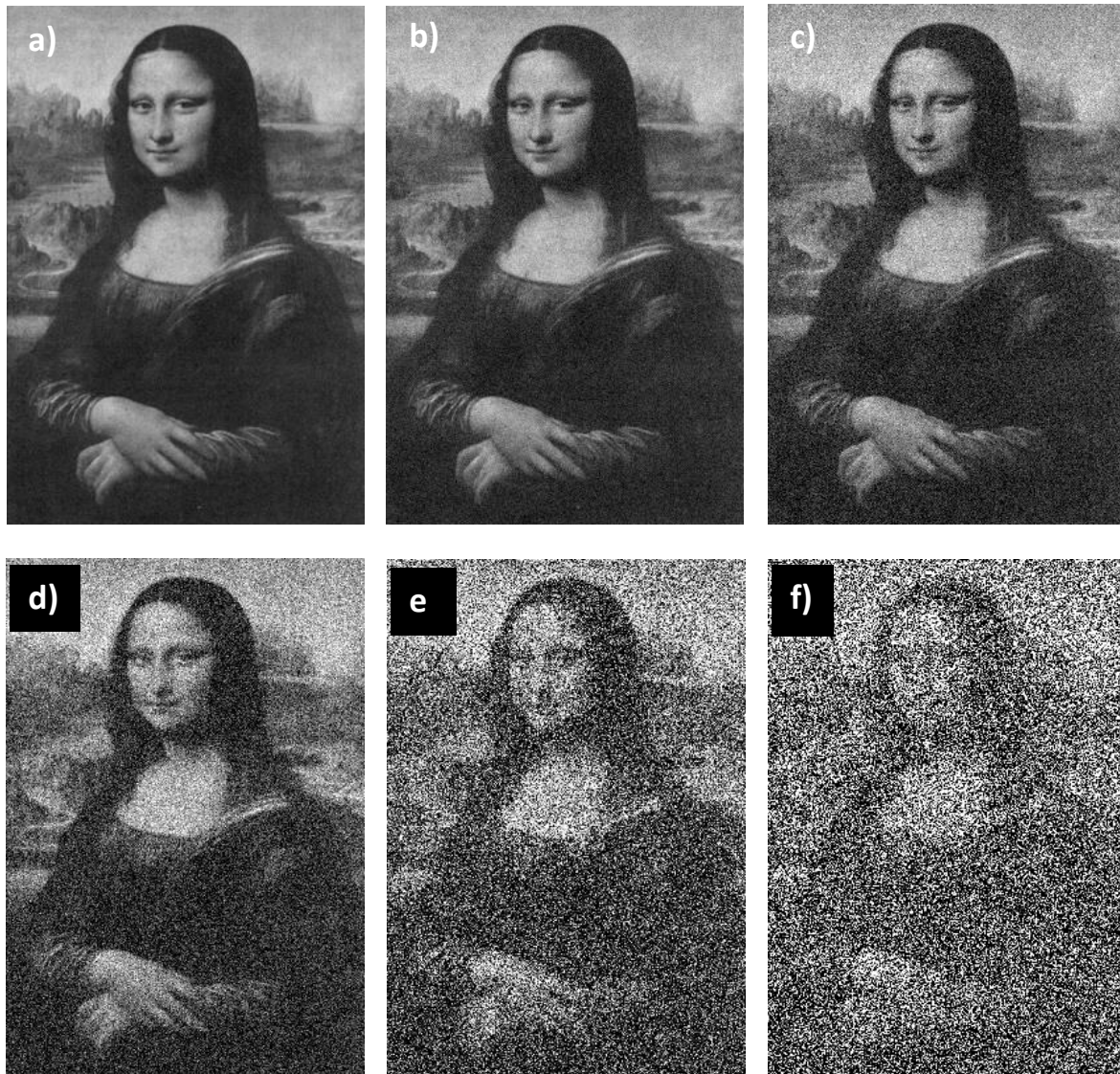


Fig. 14.13. Samma motiv avbildat med olika signal-brus-förhållande (SNR). a) SNR = 50. b) SNR = 20. c) SNR = 10. d) SNR = 5. e) SNR = 2. f) SNR = 1.

Kvantverkningsgraden anger hur stor procentuell andel av de fotoner som faller in mot en pixel som verkligen detekteras, dvs skapar elektroner. I en perfekt sensor skulle 100% av fotonerna detekteras, men i praktiken kan detta aldrig uppnås av flera skäl. En del ljus kommer att reflekteras eller absorberas i de elektroder och andra skikt som ligger ovanför utarmningsområdet (Fig. 14.3). En annan

förlustmekanism är att fotoner kan passera genom hela utarmningsområdet utan att absorberas. Denna förlustmekanism blir mer utpräglad ju längre våglängden är, och för $\lambda > 1000$ nm kommer praktiskt taget inga fotoner att absorberas eftersom deras energi är för låg för att slå loss elektroner. Kurvan i Fig. 14.14 är bara ett typexempel på hur kvantverkningsgraden varierar med våglängden; utseendet kan variera betydligt mellan olika tillverkare och modeller. Det finns t.ex. metoder att med fluorescerande skikt på ytan omvandla UV-fotoner till synligt ljus och därmed förbättra känsligheten i UV-området. En hög kvantverkningsgrad innebär att sensorn får en hög känslighet. Den kan då ge bilder med lågt brus och god dynamik redan vid låga exponeringar. Detta är naturligtvis alltid en fördel, men kanske framförallt i vetenskapliga tillämpningar när man ska avbilda ljussvaga fenomen, t.ex. inom astronomi eller fluorescensmikroskopi.

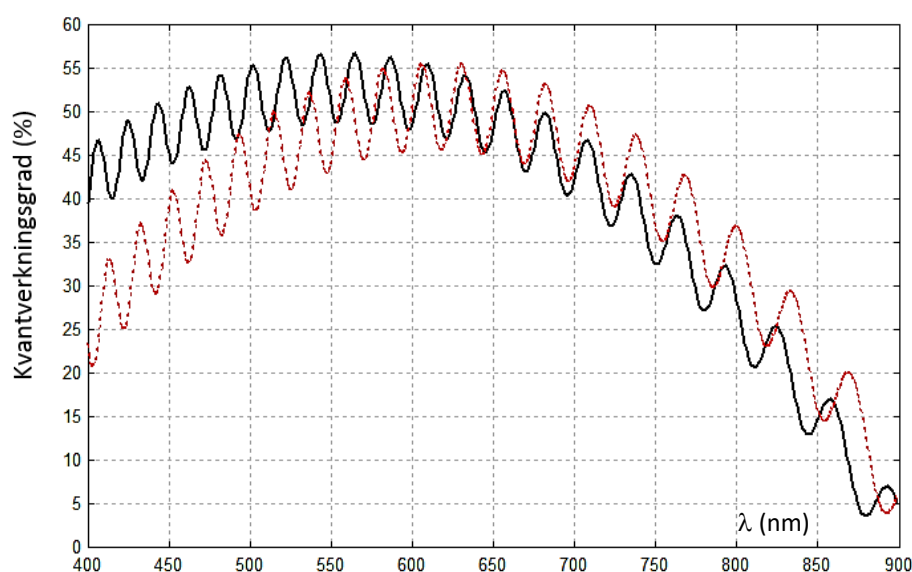


Fig. 14.14. Kvantverkningsgraden för en matrissensor kan typiskt ha det utseende som visas i den svarta heldragna kurvan. Oscillationerna beror på interferenseffekter i de tunna skikt som ligger ovanför den aktiva detektionsvolymen. Den röda prickade kurvan visar hur många elektroner som bildas i en pixel per sekund och per watt instrålad effekt (relativ skala). Skillnaden i hur kurvorna varierar med våglängden beror på att fotonerna är mer energirika vid korta våglängder.

Det finns ett radikalt grepp att förbättra kvantverkningsgraden, speciellt för korta våglängder. Metoden kallas "back side illumination" (BSI) och innebär att man gör det p-dopade materialet mycket tunt och belyser sensorn från "undersidan", Fig. 14.15. Härigenom kan man uppnå en kvantverkningsgrad på 70% eller mer över större delen av det synliga våglängdsområdet. I konsumentprodukter har BSI mest använts för små sensorstorlekar, men under 2015 introducerades för första gången en kamera (Sony A7RII) med en BSI-sensor i fullformat.

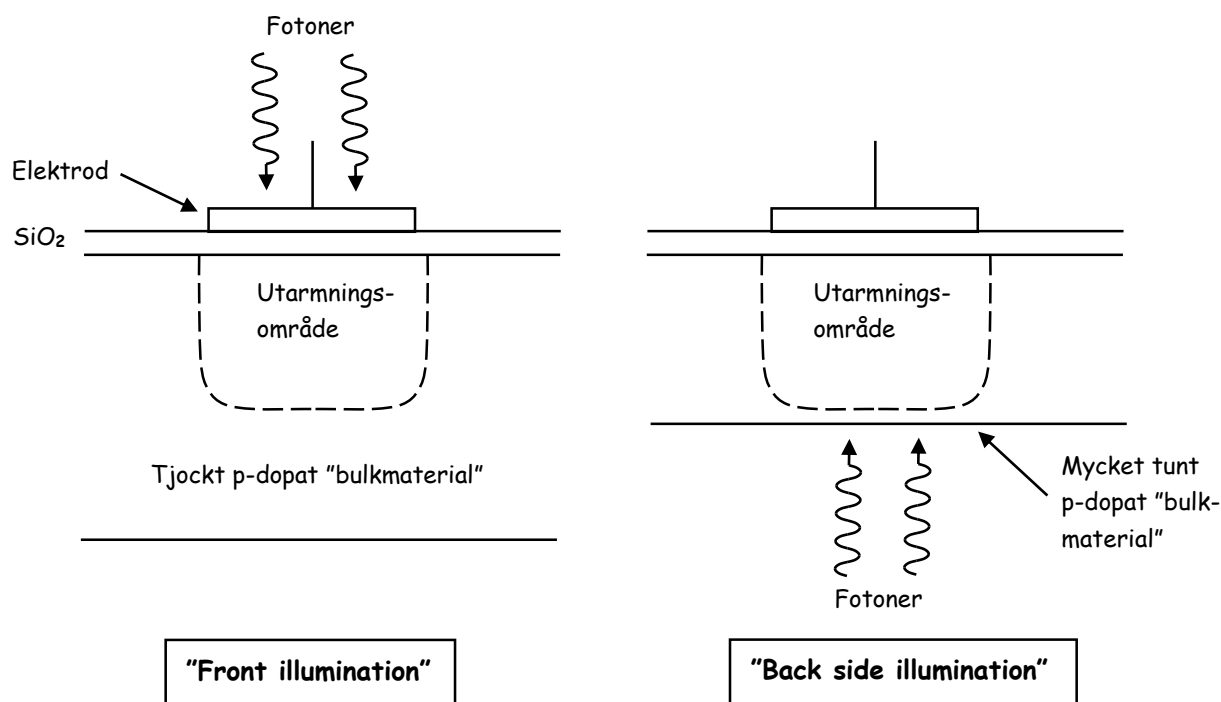


Fig. 14.15. Genom att belysa en pixel från "baksidan" kan man undvika reflektion och absorption i elektroder och kiseldioxidlager. Substratet ("bulkmaterial") görs mycket tunt för att minimera ljusabsorption.

14.6. Linearitet och gammakorrektion

De sensorelement (photogates eller fotodioder) som används i matrissensorer har i allmänhet en mycket god linearitet, dvs. utsignalen kommer att växa linjärt med exponeringen. Detta skiljer sig markant från fotografisk film som är kraftigt icke-linjär. Men om man studerar pixelvärdena i bilden från en digitalkamera, så kommer man att upptäcka att dessa vanligen inte är proportionella mot exponeringen. Detta beror på att man i kameraelektroniken gör en så kallad gammakorrektion av pixelvärdena*. Det finns flera skäl till varför man "förstör" sensors goda linearitet. Första skälet (som idag är historiskt) är att man ville att bilderna skulle återges med en naturtrogen grå- och färgskala på en CRT (Cathode Ray Tube, dvs. elektronstrålerör för TV eller datorskärm). En CRT är nämligen kraftigt icke-linjär, och skärmens luminansvärde, L , beror (något förenklat) på styrgallerspänningen (i princip pixelvärdet p) enligt formeln $L = \text{konst} \cdot p^\gamma$, där γ -värdet vanligen är mellan 2 och 2.5. Sedan gammalt har man därför använt en så kallad gammakorrektion i videosignalen som matar CRT:n. Och detta system tillämpar man också på digitala bilder. Sålunda omräknas de ursprungliga pixelvärdena, p_0 , till nya värden, p , genom en funktion som ungefärligen kan beskrivas med formeln

$$p = \text{konst} \cdot p_0^{\gamma'} \quad (14.3)$$

* Bilder som levereras av kameran i RAW-format brukar inte vara gammakorrigerade. De uppvisar därför en god fotometrisk linearitet, se <https://www.kth.se/social/files/5c88c38056be5bb29067e3b0/Tonreproduktion.Sony%20A7.pdf>.

där $\gamma' \approx \frac{1}{\gamma} \approx 0.3 - 0.5$. Genom detta arrangemang kommer CRT-skärmens luminansvärden att bero approximativt linjärt på motivets luminansvärden. I Fig. 14.16 visas både CRT-karakteristiken och gammakorrektionens funktion.

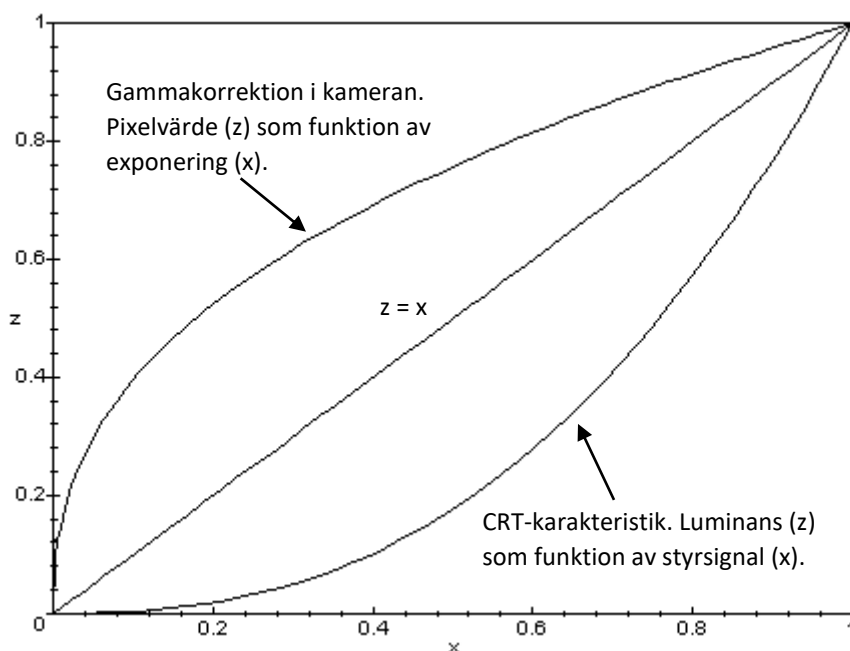


Fig. 14.16. Figurer på CRT-karakteristik och kamerans gammakorrektion av pixelvärdena. Gammavärden på 2.5 (CRT-skärm), 1 och 0.4 (kamera) illustreras.

Eftersom CRT-skärmar har ersatts av LCD- eller LED-skärmar behövs inte längre gammakorrektion för att få god linearitet på skärmen. Det finns emellertid även andra skäl till att utföra gammakorrektion av pixelvärdena. Ett skäl är att man med gammakorrektion kan klara sig med färre bitar i varje pixel, utan att vi med våra ögon kan uppleva några "hack" i en kontinuerligt varierande gråskala (se kap. 14.7). Skälet till det står att finna i hur våra ögon fungerar. För att vi ska uppleva en skillnad mellan två grånivåer, så krävs att luminansvärdena skiljer sig med ett visst **procenttal**. Detta innebär att vi i absoluta tal blir mer känsliga för luminansförändringar vid låga luminansvärden än vid höga. Den diskretisering som digitaliseringen medför ska därför ha en tätare indelning vid låga ljusvärden än vid höga. Detta är precis vad som uppnås om man gör en gammakorrektion. Ett ytterligare skäl att utföra gammakorrektion är att man bättre kan utnyttja sensorns dynamiska vidd i de fall man jobbar med 8-bitars bilder (t.ex. jpeg-bilder). Detta kommer att beröras lite närmare i kap. 14.7. För att förhindra att den gammakorrigerade kurvan (Fig. 14.16) blir alltför brant vid låga x-värden, brukar man ibland begränsa lutningen nära origo. Även andra avvikelser från det matematiska uttrycket för gammakorrigering som ges av ekv. 14.3 förekommer ofta, vilket vi snart ska se exempel på.

I definitionerna på färgrymderna sRGB och AdobeRGB (se kap. 12.6) finns beskrivet hur gammakorrektionen av jpeg-bilder ska utföras. Med användande av

samma normerade skalor som i Fig. 14.16, så gäller följande överföringsfunktion för sRGB

$$\begin{aligned} z &= 12.92 \times x, \text{ för } 0 \leq x \leq 0.00313 \\ z &= 1.055 \times x^{\frac{1}{2.4}} - 0.055, \text{ för } 0.00313 < x \leq 1.0 \end{aligned} \quad (14.4)$$

medan AdobeRGB föreskriver en ren gammakurva av den typ som ges av ekv. 14.3, och med ett gammavärde av $\frac{1}{2.2}$ (0.45). I praktiken ser gammakorrektionen för sRGB och AdobeRGB mycket lika ut, utom alldeles nära origo där lutningskoefficienten begränsas till 12.92 för sRGB medan den går mot oändligheten för AdobeRGB.

Om man undersöker hur gammakorrektionen implementeras i verkliga kameror, så kan man hitta väsentliga skillnader jämfört med definitionerna i ekv. 14.4. I Fig. 14.17 visas tonöverföringen för kameran Sony A7. Som synes är avvikelserna från specifikationen för sRGB avsevärd. Publicerade resultat från andra kameror visar liknande resultat. Effekten av dessa avvikelser är att kameran kommer att förstärka upp kontrast och ljushet i mörkare bildpartier*, medan kontrasten dämpas i ljusare partier. Det betyder att man tydligare ser detaljer i mörka partier, medan återgivningen av ljusa partier blir mer begränsad. En annan effekt är att man får en ”mjuk botten” av signalen när man närmar sig överexponering. Detta gör att man undviker en abrupt klippning av signalen vid en viss ljusnivå, vilket kan ge fula effekter i bilden.

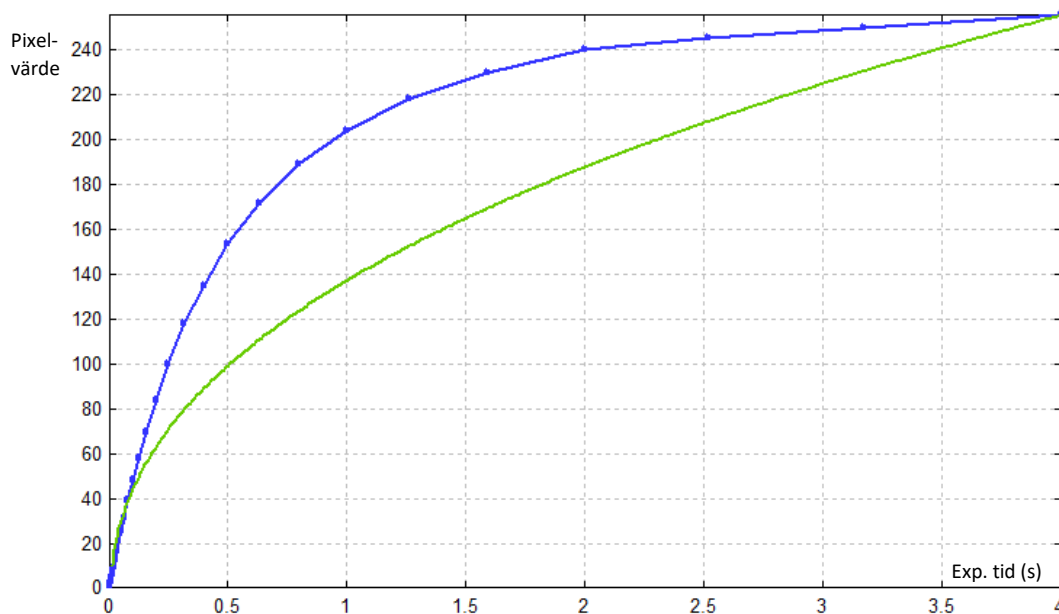


Fig. 14.17. Gröna kurvan visar gammakorrektion för jpeg-bilder enligt specifikationen för sRGB, medan blåa kurvan visar uppmätta värden för kameran Sony A7. Exponeringen varierades genom att ändra exponeringstiden vid konstant bländartal.

* Utom i de allra mörkaste, eftersom man inte vill förstärka upp mörkerbruset alltför mycket.

Den blå kurvan i Fig. 14.17 gäller för jpeg-bilder som levereras direkt ur kameran. Om man skapar jpeg-bilder via en RAW-konvertering skiljer sig kurvan bara marginellt i det aktuella fallet (bottningen blev ännu något mjukare).

14.7. Analog/digital omvandling

När ljus träffar en elektronisk sensor skapas i varje pixel fria laddningar (elektroner) i direkt proportion till exponeringen. Denna laddningsmängd konverteras elektroniskt till en spänningsnivå. För att kunna lagras digitalt måste emellertid spänningsnivån konverteras till ett binärt tal, vilket görs i en så kallad analog-till-digital omvandlare (på engelska: analogue-to-digital converter = ADC)[†]. En viktig parameter för en ADC är antalet bitar. Vanliga bit-antal i olika tillämpningar är 8, 10, 12, 14 och 16. I jpeg- och tiff-bilder som kommer från digitalkameror brukar antalet bitar vara 8 (i färgbilder 8 bitar i varje färgkanal). I bilder som levereras i RAW-format brukar antalet bitar ofta vara 12 eller 14.

Antalet bitar påverkar hur många olika grånivåer som kan återges i varje pixel. Detta illustreras lite förenklat i Fig. 14.18. Med 8-bitars ADC kan vi återge $2^8 = 256$ olika nivåer (0, 1, 2, ..., 255). För 10, 12, 14 och 16 bitar blir motsvarande antal nivåer 1024, 4096, 16384 och 65536. Ju fler bitar vi har i ADC:n, desto noggrannare kan alltså grånivåerna (eller färgerna) återges. Vid lågt antal bitar kan man visuellt uppleva en hackighet i bildens gråvärden på grund av att steget mellan två nivåer blir för stort. Hur ska man då kunna avgöra hur många bitar som krävs i olika sammanhang? En analys av detta görs i kap. 14.7.1 och 14.7.2. Men vi ska här göra en sammanfattning av resultaten.

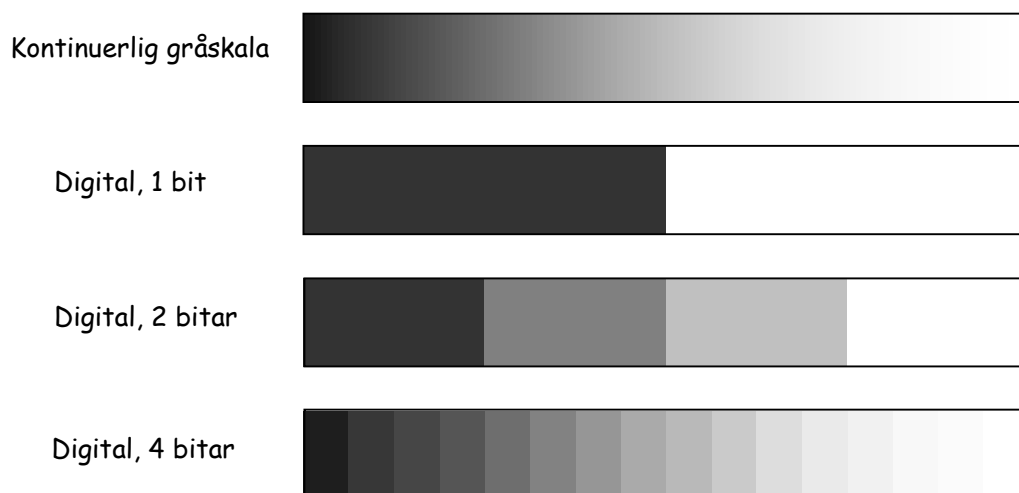


Fig. 14.18. Ju fler bitar som används i varje pixel, desto fler grånivåer kan återges. Om antalet bitar är n , så kan 2^n grånivåer återges. Det vanligaste för digitalbilder är $n = 8$, vilket är fallet för t.ex. jpeg-bilder.

[†] CMOS-sensorer har ofta flera tusen ADC:er som arbetar parallellt (en för varje pixelrad), vilket gör utläsningen av data snabbare.

Man kan lägga (åtminstone) två aspekter på hur många bitar per pixel som behövs. Den första är att bilden inte för ögat ska uppvisa några störande steg i gråskalan såsom är fallet i Fig. 14.18. Den andra är att antalet bitar inte alltför drastiskt ska påverka dynamiska vidden (se kap. 14.5) hos sensorn, dvs förmågan att samtidigt kunna återge detaljer i både mörka och ljusa partier av motivet.

När man undersöker den första aspekten, inga störande steg i gråskalan, så kommer man fram till att det lägsta acceptabla antalet bitar är ungefär 8 när bilderna visas på en dator- eller TV-skärm eller som en utskrift på fotopapper. Och detta endast under förutsättning att man använder en bra gammakorrektion (detta undersöks i kap. 14.7.2). Utan gammakorrektion behövs betydligt fler bitar. Vad gäller dynamiska vidden, så är det återigen mycket viktigt att använda en bra gammakorrektion för att inte förlusten ska bli för stor när man reducerar antalet bitar. Vi såg i kap. 14.5 att en typisk fullformatkamera kan ha en dynamisk vidd på ca. 15 000 i 14-bitars RAW-data. Om vi omvandlar 14-bitsignalen till 8 bitar med linjär skalning (alltså utan gammakorrektion), skulle vi få en dynamisk vidd av strax under 900. Detta är oacceptabelt lågt i många praktiska sammanhang. Om vi istället använder en bra gammakorrektion, så kan vi få en dynamisk vidd i 8-bitsignalen som är ungefär hälften av vad den är i 14-bitsignalen. Man vinner alltså nästan en tiopotens i dynamisk vidd i 8-bitsignalen genom gammakorrektion (för detaljer, se kap. 14.7.1). Den förlust som trots allt uppstår beror på digitaliseringsbruset som uppstår när man reducerar antalet bitar från 14 till 8.

14.7.1. Hur bildens dynamiska vidd påverkas av antalet bitar och gammakorrektion.

Dynamiska vidden avhandlades i kap. 14.5, och vi såg där att för en typisk fullformatkamera kan den ligga runt 15 000. Detta gäller för 14-bitars bilder i RAW-format. Ofta använder man emellertid jpeg-bilder som innehåller bara 8 bitar per färgkanal (pixelvärden 0-255). Vi ska nu titta på vad detta kan förväntas innebära för dynamiska vidden.

Vi kan börja med att titta på fallet att vi inte gör någon gammakorrektion, så att 8-bitarsbildens pixelvärden varierar linjärt med exponeringen precis som fallet är med RAW-bilden. Det teoretiskt bästa värdet vi då kan få, är om allt elektroniskt brus är försumbart så att endast digitaliseringsbruset finns kvar. Detta digitaliseringsbrus blir för en ideal ADC lika med $\frac{1}{\sqrt{12}}$ enheter*. Under dessa ideala omständigheter skulle vi få en dynamisk vidd av $\frac{255}{\frac{1}{\sqrt{12}}}$, vilket är strax under 900. Vi har alltså förlorat över en tiopotens i dynamisk vidd när vi går från 14 till 8 bitar beroende på digitaliseringsbruset som uppstår. Denna förlust kan minskas avsevärt med en icke-linjär skalning av pixlarna, dvs genom gammakorrektion. Fig. 14.19 visar ett tänkt fall där vi omvandlar 14-bitars RAW-data till 8-bitars gammakorrigerade pixelvärden.

* Detta finns härlett i ett kompendium tillgängligt på

<https://www.kth.se/social/files/57d11da0f276542d9216413a/Comp.ImagingPhysics.new.pdf>

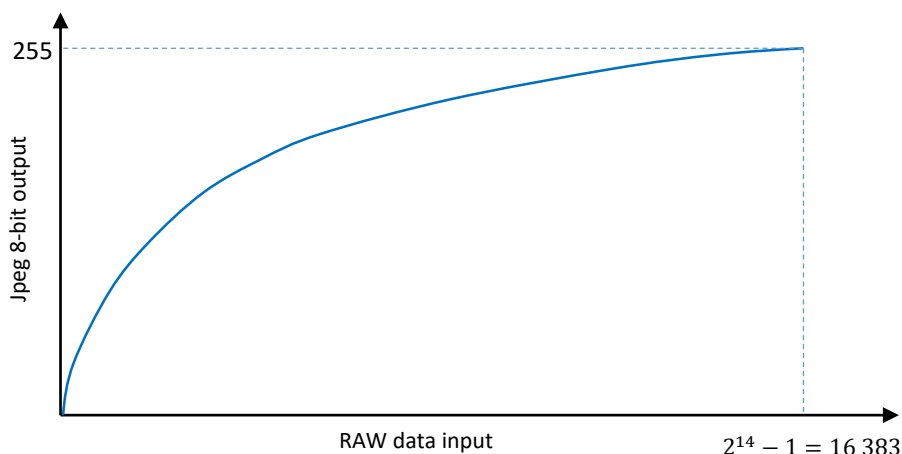


Fig. 14.19. Vid omvandling från RAW-data till jpeg minskas antalet bitar till 8. För att inte allt för stora förluster ska göras vid omvandlingen används en icke-linjär så kallad "gammakurva".

Gammakorrektionen i sig påverkar inte dynamiska vidden, eftersom den inte påverkar hur mycket utsignalen från sensorn behöver öka över mörkernivån för att överstiga read-out noise; brus och signal skalar ju på samma sätt via gammakurvan. Däremot kommer digitaliseringsbruset som uppstår vid 8-bitarskonverteringen att få mindre inverkan vid användning av gammakorrektion. Konverteringen från 14 till 8 bitar sker genom datorberäkningar (ingångsdata är ju redan digitala), och därför kan vi nog anta att digitaliseringsbruset i 8-bitarssignalen ligger nära det teoretiskt minsta, dvs $\frac{1}{\sqrt{12}}$ enheter. Hur stor inverkan digitaliseringsbruset får på dynamiska vidden beror på gammakurvans lutning nära origo, närmare bestämt vid den nivå som motsvarar mörkervärdet (vilket i sin tur beror på mörkersignal och elektronisk offset). Fig. 14.20 visar en delförstoring nära origo av samma gammakurva som i Fig. 14.19.

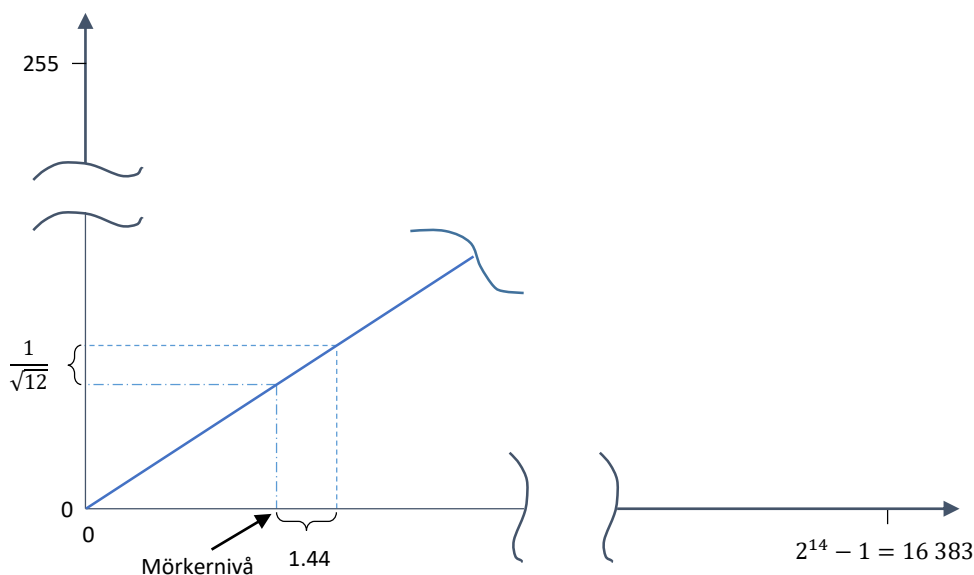


Fig. 14.20. Uppförstorad detalj av området nära origo i Fig. 14.19. Gammakorrektion enligt sRGB antas. En ökning av 1.44 enheter över mörkernivån i 14-bitsignalen motsvarar precis digitaliseringsbruset i 8-bitsignalen. Se texten för detaljer.

Ju högre lutningskoefficient kurvan i Fig. 14.20 har, desto mer kommer små luminansvariationer att förstärkas upp när vi konverterar till 8-bitars pixelvärde. Vi kan ta ett konkret exempel för att illustrera detta. Anta att vi använder en gammakorrektion enligt sRGB, dvs. ekv. 14.4. Nära origo har denna en konstant lutning 12.92 för normerade skalor på axlarna (jfr. Fig. 14.16). Om vi istället använder axelskalor enligt Fig. 14.19 och 14.20 så blir motsvarande lutning $12.92 \times \frac{255}{16383} = 0.20$. Detta innebär att om vi ökar 14-bitsignalen med 5 digitala enheter (av totalt 16 383) så ökar 8-bitsignalen med 1 enhet. För att öka 8-bitsignalen med $\frac{1}{\sqrt{12}}$ enheter (dvs lika mycket som digitaliseringsbruset), så ska alltså 14-bitsignalen ökas $\frac{5}{\sqrt{12}} = 1.44$ enheter. Detta innebär att digitaliseringsbruset i 8-bitsignalen motsvarar ett brus av 1.44 enheter i 14-bitsignalen. Om detta är det enda bruset skulle det ge en dynamisk vidd av $\frac{16383}{1.44} = 11\,300$. Genom gammakorrektion har vi alltså lyckats öka den teoretiskt maximala dynamiska vidden från knappt 900 till över 11 000 (i bägge fallen antog vi att endast digitaliseringsbrus var närvarande). I praktiken måste även read-out noise i RAW-bilden tas med i beräkningen, och som vi såg i kap. 14.5 ligger det typiskt på ungefär samma nivå som digitaliseringsbruset vid 8-bitskonverteringen (närmare bestämt 1.21 enheter för Sony A7 som vi tittade på). Om vi adderar dessa brusbidrag får vi en brusnivå på $\sqrt{1.44^2 + 1.21^2} = 1.9$ enheter*. Då hamnar dynamiska vidden i 8-bitarsbilden med gammakorrektion på $\frac{16383}{1.9} = 8700$. Detta är nästan exakt en tiopotens högre än man kan uppnå för en 8-bitarsbild med linjär skalning.

Nu får man inte ta de här angivna siffervärdena alltför bokstavligt. Dynamiska vidden beror på lutningskoefficienten nära origo för den typ av gammakorrektion som används, och som vi sett i Fig. 14.17 kan gammakurvor variera en hel del. För Sony A7 har gammakurvan en lutning av 7.6 med normerade axlar, inte 12.92 enligt sRGB som vi räknade med ovan. För Sony-kameran innebär det 70% högre digitaliseringsbrus än ovan, dvs. ca. 2.5 enheter istället för 1.4. Totalt brus blir då 2.7 enheter, och dynamiska vidden $\frac{16383}{2.7} \approx 6000$. Olika kameror och olika RAW-konverterare kan säkerligen ge en dynamisk vidd som varierar en hel del för jpeg-bilder. Utgående från de numeriska resultat vi erhållit, kan vi dock som en grov tumregel säga att vi kan förvänta oss ungefär en halvering av dynamiska vidden när vi går från 14-bitars RAW-format till 8-bitars jpeg.

Varför ökar man inte på gammakurvans lutning ordentligt nära origo (större än både 7.6 och 12.92), detta skulle ju innebära en högre dynamisk vidd? Men då förstärks också sensorns read-out noise upp mer, och kan bli störande i bilden. Dessutom kan mörkersignal och DC-offset i sensor och elektronik resultera i för stor offset i 8-bitsignalen, och även störande steg i gråskalan kan dyka upp. Det sistnämnda kommer att diskuteras i nästa delkapitel.

* Brus adderas inte linjärt, utan som roten ur summan av kvadraterna (se ekv. 14.2).

14.7.2. Att undvika synliga steg i tonreproduktionen.

Som vi såg i Fig. 14.18 så återges skiftningar i bildernas grånivå på ett bättre sätt när antalet bitar per pixel ökar. Vid för få bitar får man plötsliga steg i grånivån som inte har någon motsvarighet i motivet. Sådana osnygga steg får man naturligtvis även i färgbilder, vilket illustreras i Fig. 14.21. Denna steg-effekt kallas "posterization" eller "banding" på engelska, och utnyttjas ibland för att skapa konstnärliga effekter. Vi ska i fortsättningen använda termen posterization för denna effekt.

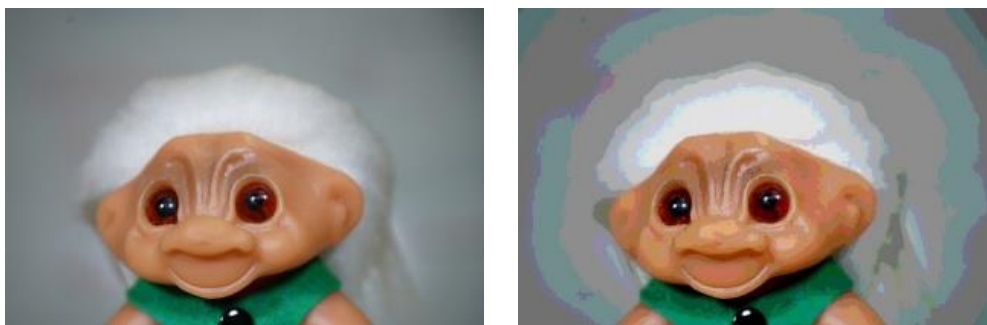


Fig. 14.21. Illustration av "posterization". Vänstra bilden innehåller 256 intensitetsnivåer per färgkanal, medan den högra endast har 10 nivåer per kanal.

I normala bilder vill man undvika synlig posterization, och vi ska nu undersöka vad som krävs för detta. Ordet "synlig" i föregående mening är värt att notera, eftersom posterization-effekten alltid är närvarande i digitala bilder på grund av diskretiseringen av intensitetsnivåerna. Det gäller bara att göra steget mellan närliggande nivåer så litet att ögat inte kan uppfatta det.

Detta för oss in på ögats förmåga att se skillnader i ljusnivåer. Som nämnades i kap. 14.6 så ska nivån ändras med ett visst procenttal för att vi ska se en skillnad. Undersökningar har visat att det krävs cirka 2% skillnad för att det ska vara synligt. Vid låga ljusnivåer krävs dock en successivt allt större procentuell skillnad för att vi ska upptäcka den (gränsen för när detta inträffar ligger runt luminansvärdet 1 cd/m² för en dator- eller TV-skärm). För att ta hänsyn till detta brukar man ibland undersöka både 2%-fallet och 5%-fallet.

Låt oss först titta på fallet att vi har en ren gammakurva av den typ som ges av ekv. 14.3, och med ett gammavärde av $\frac{1}{2.2}$ (0.45). Detta är fallet om vi följer den standard som gäller för AdobeRGB. Precis som i föregående kapitel ska vi anta att vi jobbar med jpeg-bilder, vilket innebär 8 bitar per pixel (i varje färgkanal)*. Pixelvärdena, p , kommer då att variera mellan noll och 255. Gammakorrektionen för AdobeRGB kan då skrivas

$$p = 255x^{0.45} \quad (14.5)$$

Där x -axeln är normerad till 0-1, och står för RAW-pixelvärde vilket motsvarar motivluminansvärdet med linjär skalning. Vi antar att bilden visas på en dator-

* Resonemangen i kap. 14.7.1 och 14.7.2 kan lätt anpassas för andra bit-antal om detta skulle vara önskvärt.

skärm, och vill då ha en skärmluminans som är proportionell mot motivluminansen (man gör ofta avsteg från detta i praktiken, men det är en bra utgångspunkt att bilden ska stämma överens med motivets grå- och färgskala). För att få detta ska skärmen omvandla pixelvärden, p , till luminansvärden, L , via ekvationen

$$L = \frac{p^{2.2}}{255^{2.2}} \quad (14.6)$$

Där L är normerad till intervallet 0-1. Vi ska nu undersöka hur stor procentuell skillnad vi får i skärmluminans när pixelvärdet ändras med en enhet, dvs. vi ska beräkna $\frac{\Delta L}{L}$ för $\Delta p = 1$. Genom att derivera ekv. 14.6 med avseende på p , och omforma uttrycket under förutsättningen att $\frac{dL}{dp} \approx \frac{\Delta L}{\Delta p}$ för $\Delta p = 1$, så erhålls

$$\frac{\Delta L}{L} = \frac{\Delta p}{255 \times 0.45 \times L^{0.45}} = \frac{1}{255 \times 0.45 \times L^{0.45}} \quad (14.7)$$

Vilket visar relativa förändringen i skärmluminans när pixelvärdet ändras en enhet. Om vi sätter som gräns att $\frac{\Delta L}{L} \leq 0.02$ så får vi att $L \geq 0.16$. Dynamiken i bilden blir då $\frac{1}{0.16} = 6.3$ (eller 2.7 bländarsteg som man ofta säger i fotosammanhang). Detta låter ju extremt dåligt om man betänker att dynamiska vidden för sensorn brukar ligga i storleksordningen 10 000 eller 13 bländarsteg (kap. 14.7.1). Nu är som redan påpekats 2%-gränsen kanske lite väl snäv när man kommer ner till låga L -värden, och om vi istället sätter $\frac{\Delta L}{L} \leq 0.05$ så får vi $L \geq 0.021$, dvs en dynamik av $\frac{1}{0.021} = 48$, vilket motsvarar 5.6 bländarsteg. Bättre, men fortfarande långt ifrån dynamiska vidden för sensorn (och det mänskliga ögat). Som parentes kan nämnas att utan gammakorrektion (ersätt 0.45 med 1.00 i ekv. 14.7) så blir dynamiken i 2%-fallet 5.1 och i 5%-fallet 13. Så gammakorrektionen förbättrar bevisligen dynamiken, men resultatet verkar ändå skrämt. Men innan vi hemfaller åt melankoli, bör vi ställa oss två frågor. För det första, vilken dynamik klarar en datorskärm (och utskrifter) av att visa? För det andra, kan man med en annan typ av gammakorrektion än AdobeRGB erhålla en bättre dynamik? Låt oss börja med den första frågan.

Enligt flera undersökningar kan man förvänta sig att en normal dator- eller TV-skärm har en dynamik (mörkaste till ljusaste som kan återges) av 50-100, dvs ca. 6 bländarsteg (och detta gäller bara om omgivningen är mörk). Det finns betydligt bättre skärmar (dyra), men typiskt hamnar man nära det angivna intervallet. För utskrifter på papper, så brukar man tala om svärtningsomfång, och det ligger typiskt på 0-2, vilket motsvarar en reflektans mellan 1% och 100%, alltså en dynamik av 100. Detsamma gäller för fotopapper framställda med analog fotografisk teknik. Jämför vi dynamik-siffrorna för skärmar och utskrifter med beräkningarna ovan, så ser vi att en jpeg-bild enligt AdobeRGB nästan klarar av samma dynamik utan posterization. Av det som sagts förstår vi att normala "bildvisningsmedier" inte på långa vägar klarar av dynamiken hos en kamerasensor eller det mänskliga ögat. Då kanske någon undrar vad det är för vits med hög dynamisk vidd i bildfilen, om displayen bara klarar av bråkdelen av den vidden. Svaret är att displayen bara klarar att visa en liten del av bilddynamiken samtidigt, men genom omskalning kan man visa upp allt som finns i bildfilen. Man kan t.ex. skala om bilden så att informationen i skuggpartierna visas tydligt medan ljusa partier då blir helt kritvita

och bottnade. Sedan kan välja att selektivt studera mellanljusa respektive ljusa partier etc. Men man kan inte se all information som finns i bildfilen samtidigt på en bildskärm eller utskrift.

Nu till den andra frågan, kan man i den visade bilden erhålla större dynamik utan posterization med en annan typ av gammakorrektion? Nära till hands är då att prova gammakorrektion enligt sRGB (ekv. 14.4). Om man gör motsvarande utredning som för AdobeRGB, kommer man fram till att sRGB med 2%-gräns ger en dynamik av 6.9 (2.8 bländarsteg) och med 5%-gräns 63 (6.0 bländarsteg). En ganska marginell förbättring alltså. Om man tittar efter vad problemet är med de gammakorrekationer som hittills studerats, så kommer man fram till att det beror på ett icke-optimalt utnyttjande av antalet digitala nivåer. För höga L -värden ligger nivåerna onödigt tätt, dvs. $\frac{\Delta L}{L}$ blir onödigt litet. Man kan då fråga sig om det finns någon gammakorrektion som ger konstant $\frac{\Delta L}{L}$ oberoende av L -värdet. Betrakta fallet att vi ersätter gammakorrekturen i ekv. 14.5 med $p = 255(1 + \ln x)$. Då får vi istället för ekv. 14.6 att $L = e^{\left(\frac{p-255}{255}\right)}$, och $\frac{\Delta L}{L} = \frac{\Delta p}{255} = \frac{1}{255} = \text{konstant} = 0.4\%$ när pixelvärdet ändras en enhet. En bra idé kan det tyckas, men det skulle kräva att man använder negativa pixelvärden för $L < \frac{1}{e}$ och dessa behöver gå mot $-\infty$ för att komma ner till $L = 0$. Inte så realistiskt. Men vid lite högre luminanser skulle en logaritmisk funktion av typen ovan kunna fungera bra. Vid lägre luminansvärden kan man utnyttja en klassisk gammakorrektion av samma typ som i ekv. 14.5. På så sätt kan man utnyttja antalet digitala nivåer på ett mer optimalt sätt för att uppnå ökad dynamik i bildvisningen utan posterization. En sådan typ av gammakorrektion finns beskriven i en publikation från BBC*. Om den tillämpas i vårt fall skulle ekv. 14.5 ersättas med

$$p = \frac{255\sqrt{12x}}{2}, \text{ för } 0 \leq x \leq \frac{1}{12}$$

$$p = 255[0.1788\ln(12x - 0.2847) + 0.5599], \text{ för } \frac{1}{12} < x \leq 1 \quad (14.8)$$

Denna typ av gammakorrektion kallas ofta för "hybrid". Det visar sig att om man plottar upp denna hybridkurva, så stämmer den mycket bättre med den uppmätta kurvan från Sony A7 än kurvorna för sRGB och AdobeRGB. Mycket tyder därför på att kamerafabrikanterna har frångått sRGB och AdobeRGB vad gäller gammakorrektion för att minska risken för posterization i bilderna. Hur mycket bättre blir då dynamiken med hybridkurvan som kombinerar en logaritmisk funktion för höga L -värden med en klassisk gammafunktion för låga L -värden? Resultatet av beräkningarna visas i tabell 14.1. Tabellen visar att en kombination av logaritmisk och klassisk gammakorrektion (hybrid) ger den klart bästa dynamiken. Risken för posterization borde vara minimal under normala omständigheter om den används. Om man däremot använder en skärm med extra hög dynamisk vidd så räcker det helt enkelt inte med 8 bitar, man måste öka på antalet bitar och frångå jpeg-formatet.

* <https://www.bbc.co.uk/rd/publications/whitepaper309>

Tabell 14.1

Gammakorrektion	Dynamik, 2% gräns	Dynamik, 5% gräns	Motsvarande antal bländarsteg
Ingen (linjär skalning)	5.10	12.8	2.4 resp. 3.7
Enligt AdobeRGB	6.33	48.5	2.7 resp. 5.6
Enligt sRGB	6.94	62.6	2.8 resp. 6.0
Hybrid (ekv. 14.8)	19.5	122	4.3 resp. 6.9

Tabellen visar hur stor bilddynamik som kan utnyttjas utan att posterization (steg i grå- och färgskalan) blir störande. Två olika maximala steggränser är medtagna, 2% och 5%. I de flesta praktiska fall torde 5%-gränsen vara tillfyllest, eftersom 5% steg bara erhålls vid de allra lägsta luminanserna där ögat inte är så steg-känsligt. Som jämförelse kan nämnas att en ordinär dator- eller TV-skärm klarar av att visa en dynamik på 50-100, motsvarande ca. 6 bländarsteg.

En sak som vi inte gått in på överhuvudtaget, är att skärmar i allmänhet inte är inställda för att ge en linjär återgivning av motivets luminansvärden såsom antagits i t.ex. ekv. 14.6. Man brukar kalla detta för att man har ett system-gamma som avviker från ett. Detta medför ytterligare en parameter i systemet som försvårar analysen, eftersom önskvärt systemgamma beror på t.ex. omgivningsbelysningen när man betraktar en bildskärm. Vi går inte in på detta. De resultat som presenterats torde ändå ge en god fingervisning om vilka systemprestanda som kan förväntas.

Förutom att införa någon typ av gammakorrektion, så är ett huvudsyfte med jpeg-formatet att komprimera bilddata vid lagringen så att bilder upptar mycket mindre minnesutrymme. Detta kommer att beröras i kap. 23.2.

14.8. Färgregistrering

Såsom detektormatriserna hittills beskrivits kommer fotoner av alla våglängder att registreras av varje pixel. Detta gör att man inte får någon färginformation, utan bara en svartvit bild. Nästan undantagslöst vill man dock få en färgbild. Detta kan i princip åstadkommas på flera olika sätt. Ett sätt är att göra tre exponeringar av samma motiv där man använder tre olika färgfilter (rött, grönt och blått) framför objektivet. Ur de tre bilderna kan man sedan rekonstruera motivets färger. Denna metod används i vissa vetenskapliga sammanhang (t.ex. astronomi och mikroskopi), men är alltför opraktisk för vanlig fotografi. En annan metod är att inne i kamerahuset dela upp ljuset spektralt med hjälp av stråldelare så att man får tre stycken bilder i rött, grönt och blått som projiceras på varsin detektormatris (se Fig. 18.3 där motsvarande metod används vid infrarödfotografering). Denna metod är också opraktisk för vanlig fotografi, eftersom den innebär en klumpig och dyr kamera*. Istället belägger man pixlarna med olika färgfilter som bara transmittar rött, grönt eller blått ljus, Fig. 14.22. Det finns också matriser med gul-, magenta- och cyanfilter (samt en del andra varianter), vi ska återkomma till detta senare. Det

* I videokameror används däremot denna teknik ofta, eftersom man där använder små och mycket billigare sensormatriser.

man i samtliga fall är ute efter är emellertid att registrera hur stora röd-, grön- och blåkomponenter man har i det detekterade ljuset. Med kännedom om dessa komponenter kan man återskapa motivets färger på det sätt som beskrivits i kap. 12.

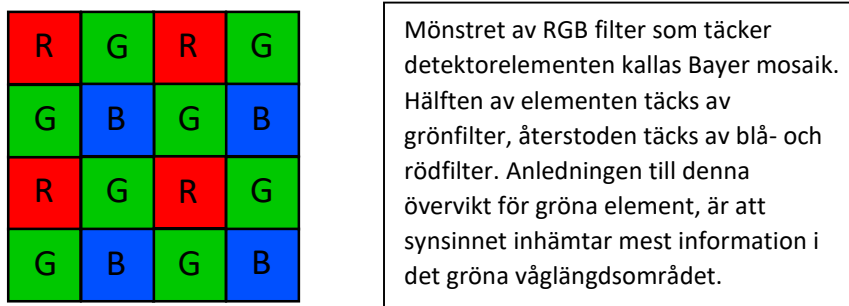


Fig. 14.22. Genom att täcka detektorpixlarna med filter i olika färger kan man registrera motivets färger.

Trots att varje pixel i sensorn bara registrerar en delfärg, kommer den digitala bilden som erhålls från kameran att innehålla RGB-värden för varje pixel. Hur är detta möjligt? Jo, man fuskar helt enkelt. De färgkomponenter som inte registreras av en pixel interpoleras fram ur grannpixlarnas värden med hjälp av kamerans inbyggda processor, se Fig. 14.23 och kap. 14.9. Det finns olika algoritmer för att göra detta, men självklart finns det risk att det uppstår felaktigheter. Dessutom kommer interpolationen att minska skärpan i bilderna. Många kameror kan leverera bilder i så kallat RAW-format. Detta innebär att ingen färginterpolation är gjord, och pixeldata erhålls med fler bitar (ofta 12 eller 14 istället för 8). Färginterpolationen kan då göras av användaren på en yttre, kraftfullare dator vilket kan ge bättre resultat (men fortfarande finns risk för felaktigheter). Det kan nämnas att det också finns en matrissensor (Foveon) som i varje pixel registrerar separat röd-, grön- och blåintensitet. Detta sker, precis som i färgfilm, i tre olika skikt på djupet. Man slipper då göra färginterpolation med de felaktigheter som det kan medföra. Nackdelen med Foveon är att det är mycket stor överhörning mellan RGB-värdena direkt från kameran. Därför måste man tillämpa en kraftig datorbearbetning för att få fram "renare" RGB-värden, vilket tar tid och dessutom höjs brusnivån markant. Det sistnämnda gör att fotografering vid dåliga ljusförhållanden blir mycket besvärligt.

I kapitel 13, ljuskällor, nämndes att man i en digitalkamera kan ställa in ljuskällans färgtemperatur (vitbalansvärde) för att få bra färgåtergivning med olika ljuskällor. Det som händer i kameran vid denna inställning är att man ändrar digitala förstärkningarna i röd-, grön- och blåkanalerna. Om man ställer in färgtemperaturen på ett lågt värde (t.ex. 2800K som är typiskt för en glödlampa) så är förstärkningen hög i blåkanalen, medelhög i grönkanalen och låg i rödkanalen. På detta sätt kompenserar man för att ljuskällan innehåller mycket rött ljus och lite blått ljus (se fig. 13.3). För bilder som fotograferas i RAW-format görs denna justering efteråt vid datorbearbetningen.

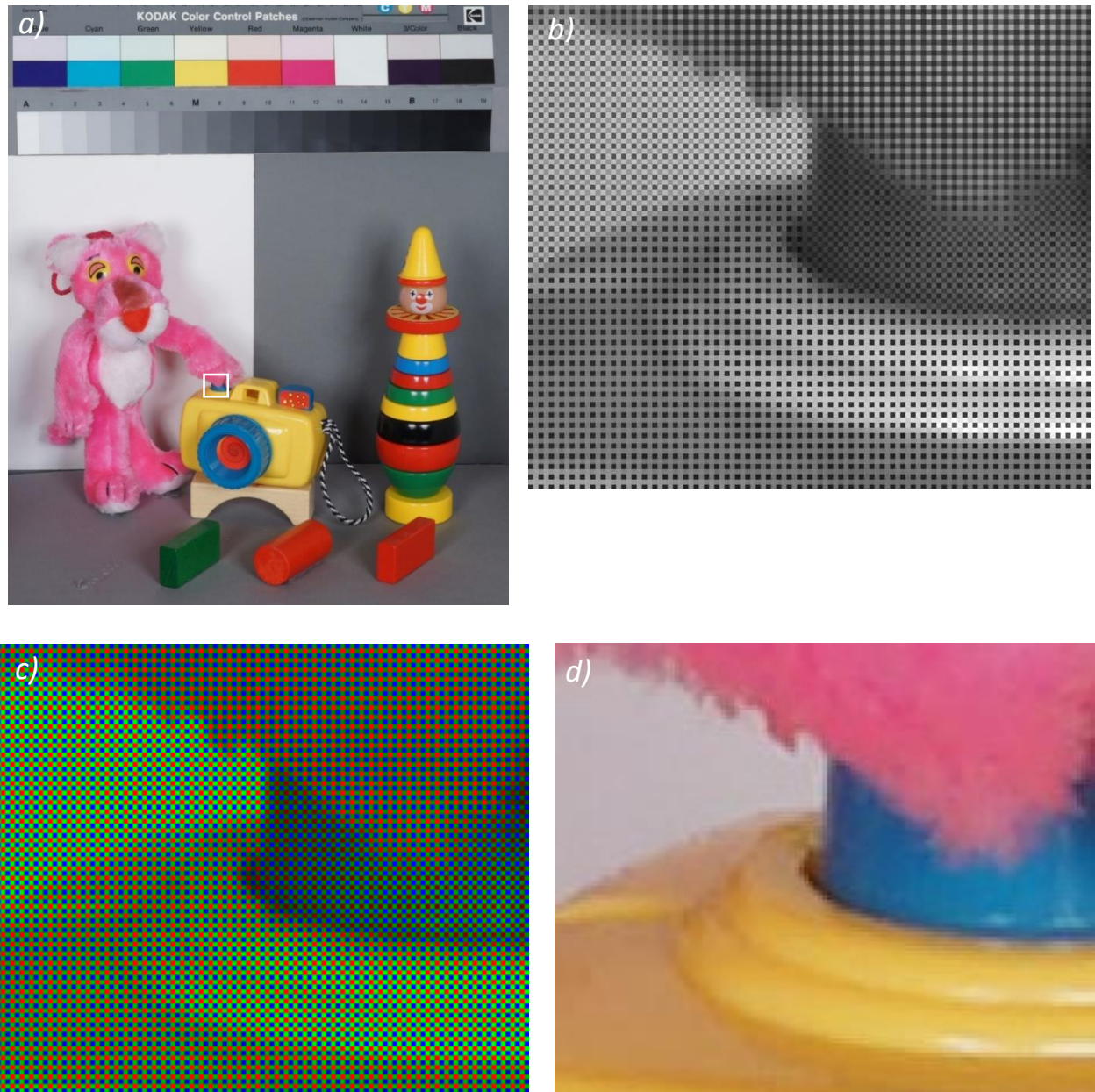


Fig. 14.23. Illustration av bildens utseende i olika stadier för en kamera med RGB Bayermask (det som visas är autentiska bilddata från en Sony A7). a) Motivet som fotograferades. Området markerat med vit fyrkant detaljstuderas i bilderna b-d. b) Detalj ur RAW-bilden direkt från kamerans sensor. Bilden är svartvit, eftersom varje pixel bara ger ett intensitetsvärde svarande mot antalet detekterade fotoner. c) Eftersom man vet vilka pixlar som är utrustade med R-, G- och B-filter kan b-bildens gråvärden översättas till RGB-intensiteter, men fortfarande innehåller varje pixel bara ett färgvärde. Om man zoomar in i bilden framträder strukturen i Bayer-mosaiken från Fig. 14.22 tydligt. d) Genom interpolationsräkningar kan de färgvärden som saknas i pixlarna "gissas" på ett intelligent sätt, så att varje pixel i den slutliga bilden (här i jpeg-format) innehåller alla tre färgkomponenter. Genom att ta ut bilden från kameran i RAW-format och utföra interpolationsräkningarna på en yttre kraftfullare dator, istället för på kamerans inbyggda processor, kan man i allmänhet nå ett bättre resultat.

14.8.1. Alternativa filtermasker

Som tidigare nämntes kan pixlarna täckas av filter med andra färger än R, G och B (även om det är dessa färgkomponenter man i slutändan är ute efter). En variant som ligger nära till hands är då att använda färgerna cyan, magenta och gult (CMY-filter), se kap. 12.3. CMY-värden kan omvandlas till RGB-värden genom:

$$\begin{aligned} R &= \frac{Y + M - C}{2} \\ G &= \frac{Y + C - M}{2} \\ B &= \frac{M + C - Y}{2} \end{aligned}$$

En motivation som ibland framförs för att använda CMY- snarare än RGB-filter, är att man då detekterar flera fotoner vilket är önskvärt om motivet är ljussvagt. Denna ökning beror på att CMY-filter transmitterar ca. 2/3 av synliga spektrat vardera, medan RGB-filter bara transmitterar ca. 1/3 (se Fig. 12.3 och 12.5). I medeltal kommer alltså bara 1/3 av antalet fotoner som faller in mot en RGB-sensor att detekteras, medan motsvarande siffra för en CMY-sensor är 2/3. Man kan därför undra varför inte alla kameror har sensorer med CMY-filter. Svaret är att det inte bara är signalstyrkan från sensorn som är av intresse, utan även brusnivån. Som vi sett i kap. 14.5 så är det signal/brus förhållandet (SNR) som bestämmer ”grynigheten” i en bild, och därmed hur klart man ser små detaljer. Vi såg också att när fotonbrus är den dominerande bruskällan, så blir SNR lika med kvadratroten ur antalet detekterade fotoner. Detta skulle innebära att var och en av C-, M- och Y-signalerna har ett SNR som (i medeltal) är $\sqrt{2}$ gånger högre än för R-, G- och B-signalerna. Men problemet är att när man ur CMY-värdena räknar fram RGB-värden enligt ekvationerna ovan, så kommer bruset i de tre signalerna att adderas (inte linjärt, utan kvadratisk*). Detta medför i slutändan att en bild registrerad med CMY-filter kommer att ha ett lägre SNR än en bild registrerad med RGB-filter. För detaljer om SNR-beräkningar, se kap. 14.8.2. Man får alltså en förlust i SNR, samtidigt som man är tvungen att utföra extra beräkningar i det fall man använder CMY-filter. Naturligtvis påverkas de verkliga SNR-värdena av filtrens transmissionskurvor (som inte är ideala), samt av andra bruskällor, men resultatet visar ändå att det inte finns något skäl att förvänta sig en vinst i SNR genom att använda CMY-filter. Men, kanske någon invänder, kan man inte utnyttja CMY-signalerna direkt för att skapa färgbilder utan att omvandla till RGB? Svaret är nej. Visserligen används CMY-färger vid analog färgfotografi och när man trycker bilder i böcker och tidningar, men vid dessa tillämpningar lägger man tre färgskikt på varandra (se kap. 12.3). Detta innebär att man kombinerar tre CMY-värden för att skapa RGB-värden, precis som när vi tidigare räknade fram RGB-värden. I slutändan behövs alltid RGB-värden, eftersom de ger information om ljusets intensitet i de röda, gröna och blå våglängdsområdena, dvs de områden som

* Totala brusnivån, σ_{tot} , mätt som standardavvikelse (RMS), ges av $\sigma_{tot} = \sqrt{\sigma_C^2 + \sigma_M^2 + \sigma_Y^2}$, där σ_C , σ_M och σ_Y är brusnivåerna i C-, M- och Y-signalerna.

påverkar de tre typerna av tappar i vår näthinna. CMY-värdena däremot innehåller ”oren” färginformation, dvs blandningar av flera olika våglängdsområden. Det är därför de måste räknas om.

Det finns dock ett sätt att betydligt förbättra SNR vid färgfotografering, och det är att använda tre matris-sensorer, en för varje delfärg RGB, och att spektralt dela upp ljuset i dessa tre våglängdsband i kameran (samma princip som Fig. 18.3, men för färgbanden RGB). På detta sätt kommer i medeltal tre gånger så många fotoner att detekteras, vilket ger ett SNR som är $\sqrt{3} = 1.73$ gånger högre än med RGB-filter. Som tidigare nämndes blir dock sådana kameror dyra, stora och tunga.

Förutom filtermasker i RGB och CMY, kan även andra färgkombinationer användas. Den vanligaste varianten är troligen CYGM, vilken är samma som CMY men med extra G-pixlar inlagda. Men precis som för CMY, ger CYGM sämre SNR i samtliga framräknade färgkanaler (RGB) jämfört med en RGB-mask. CMY- och CYGM-masker har tidigare använts av bl.a. Kodak, Nikon och Canon, men i moderna digitalkameror är RGB-masker standard. I nästa delkapitel, 14.8.2, görs en kvantitativ jämförelse av SNR för några olika typer av filtermasker som används, eller har använts, i digitalkameror.

Men innan vi går in på SNR ska vi titta lite på en annorlunda variant av Bayer filtermask som kallas Quad Bayer, och där pixlarna är ordnade i grupper om 2 x 2 i samma färg, Fig. 14.24.

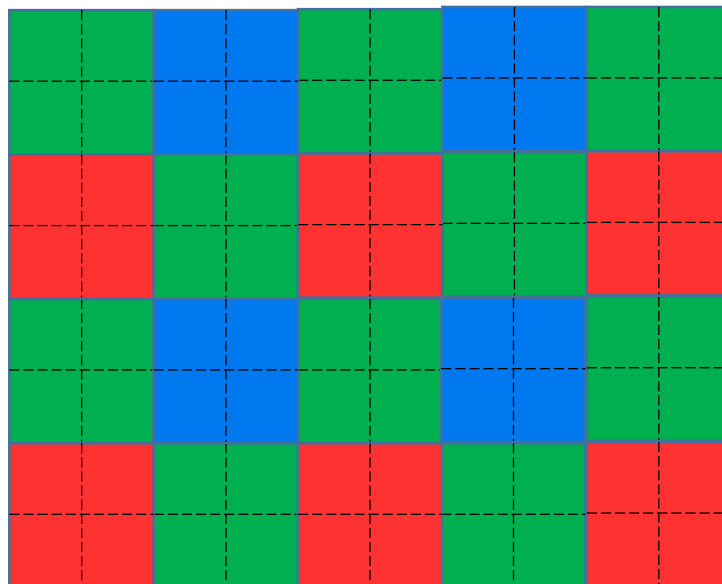


Fig. 14.24. Quad Bayer filtermask. Jämfört med standard Bayer, Fig. 14.22, är varje pixel uppdelad i 4 mindre delpixlar.

En Quad Bayer sensor kan användas i 3 olika moder:

1. Signalerna från pixlarna i varje 2×2 grupp summeras så att resultatet blir detsamma som från en enda stor pixel. Denna mod används vid dåliga ljusförhållanden då maximal känslighet krävs.
2. Pixlarna inom en 2×2 grupp exponeras samtidigt, men med olika exponeringstider. Vanligen får två av pixlarna en låg exponering och två får en hög exponering. Denna mod används för att öka kamerans dynamiska vidd.
3. Pixlarna i varje 2×2 grupp exponeras samtidigt och med samma exponeringstid. Alla delpixlarnas värden läses ut och färginterpolation utförs så att varje liten delpixel får RGB-värden. Denna mod används för att få maximal bildupplösning.

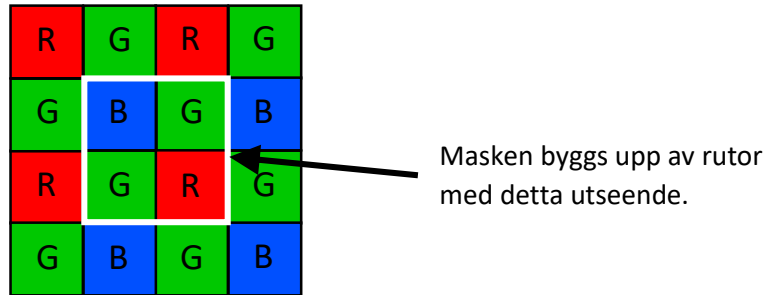
Quad Bayer används mest i kameror med liten sensoryta, typ mobilkameror. På grund av den lilla sensorstorleken får dessa kameror ofta problem vid dåliga ljusförhållanden och vid motiv som har stort kontrastomfång, se kap. 14.5. Vid dåliga ljusförhållanden kan man använda mod 1 som ger bättre signal-brus förhållande genom färre men större pixlar. Vid motiv som är tillräckligt ljusa, men som har stort kontrastomfång, kan man använda mod 2. Denna mod ger två bilder med olika kraftig exponering. Den minst exponerade används för att återge ljusa motivpartier, och den mest exponerade för att återge mörka partier. Sådana HDR-bilder (High Dynamic Range bilder) kan åstadkommas med många vanliga kameror genom att man gör två exponeringar efter varandra. Men problemet är då att rörliga föremål kan återges felaktigt på grund av att de rört sig mellan exponeringarna. Detta undviks genom Quad Bayer tekniken genom att bilderna tas samtidigt (men längsta exponeringstiden måste naturligtvis vara så kort att rörelseoskärpa undviks). Mod 3 slutligen används vid goda ljusförhållanden och med motiv utan alltför stort kontrastomfång. Man får då maximal upplösning motsvarande en normal Bayer sensor med pixlar i samma storlek som delpixlarna i Quad Bayer. Men färginterpolationen blir ogynnsammare i Quad Bayer fallet (färre grannpixlar i olika färger), vilket gör att man kan förvänta sig större interpolationsfel. Detta i kombination med begränsad SNR och dynamisk vidd för mod 3, gör att vissa kameror med Quad Bayer bara tillåter fotografering i moderna 1 och 2.

14.8.2. Jämförelse av signal-brusförhållandet för olika varianter av filtermasker.

Vi ska nu titta på vilket signal-brus förhållande (SNR) som är teoretiskt möjligt att uppnå för sensorer med olika varianter av färgmasker. Vi ska då anta att fotonbrus är den helt dominerande bruskillan, vilket är fallet i någorlunda ljusa bildpartier. Vidare ska vi anta att alla filter är ideala, dvs att de transmitterar de röda, gröna och blå spektralområdena som de ideala ”fyrkanter” som visas i Fig. 12.3. R-, G- och B-filter transmitterar alltså en sådan ”fyrkant” var, medan C-, M- och Y-filter transmitterar två stycken var (Fig. 12.5). Vi antar också att sensorn träffas av homogent ”vitt” ljus av sådan styrka att (i medeltal) N_0 fotoner detekteras i varje

pixel under exponeringstiden med användande av ett R-, G- eller B-filter* (och följaktligen $2N_0$ fotoner med C-, Y- eller M-filter♥). Vid analysen studeras SNR i de erhållna RGB-värdena för olika filtermasker jämfört med det SNR som skulle erhållas om inga fotoner gick förlorade på grund av absorption i filter. Vi börjar med att titta på den vanligaste typen, RGB Bayer mask.

RGB Bayer mask



Inom vita rutan finns $4N_0$ detekterbara fotoner (i medeltal) i vart och ett av färgbanden R, G och B. Idealt vill vi detektera alla dessa. Som beskrevs i kap. 14.5 gäller då att standardavvikelsen (bruset) blir lika med kvadratroten ur detta maxantal, alltså $\sqrt{4N_0}$. SNR är lika med medelvärde dividerat med standardavvikelse, vilket då blir $\sqrt{4N_0} = 2\sqrt{N_0}$. Låt oss nu titta på de verkliga SNR-värdena.

Hälften av pixlarna detekterar gröna fotoner. Inom vita rutan detekteras därför $2N_0$ fotoner (i medeltal, jag skippar denna kommentar i fortsättningen). SNR blir då $\sqrt{2N_0}$. Vi får alltså

$$\frac{SNR_G}{SNR_{ideal}} = \frac{\sqrt{2N_0}}{2\sqrt{N_0}} = \frac{1}{\sqrt{2}} = 0.71$$

Vi kan alltså få maximalt 71% av idealt SNR i grönkanalen med en Bayer RGB mask. För röda och blå pixlar gäller att de detekterar en fjärdedel av maximala antalet fotoner, dvs N_0 . Detta ger

$$\frac{SNR_{R,B}}{SNR_{ideal}} = \frac{\sqrt{N_0}}{2\sqrt{N_0}} = 0.50$$

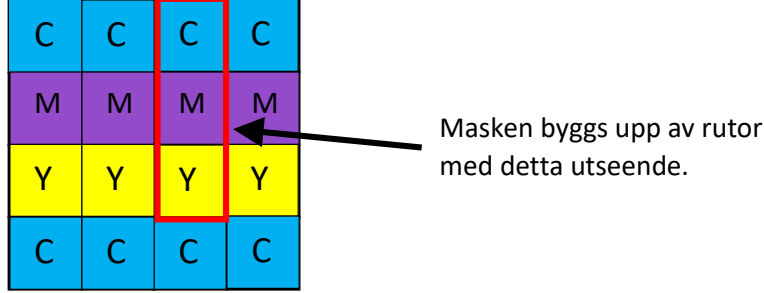
Vi kan alltså få maximalt 50% av idealt SNR i röd- och blåkanalerna med en Bayer RGB mask. I medeltal detekteras en tredjedel av fotonerna i denna mask, och man kan därför säga att vi i medeltal (sett över alla kanaler) får

* På grund av kvantverkningsgradens våglängdsberoende (Fig. 14.14) kommer då antalet *infallande* fotoner inte att se ut riktigt som det idealt vita i Fig. 12.3.

♥ Vi antar att exponeringen är lika hög för samtliga betraktade filtermasker, samt att vi aldrig överskrider "well capacity" för någon pixel. Lite senare ska vi se hur saker förändras om vi för de olika filterfallen exponerar sensorn just så kraftigt att man uppnår precis "well capacity", dvs att vi uppnår maximalt SNR oberoende av hur hög exponering som krävs.

$$\frac{SNR_{medel}}{SNR_{ideal}} = \frac{1}{\sqrt{3}} = 0.58$$

CMY-mask



Figuren som visas är bara en möjlig variant av CMY-mask, men andra layouter ger samma resultat vad gäller SNR. En vanligare variant av mask är CYYM, som innehåller dubbelt så många Y-pixlar jämfört med C och M. Denna typ kommer att behandlas i nästa avsnitt, men vi börjar med CMY, eftersom den är lite enklare att hantera begreppsmässigt.

Inom röda rutan finns $3N_0$ detekterbara fotoner i vart och ett av färgbanden R, G och B. Idealt vill vi detektera alla dessa. Då blir standardavvikelsen blir lika med $\sqrt{3N_0}$, och SNR blir också lika med $\sqrt{3N_0}$. Låt oss nu titta på de verkliga SNR-värdena.

Antalet detekterade fotoner i varje CMY-pixel blir $2N_0$, och standardavvikelsen i detta antal blir $\sqrt{2N_0}$. Som beskrevs i kap. 14.8.1 erhålls R-, G- och B-värden genom att kombinera C-, M- och Y-värden. Sålunda kan vi få rödvärdet ur $Y+M-C$. Denna summa har medelvärde $2N_0 + 2N_0 - 2N_0 = 2N_0$. Vid summeringen kommer standardavvikelseerna, σ , att adderas enligt $\sigma = \sqrt{\sigma_Y^2 + \sigma_M^2 + \sigma_C^2} = \sqrt{2N_0 + 2N_0 + 2N_0} = \sqrt{6N_0}$. Detta ger

$$SNR_R = \frac{2N_0}{\sqrt{6N_0}} = \frac{\sqrt{2N_0}}{\sqrt{3}}$$

Samma resultat kommer att erhållas för G- och B-värdena. Detta ger oss

$$\frac{SNR_{R,G,B}}{SNR_{ideal}} = \frac{\sqrt{2N_0}}{\sqrt{3}\sqrt{3N_0}} = \frac{\sqrt{2}}{3} = 0.47$$

Detta är lägre än för en Bayer RGB-mask trots att man detekterar dubbelt så många fotoner. Det beror på att man förstärker bruset när man räknar fram RGB-värden ur CMY-värden.

CYYM-mask (Kodak)

C	Y	C	Y
Y	M	Y	M
C	Y	C	Y
Y	M	Y	M

Masken byggs upp av rutor med detta utseende.

Inom röda rutan finns $4N_0$ detekterbara fotoner i vart och ett av färgbanden R, G och B. Idealt vill vi detektera alla dessa. Då blir standardavvikelsen blir lika med $\sqrt{4N_0}$, och SNR blir också lika med $2\sqrt{N_0}$. Låt oss nu titta på de verkliga SNR-värdena.

Antalet detekterade fotoner i varje CMY-pixel blir $2N_0$, och standardavvikelsen i detta antal blir $\sqrt{2N_0}$. Som för CMY-fallet erhålls R-, G- och B-värden genom att kombinera C-, M- och Y-värden. Sålunda kan vi få rödvärdet ur $Y + Y + 2M - 2C$. Denna summa har medelvärde $2N_0 + 2N_0 + 4N_0 - 4N_0 = 4N_0$. Vid summeringen kommer standardavvikelserna, σ , att adderas enligt

$\sigma = \sqrt{\sigma_Y^2 + \sigma_Y^2 + 4\sigma_M^2 + 4\sigma_C^2} = \sqrt{2N_0 + 2N_0 + 8N_0 + 8N_0} = \sqrt{20N_0}$. (När en signal multipliceras med 2 så multipliceras också standardavvikelsen med 2.) Detta ger

$$SNR_R = \frac{4N_0}{\sqrt{20N_0}} = \frac{2\sqrt{N_0}}{\sqrt{5}}$$

Motsvarande utredning för G- och B-värdena ger samma resultat, vilket medför att vi får

$$\frac{SNR_{R,G,B}}{SNR_{ideal}} = \frac{2\sqrt{N_0}}{2\sqrt{N_0}\sqrt{5}} = \frac{1}{\sqrt{5}} = 0.45$$

Detta är ungefär samma som för en CMY-mask (den hade värdet 0.47), och återigen hamnar vi alltså sämre än en RGB Bayer mask.

(Det kan tyckas märkligt att CYYM-masken inte ger högre SNR i de framräknade R- och G-värdena än för B-värdet, eftersom man detekterar fler röda och gröna fotoner än blå. Men man kan faktiskt visa att SNR för R- och G-värdena, samt B-värdet, sjunker kontinuerligt ju större övervikt man har för Y-pixlar. Detta beror på att det krävs en allt högre förstärkning av C- och M-signalerna för att kunna separera rött från grönt i Y-signalen, och detta gör att brusnivån höjs mer och mer. Däremot kommer nästa mask, CYGM, att ge bättre SNR i grönt eftersom den innehåller även rena G-pixlar.)

CYGM-mask (Nikon, Canon, Sony)

C	Y	C	Y
G	M	G	M
C	Y	C	Y
G	M	G	M

Masken byggs upp av rutor med detta utseende.

Denna mask är identisk med CYYM, förutom att hälften av de gula pixlarna ersatts med gröna. Inom röda rutan finns $4N_0$ detekterbara fotoner i vart och ett av färgbanden R, G och B. Idealt vill vi detektera alla dessa. Då blir standardavvikelsen blir lika med $\sqrt{4N_0}$, och SNR blir också lika med $2\sqrt{N_0}$. Låt oss nu titta på de verkliga SNR-värdena.

Antalet detekterade fotoner i varje CMY-pixel blir $2N_0$, och standardavvikelsen i detta antal blir $\sqrt{2N_0}$. För G-pixlarna blir motsvarande siffror N_0 och $\sqrt{N_0}$. R-, G- och B-värden erhålls som tidigare genom kombination av pixelvärdena. Grönt är ju lite speciellt, så vi börjar med det. Grönvärdet kan fås ur $G + Y + C - M$. Denna summa har medelvärde $N_0 + 2N_0 + 2N_0 - 2N_0 = 3N_0$. Vid summeringen kommer standardavvikelseerna att adderas enligt $\sigma = \sqrt{\sigma_G^2 + \sigma_Y^2 + \sigma_M^2 + \sigma_C^2} = \sqrt{N_0 + 2N_0 + 2N_0 + 2N_0} = \sqrt{7N_0}$. Detta ger

$$SNR_G = \frac{3N_0}{\sqrt{7N_0}} = \frac{3\sqrt{N_0}}{\sqrt{7}}$$

$$\frac{SNR_G}{SNR_{ideal}} = \frac{3\sqrt{N_0}}{2\sqrt{N_0}\sqrt{7}} = \frac{3}{2\sqrt{7}} = 0.57$$

Rödvärdet fås ur $Y + M - C$. Antalet detekterade fotoner blir $2N_0 + 2N_0 - 2N_0 = 2N_0$. Standardavvikelsen blir $\sigma = \sqrt{\sigma_Y^2 + \sigma_M^2 + \sigma_C^2} = \sqrt{2N_0 + 2N_0 + 2N_0} = \sqrt{6N_0}$. Samma siffror kommer att gälla för blåvärdet. Detta ger

$$SNR_{R,B} = \frac{2N_0}{\sqrt{6N_0}} = \frac{\sqrt{2N_0}}{\sqrt{3}}$$

$$\frac{SNR_{R,B}}{SNR_{ideal}} = \frac{\sqrt{2N_0}}{2\sqrt{N_0}\sqrt{3}} = \frac{1}{\sqrt{6}} = 0.41$$

Som förväntat ligger grönvärdets SNR bättre än för CMY- och CYYM-masker, medan röd- och blå-värdena har sämre SNR. I samtliga fall ligger emellertid värdena lägre än för en RGB Bayer mask.

Maximalt SNR för de olika filterfallen vid kraftig exponering

I ovanstående härledningar har vi utgått från en lika hög exponering, och sett hur högt SNR vi får för olika typer av filtermasker. Resultaten är av intresse när antalet fotoner är begränsat, t.ex. vid låga ljusnivåer eller när kort exponeringstid är nödvändig. Ofta är emellertid tillgången på fotoner god, och då spelar det mindre roll om filtren stjälar mycket ljus – man bara ökar på exponeringen tills sensorn nästan bottnar. Detta är fallet vi ska titta på nu. Vi antar att varje pixel kan detektera $N_{\max}$ fotoner, dvs att detta är ”well capacity”. Som tidigare jämförs erhållna SNR-värden med SNR_{ideal} , som i detta fall sätts lika med SNR som erhålls om varje pixel, i varje färgband RGB, detekterar $N_{\max}$ fotoner (detta skulle vara fallet för en kamera med tre sensorer enligt Fig. 18.3, men för färgbanden RGB). Härledningarna blir mycket likartade de som gjordes i tidigare avsnitt. Skillnaden blir att för masker innehållande CMY-filer byts $2N_0$ mot $N_{\max}$ vid beräkning av signal och brus (för CYGM-masken blir då G-pixeln bara halvfull innan övriga pixlar bottnar). För RGB-masken byts bara N_0 mot $N_{\max}$. Detta innebär att $\frac{SNR}{SNR_{ideal}}$ bibehålls för RGB-masken, medan den reduceras med faktorn $\frac{1}{\sqrt{2}}$ för alla masker innehållande CMY-filer. Alltså en ytterligare försämring för CMY-filer jämfört med RGB-filer. Vi får:

$$\text{RGB Bayer mask: } \frac{SNR_G}{SNR_{ideal}} = \frac{1}{\sqrt{2}} = 0.71 \text{ och } \frac{SNR_{R,B}}{SNR_{ideal}} = 0.50$$

$$\text{CMY-mask: } \frac{SNR_{R,G,B}}{SNR_{ideal}} = \frac{1}{3} = 0.33$$

$$\text{CYYM-mask: } \frac{SNR_{R,G,B}}{SNR_{ideal}} = \frac{1}{\sqrt{10}} = 0.32$$

$$\text{CYGM-mask: } \frac{SNR_G}{SNR_{ideal}} = \frac{3}{2\sqrt{14}} = 0.40 \text{ och } \frac{SNR_{R,B}}{SNR_{ideal}} = \frac{1}{2\sqrt{3}} = 0.29$$

Sammanfattning angående SNR för olika filtermasker

Som vi sett ur ovanstående analys, så ger alla studerade filtermasker som inbegriper CMY-filer ett lägre SNR än en RGB Bayer mask. Detta trots att CMY-filer gör att man detekterar fler av de infallande fotonerna. Problemet med CMY-varianterna, är att man vid omräkningen till RGB-värden kommer att förstärka upp bruset i signalen avsevärt. Det ska emellertid betonas att härledningarna som gjorts gäller under idealiserade förhållanden. Vi har t.ex. antagit ideala filteregenskaper, samt att andra brusbidrag än fotonbrus är försumbara (vilket i och för sig brukar vara fallet utom i mycket mörka bildpartier). Man ska därför inte ta siffervärdena ovan allt för bokstavligt, men de bör ändå kunna utgöra en god vägledning.

Om man tittar på sensorer i moderna digitalkameror så är RGB Bayer masker standard. Äldre kameror använde ibland andra typer. Kodak har t.ex. använt CYYM-masker i sina kameror, och Nikon, Canon och Sony har använt CYGM-masker i sina tidiga kameror. Kodaks kameratillverkning är numera nedlagd, och de övriga företagen har samtliga övergått till RGB-filter.

14.9. Färginterpolation

Som tidigare nämnts måste de färgkomponenter som inte direkt registreras av en pixel interpoleras fram ur värdena på grannpixlar. Låt oss anta att vi har en färgmask av Bayer-typ (den absolut vanligaste) och titta på en 3x3-omgivning runt en pixel vars värde vi vill interpolera fram. Vi får då 4 möjliga interpolationsfall som återges i Fig. 14.25 (jämför med Fig. 14.22).

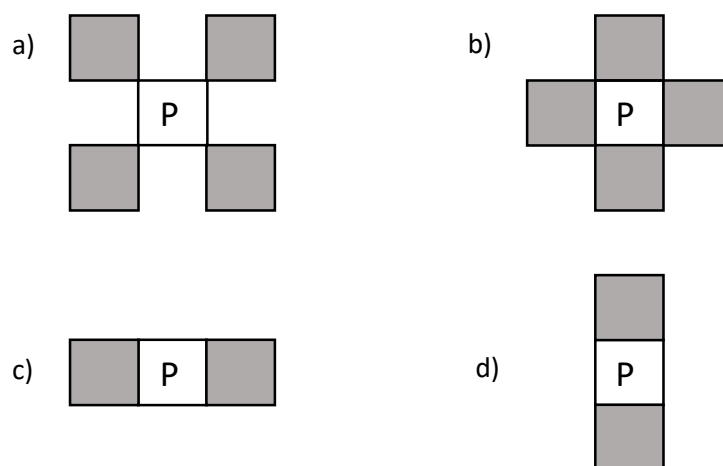


Fig. 14.25. Fyra interpolationsfall kan uppstå för en Bayer färgmask (Fig.14.22) om man betraktar en 3x3 omgivning till varje pixel. P markerar en pixel vars värde ska interpoleras fram ur de gråmarkerade pixlarnas värden. a) gäller för interpolation av blåpixlarnas rödvärden, samt för rödpixlarnas blåvärden. b) gäller för interpolation av grönvärdet för blå- och rödpixlar. c) och d) gäller för interpolation av grönpixlarnas blå- och rödvärden.

Den interpolationsmetod som kräver minst datorkraft kallas "nearest neighbour" (närmaste granne) och innebär helt enkelt att man tar värdet hos närmaste grannpixel som detekterar den önskade färgen. Den metoden passar illa för en Bayer-mask eftersom vi enligt Fig. 14.25 har två eller fyra närmaste grannar. Dessutom ger den ett mediokert bildresultat.

Den enklaste interpolationsmetod som förtjänar att övervägas är bilinjär interpolation. Den innebär att man i fallen a) och b) räknar fram medelvärdet av de fyra grå pixlarna, och i fallen c) och d) medelvärdet av de två grå pixlarna. Bilinjär interpolation ger betydligt bättre resultat än nearest neighbour, men fortfarande blir det en hel del felaktigheter vid t.ex. skarpa kanter och täta linjemönster. Att helt undvika sådana felaktigheter är omöjligt eftersom primärdata är undersamplat

(sampling behandlas i kap. 22), men man kan med mera avancerade metoder än bilinjär interpolation minska ner dessa felaktigheter.

Ett naturligt steg för att förbättra interpolationen är att betrakta en större omgivning runt varje pixel, samt utnyttja icke-linjär interpolation. Ett exempel på detta är bikubisk interpolation som vi inte ska gå närmare in på, utan vi konstaterar bara att det blir en ”mjukare” och inte så ”hackig” intensitetsfördelning i bilden. Sedan kan man ytterligare förfinas interpolationen genom att ta hänsyn till bildinnehållet, t.ex. att inte interpolera så kraftigt över kanter utan mera i en riktning parallellt med kanten. Man kan även utnyttja att det ofta finns en bildmässig korrelation mellan de olika färgkanalerna, se kap. 22.7.2. Ju mer sofistikerade metoder som används desto mer datorkraft krävs emellertid. Visserligen har kamerornas inbyggda processor blivit betydligt kraftfullare med tiden, men den sätter ändå en begränsning för hur avancerad interpolation som kan göras i de jpeg-bilder som levereras direkt ur en kamera. Om man tar ut bilderna i RAW-format så kan man emellertid göra färginterpolationen (och mycket annat) på en yttre kraftfullare dator, och därmed få ett bättre resultat.

Oavsett hur sofistikerad interpolation som görs, så innebär den alltid att resultatet blir sämre än om man i varje pixel hade registrerat alla tre färgkomponenterna. Försämringen yttrar sig på två sätt. Dels kommer felaktiga färger att kunna uppstå på vissa pixlar, t.ex. intill skarpa kanter, dels kommer bilden att bli lite oskarpare eftersom interpolationen medelvärdesbildar över flera pixlar (det senare berörs närmare i kap. 17.12). I Fig. 14.26 visas ett exempel på felaktiga färger som uppstått på grund av interpolation.

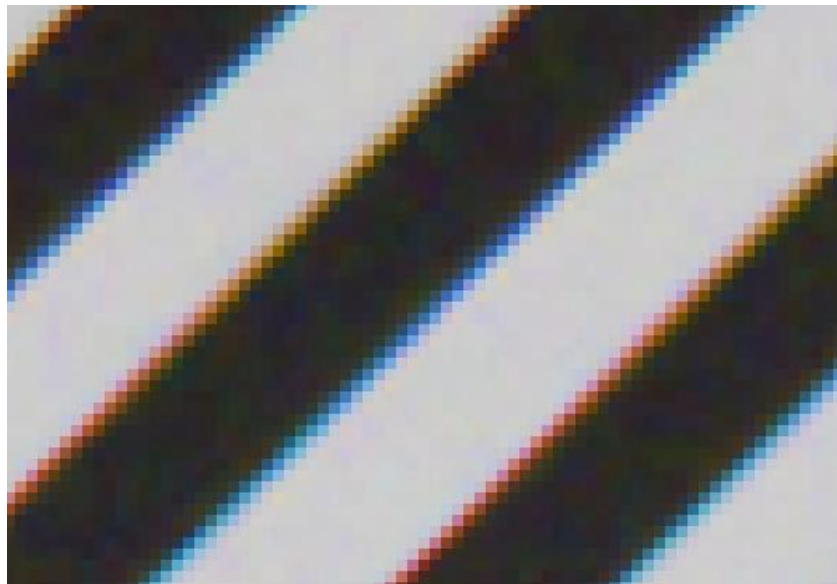


Fig. 14.26. Skarpa svartvita kanter återges ofta färgade i bilder där man utnyttjat färginterpolation. Men fenomenet är oftast störande bara när man zoomar in kraftigt i bilden så att man ser enskilda pixlar. Dessutom har färgmättnaden överdrivits i bilden ovan för att tydligt visa fenomenet.

I kapitel 14.8.2 tittade vi på signal-brusförhållandet (SNR) vid användande av olika typer av filtermasker. Vi fann då att Bayer-masken var fördelaktig vad gäller SNR.

Det är då naturligt att fråga sig hur färginterpolationen påverkar SNR. Som tidigare nämnts innebär interpolationen en medelvärdesbildning, och därmed kommer den att minska bruset och öka SNR. En närmare studie av detta kompliceras emellertid av att bruset i framinterpolerade pixelvärden kommer att uppvisa en korrelation, vilket gör att de analysmetoder som användes i kap. 14.8.2 inte är giltiga. Dessutom beror brusreduceringens storlek på vilken interpolationsmetod som används. Vi nöjer oss därför med att konstatera att bruset reduceras på grund av interpolationen.

14.10. Känslighet (ISO-tal)

ISO-tal används för att beskriva hur ljuskänsliga fotografiska filmer är. En films ISO-tal bestäms utgående från hur hög exponering som behövs för att få en godtagbar kontrast i den framkallade filmen. Denna metod är inte användbar för digitalkameror, eftersom CCD- och CMOS-sensorer ger god kontrast även för extremt låga exponeringar (till skillnad mot film). Det som istället sätter gränsen är bruset man får i digitalkamerabilder vid låga exponeringar. Tyvärr har man inte kunnat enas om en enda fundamental definition vad gäller ISO-talet för digitalkamerasensorer. Det finns t.ex. två olika ISO-tal som bygger på brusnivån, och ytterligare ett som bygger på hur hög exponering som behövs för att botten sensorn. Men dessa ISO-tal används mest i vetenskapliga sammanhang. Det vanligaste är att ISO-tal som anges av kamerafabrikanterna betyder att en film med samma ISO-tal som kamerans behöver lika hög exponering för att ge en slutbild som är lika ljus som digitalkamerans. Detta kriterium är mycket otydligt, och i praktiken omöjligt att mäta. Fabrikanterna kan alltså sätta sina egna ISO-tal mera på känsla än verkliga mätningar. Men rent praktiskt betyder det alltså att om man fotograferar med 100 ISO-film en solig sommardag, och korrekt exponering erhålls med bländartal 11 och exponeringstid 1/100 s, så kan man förvänta sig korrekt exponering också med en digitalkamera inställd på samma ISO-tal, bländartal och exponeringstid.

Det väsentligaste för en fotograf att veta är att ISO handlar om ljuskänslighet och att skalan är linjär, dvs fördubblat ISO-tal medför att exponeringen kan halveras.

ISO-talet på en digitalkamera kan ställas in på en mängd olika värden. Det lägsta brukar ligga runt 100, medan det högsta kan variera mellan ca. 3000 och 100 000 beroende på kameratyp. När man använder det lägsta ISO-talet, vi kan kalla det för "nominellt" ISO-tal, och exponerar bilden korrekt, så kommer pixlarna i de ljusaste motivpartierna att ligga nära maximal utsignal. Det som händer när man ökar ISO-talet, är att man ökar den elektroniska förstärkningen av sensorsignalen. Om man ökar ISO från 100 till 1000, innebär det att sensorn bara exponeras för 10% så många fotoner som tidigare. Man utnyttjar alltså bara 10% av sensorns "well capacity" (dvs maximala antalet fotoner som kan detekteras i en pixel). Som vi sett i tidigare avsnitt innebär färre fotoner en sämre bildkvalitet. Detta illustreras i Fig. 14.27. Både SNR och dynamiska vidden kommer att försämrats vid högre ISO-tal. En sensor med stor pixelarea är emellertid användbar vid betydligt högre ISO-tal än en kamera med liten pixelarea, eftersom antalet detekterade fotoner generellt sett är större för större pixlar.

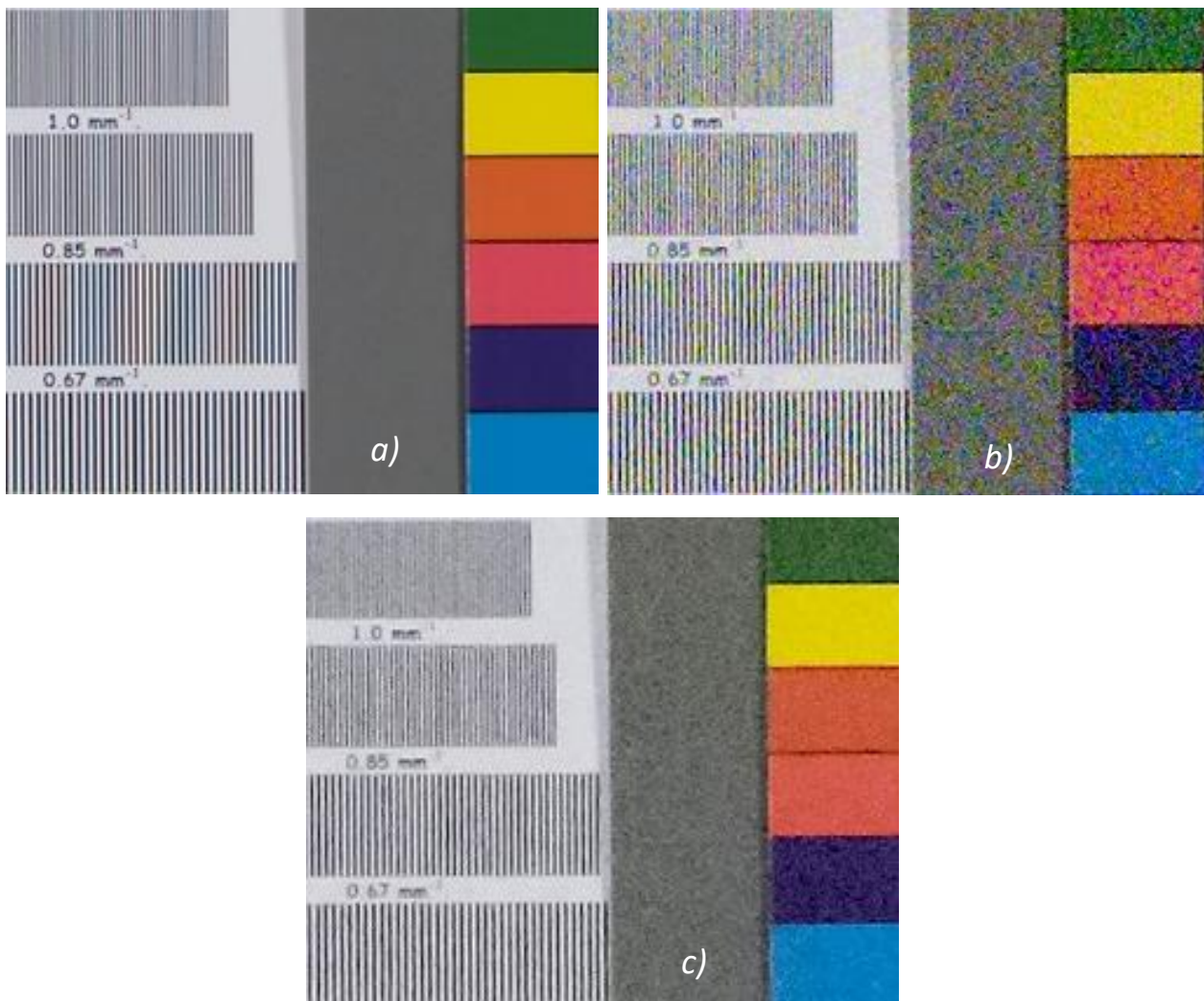


Fig. 14.27. Effekt av olika ISO-inställningar. a) ISO-inställning 100. b) ISO-inställning 25600. Det högre ISO-talet medför en kraftigt ökad brusnivå. Detta medför att det blir svårare att se små detaljer i motivet, samt att återgivningen av motivets mörkare delar försämras (bildens dynamik minskar). c) Samma bild som i b), men efter datorbehandling för att reducera inverkan av brus. Bruset är nu mycket mindre störande för ögat, men fortfarande är detaljåtergivningen, dynamiken och färgåtergivningen klart sämre än vid ISO 100. (Men det är ändå imponerande att man kan få så pass bra bildkvalitet som i c-bilden vid en ljusnivå som bara är 0.4% av vad den är i a-bilden!) Kameran som användes var en SONY A7(sensor 24 mm x 36 mm, 24 Mpixlar. Bilderna som visas är delbilder om ca. 350 x 300 pixlar.).

Ett gott råd vid fotografering är alltså att inte använda högre ISO-tal än nödvändigt. Men om ljusnivån är för låg, och risken för rörelseoskärpa gör att man inte kan öka exponeringstiden, är det en god idé att öka på ISO-talet. Som sagts ovan, är det i sådana situationer bra att ha en kamera med stor pixelstorlek, eftersom den ger en godtagbar bildkvalitet vid högre ISO-tal än om pixlarna är små.

Många kameror ger inte bara möjlighet att öka ISO-talet till högre värden än sensorns eget nominella värde, utan tillåter även att man sänker ISO-talet under detta värde. Kameran som tog bilderna i Fig. 14.27, SONY A7, har ett nominellt ISO-tal av 100. Det högsta som kan ställas in är 25 600, och det lägsta är 50. ISO-tal under 100 markeras med en särskild symbol som visar att kameran arbetar utanför sitt normala område*.

Om man använder ett ISO-tal som bara är hälften av sensorns nominella, så innebär det i praktiken att man överexponerar med en faktor två. Vad är vitsen med det kan man undra, blir inte bilden bara överexponerad (bottnad) och helt kass i ljusa partier? Nej, ofta blir den inte det. Sensorer har oftast en viss tolerans för överexponering, genom att de avviker från linearitet när man närmar sig bottning. De blir helt enkelt mindre ljuskänsliga nära bottning. Detta liknar det man i ljudsammanhang kallar ”soft clipping”, vilket används för att undvika kraftig distorsion när man bottnar en förstärkare. I praktiken visar det sig att man ofta kan exponera sensorn i en digitalkamera dubbelt, eller kanske fyrdubbelt, så mycket som exponeringsmätare eller histogram anger som korrekt exponering. Det man vinner på detta är att motivets skuggpartier kommer att återges bättre (mera detaljer och mindre brus). Det man kan förlora är kontraståtergivning i bildens ljusaste partier, eftersom dessa hamnar i sensorns ”soft clipping”-zon (eller i värsta fall bottnar helt). Man ska alltså iaktta viss försiktighet när man utnyttjar lägre ISO-tal än sensorns nominella. Det är också klokt att spara bilderna i RAW-format, annars händer ofta att pixelvärden bottnar vid kamerans jpeg-konvertering trots att sensorn inte är bottnad. Genom att ta ut bilderna i RAW-format, och göra all datorbearbetning i en RAW-konverter på en extern dator, så kan man ofta få fram mera detaljer både i motivets ljusa och mörka partier.

14.11. Videoinspelning och seriebildsupptagningar

De flesta digitalkameror kan också användas för videoinspelning. Bilderna som registreras har då vanligen betydligt färre pixlar än vid stillbildsfotografi. Många kameror tillåter också serietagning av flera stillbilder med full upplösning i snabb följd (men inte i videohastighet). Vi ska i detta delkapitel titta lite på typiska prestanda för denna typ av bildupptagningar. Låt oss börja med serietagning.

När man fotograferar t.ex. sport eller djur i rörelse kan det vara svårt att ta den önskade bilden i precis rätt ögonblick. Då är snabb seriefotografering ett utmärkt hjälpmedel. Man kan då under kort tid ta en stor mängd stillbilder i snabb följd, och sedan efteråt välja ut den bästa bilden. I en modern (2019) kamera av hög kvalitet, låt oss säga en fullformat systemkamera, kan man typiskt ta cirka 10 bilder per sekund med full pixelupplösning. Dessa bilder sparas i ett buffertminne i kameran som typiskt rymmer 50-60 bilder (alltså 5-6 sekunders serietagning). Sedan måste bilddata överföras till kamerans minneskort, vilket kan ta ca. en minut. Det kan

* Detta kallas på engelska för ”Extended ISO range”, och förekommer på vissa kameror både vid mycket låga och mycket höga ISO-tal. Olika kameror kan fungera lite olika, men ofta innebär ”extended range” att man vid ändring av ISO-talet förändrar signalförstärkningen (eller försvagningen för låga ISO) digitalt i kamerans processor (eller i en yttre dator för RAW-format). Detta sker alltså inte vid analoga förstärkningen, utan efter ADC:n (jfr. Fig. 14.12). Detta kan man tycka är lite fusk, eftersom precis samma resultat kan erhållas (åtminstone om man jobbar med RAW-format) genom att inte utnyttja extended range och istället medvetet felexponera bilden.

nämnas att åtminstone en fullformatskamera, Sony A9 II speciellt framtagen för sport- och naturfoto, tillåter serietagning med 20 bilder per sekund. Detta är möjligt eftersom sensorn har ”endast” 24 Mpixlar, medan många andra ”fullformatare” innehåller 50-60 Mpixlar. Dessutom har A9 II, förutom en vanlig mekanisk ridåslutare, även en elektronisk slutare (dvs. inga rörliga delar, se kap. 14.4) som kan användas vid serietagningar.

För videoinspelning erbjuder de flesta kameror ett antal olika videoformat. Men som jämförelse kan vi börja med att titta på specifikationen för gammaldags standard-TV (typ TV1, 2, 3, 4 etc.). Det finns här några vanligen använda och väletablerade standarder som betecknas PAL (Europa utom Frankrike, och stora delar av Afrika/Mellanöstern/Asien/Sydamerika samt Australien), NTSC (USA, Kanada, Västra Sydamerika och Japan) samt SECAM (Frankrike, forna Sovjetunionen och delar av Afrika). Vid analoga TV-sändningar används vanligen bredd/höjd förhållandet (aspect ratio) 4:3, medan man vid digital utsändning brukar använda 16:9. För att reducera den behövliga bandbredden vid marksänd TV introducerade man så kallad interlace-teknik. Det innebär att varje TV-bild som visas på skärmen bara innehåller hälften av bildinformationen. Varannan bild innehåller således bara bildlinjer med jämmt nummer, och varannan innehåller bara udda bildlinjer. Först när två delbilder har ritats upp i snabb följd ser alltså ögat (som har en viss tröghet) en komplett bild. Delbilderna ritas upp med 50 eller 60 bilder per sekund* för att inte ögat ska störas av blinkningarna vid bildbyte. En nackdel med interlace-tekniken är att bilderna kan se lite randiga ut om motivet rör sig snabbt.

För att indikera bildupplösning så används beteckningarna 576i för PAL och SECAM samt 480i för NTSC. Siffran anger hur många bildlinjer man har i vertikal riktning och ”i” indikerar att man utnyttjar interlace. Med bredd/höjd förhållande 4:3 i bilden så motsvarar alltså en PAL/SECAM bild ca. $576 \times \frac{4}{3} \times 576 \approx 0.44$ megapixlar och en NTSC bild ca. 0.31 megapixlar*. Vid bredd/höjd förhållande 16:9 blir motsvarande tal 0.59 respektive 0.41 megapixlar. I samtliga fall ser vi att TV-bilden innehåller mycket färre pixlar än bilderna från en digitalkamera (typiskt 10-60 megapixlar).

I samband med att digital-TV infördes så blev det allt vanligare att komplettera de gamla bildformaten med så kallad HD (High Definition) TV-sändningar. Sålunda gäller att man i dagsläget (2019) i Sverige hittar HD-sändningar i formaten 1080i (TV4, TV6 m. fl.) och 720p (SVT m.fl.). ”p” som står i det sistnämnda formatet betyder progressive, vilket innebär att man inte använder interlace-teknik utan att man vid varje bildbyte ritas upp alla bildlinjer. I samband med HD-video talar man ofta om 2K- respektive 4K-format. Dessa beteckningar står för att antalet pixlar på TV-bildens bredd är ca. 2000 respektive 4000. Formatet 1080i som nämns ovan skulle med bredd/höjd förhållandet 16:9 kunna betecknas som (nästan) 2K med sina 2.1 megapixlar. En 4K TV-bild innehåller ca. 8 megapixlar, och då börjar vi närma oss samma nivå som i en stillbild från en enklare digitalkamera. 4K används mest

* Bildfrekvensen var ursprungligen kopplad till nätspänningens frekvens, alltså används 50 Hz i Europa och 60 Hz i t.ex. USA.

♦ Vid analog TV-sändning är bilden kontinuerlig och inte uppdelad i diskreta pixlar, så megapixel-talen får inte tas allt för bokstavigt.

för dataspel, inspelningar på Blu-ray och vissa andra speciella applikationer. Redan nu finns TV-apparater förberedda för 8K, där varje bild innehåller 33 Mpixlar.

Eftersom antalet pixlar i en videobild oftast är mindre än antalet pixlar i sensorn, så måste en reduktion av pixelantalet kunna ske vid inspelningen. I detta sammanhang har CMOS-sensorer (kap. 14.3) en stor fördel jämfört med CCD-sensorer (kap. 14.2). CMOS-tekniken ger snabbare utläsning av data, samt gör det möjligt att läsa ut bara en delmängd av alla pixlar (i en CCD måste alla pixelvärden läsas ut). Eftersom i stort sett alla kameror numera använder CMOS-sensorer är de således väl lämpade för videoinspelning. Det finns lite olika strategier för att reducera antalet bildpixlar. Ibland utnyttjas t.ex. inte hela sensorns area, utan man väljer ut en något mindre central del av sensorn. Detta ger en "tele-effekt" i bilderna, dvs. det ser ut som om man hade en längre brännvidd på objektivet. Men vi ska här bara titta på fallet att man utnyttjar hela sensorarean även vid videoinspelning. Då finns tre huvudsakliga metoder för att reducera antalet pixlar. Dessa betecknas på engelska med "skipping", "binning" och "oversampling".

Skipping innebär som namnet antyder att man helt enkelt kastar bort pixelvärden. Låt oss titta på ett hypotetiskt fall där vi har en 6000 x 4000 (dvs. 24 Mpixel) sensor och önskar videobilder med 1500 x 1000 pixlar. Om vi utnyttjar skipping innebär det att vi bara läser ut var fjärde pixelrad i sensorn, och i de rader vi läser ut så läser vi bara ut var fjärde pixel. Detta är den absolut snabbaste metoden att reducera antalet pixlar eftersom inga beräkningar behöver utföras. Men den har nackdelar. En nackdel är att man kastar bort $\frac{15}{16} = 94\%$ av signalen från sensorn, vilket innebär ett sämre signal-brus förhållande (SNR). Dessutom kan man få stora problem med moiré-effekter på grund av den glesa samplingen vilket illustreras i Fig. 14.28 b (se kap. 22 ang. sampling och moiré). I takt med att kamerornas processor har blivit mera kraftfull har man därför övergått till binning och oversampling. Den sistnämnda kräver mest processor-kapacitet.

Binning innebär att man helt enkelt adderar signalerna från grannpixlar, vilket ger samma effekt som om man hade färre men större pixlar i sensorn. I exemplet ovan skulle det innebära att man adderar pixelvärden i en 4 x 4 omgivning. På så sätt får man en mycket högre signalnivå, och därmed högre SNR. Dessutom kommer moiré-effekterna att minska, Fig. 14.28 c. En nackdel är att skärpan kommer att bli lite sämre på grund av medelvärdesbildningen.

Oversampling brukar anses som den bästa metoden att reducera pixelantalet. Den innebär att man läser ut alla pixelvärden, men man gör mera avancerad databehandling än att bara addera grannpixlars värden. Det man försöker göra innan pixelreduktionen, är att så mycket som möjligt sudda ut linjemönster som är så täta att de kommer att ge moiré i den pixelreducerade bilden. Samtidigt vill man så lite som möjligt sudda ut linjemönster som kan återges utan moiré (se kap. 22 för en djupare förklaring av linjemönster och moiré). Några vanliga algoritmer som används vid oversampling är bikubisk och Lanczos. Vi går inte in på hur dessa fungerar, utan nöjer oss med ovanstående förklaring angående vad man vill åstadkomma. Som vi kan se i Fig. 14.28 b,c så sker en markant undertryckning vad gäller moiré-effekter när man går från skipping till binning. Genom att använda oversampling, Fig. 14.28 d, kan moiré nästan helt elimineras.

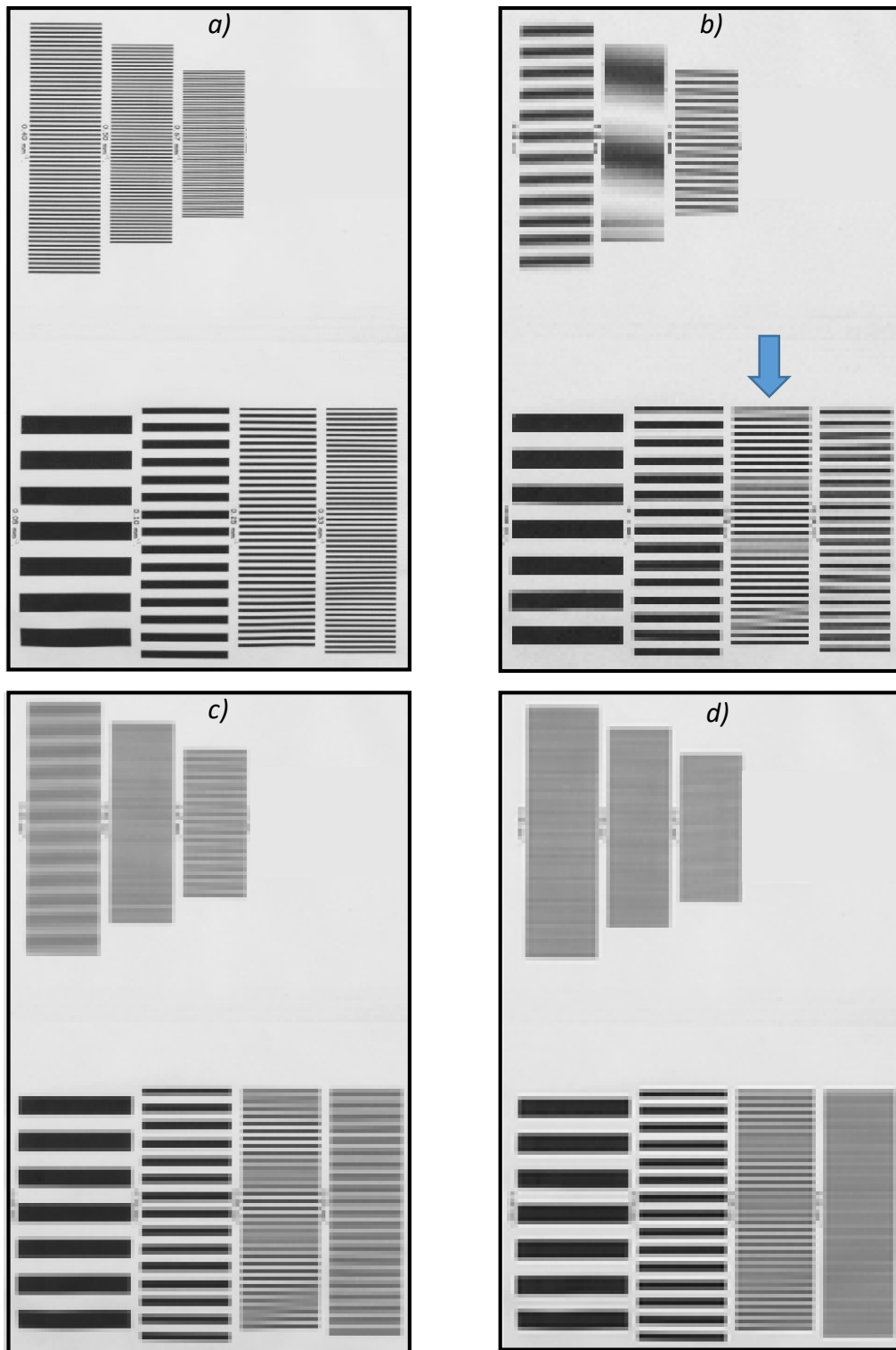


Fig. 14.28. Illustration av moiré-effekter som kan uppstå vid reduktion av pixelantalet i en bild. a) Originalbild före pixelreduktion. b) Pixelreduktion utförd med "skipping" (se texten). Pilen i bilden visar det tätaste linjemönster som kan återges med korrekt täthet efter reduktionen i antal pixlar. Observera att tätare linjemönster kan återges med god kontrast, men med fel täthet (falsk information). Idealt bör mönster som är för täta för korrekt återgivning inte återges alls. c) Pixelreduktion utförd med "binning". d) Pixelreduktion utförd med "oversampling" (Lanczos metod). Fallet d) ger den bästa bortfiltreringen av falsk information. (Sampling och moiré-effekter behandlas utförligt i kap. 22.)

Vid betraktning av Fig. 14.28 på en datorskärm är det viktigt att zooma in ordentligt så att moiré inte uppstår på grund av skärmens pixellering. Det tätaste mönster som kan återges korrekt efter pixelreduktion (blå pil i Fig. 14.28 b) uppvisar en långsam kontrastvariation (svävning) när vi förflyttar oss i vertikal riktning i Fig. 14.28 b-d. Själva linjemönstret återges dock med korrekt täthet. Svävningar förekommer ofta i digitala bilder på grund av att den analoga optiska bilden indelas i diskreta pixlar. Svävningar kan korrigeras i bildernas efterbehandling, men falsk information som uppstår för mönster tätare än blå pilen kan inte korrigeras efteråt. Att undvika falsk information på grund av moiré är särskilt viktigt när man fotograferar motiv som innehåller någon typ av linje- eller rutmönster med hög kontrast, exempelvis spjälstaket, galler, tegelväggar/tak, randiga/rutiga kläder mm.

Med hjälp av ovanstående metoder kan alltså pixelantalet i en kamerasensor reduceras så att det motsvarar t.ex. standardvideo (PAL/SECAM/NTSC), 2K eller 4K. Även andra videoformat kan erhållas från många kameror. Hur många pixlar lönar det sig att ha i en bild som visas på en skärm av viss storlek som betraktas på ett visst avstånd? Ögat har ju en begränsad upplösningsförmåga. Låt oss titta på några exempel.

Ögats upplösningsförmåga brukar anges till ca. en bågminut, dvs $1/60$ grad (upplösningsförmåga behandlas i kap. 16). Det innebär att vi kan förvänta oss att på 3.4 meters avstånd nätt och jämnt kunna urskilja ett svartvitt linjemönster där ett svartvitt linjepar upptar en bredd av en millimeter. Eftersom ca. tre meter torde vara ett ganska vanligt betraktningssavstånd vid TV-tittande, så använder vi 3.4 meters avstånd i våra beräkningar (resultaten kan ju lätt skalas om för andra avstånd). Låt oss anta att vi tittar på en TV-bild med 1080 linjer och aspect ratio 16:9 (ett mycket vanligt format vid TV-sändningar). Hur stor behöver en sådan skärm vara för att vi ska kunna se (nätt och jämnt) ett linjemönster där varannan rad är svart och varannan vit? Enligt ovan skulle det då behövas en linjebredd av 0.5 mm på skärmen, och alltså en total skärmhöjd av 540 mm. Skärmbredden blir då 960 mm. Nu brukar skärmstorlekar anges utgående från diagonalen mätt i tum, vilket för den aktuella skärmen blir 43 tum. Detta är en ganska vanlig skärmstorlek, så det verkar ju passa bra. Vill man titta på TV-bilder i 4K-format (2160 linjer) så måste man fördubbla skärmens bredd och höjd, vilket skulle ge en 86 tums skärm. Skärmar i den storleksordningen är inte så praktiska (nästan två meter breda och dessutom mycket dyra). Så för att njuta av 4K-upplösning måste man ofta sitta mycket nära skärmen, kanske på två meters avstånd eller ännu närmare beroende på aktuell skärmstorlek. Om man går nästa steg i utvecklingen och inför 8K, så skulle det innebära att man får sitta högst 1.7 meter från en 86 tums skärm för att kunna njuta fullt ut av bildernas detaljrikedom. Det är nog inte så praktiskt för TV-tittande till vardags. Upplösning i all ära, men det är nog oftare programinnehållet som avgör hur njutbart ett program är.

Låt oss nu titta lite på bildkomprimering i samband med video. Videoinspelning genererar ett mycket stort antal bilder på kort tid. Dessa bilder måste snabbt kunna lagras undan, och sedan snabbt kunna spelas upp. Det innebär att totala datamängden bör reduceras så mycket som möjligt utan att bildkvalitén blir påtagligt lidande. Det är just detta som är målet med videokomprimering. Redan på analog-TV tiden användes komprimering av signalen för att minska den behövliga bandbredden vid TV-sändningar. Detta skedde genom att använda högre

bandbredd för luminans-signalen än för färg-signalen. Denna filosofi bygger (lite förenklat) på att en suddig färgbild överlagrad på en skarp svartvit bild uppfattas som en skarp färgbild av ögat. Det räcker alltså med en lägre upplösning för färginformationen. Även i dagens digitala värld bygger bildkomprimering på samma filosofi, plus att den utökats med en mängd andra moment.

En mycket vanlig teknik för att komprimera videodata kallas MPEG-4 (MPEG står för Moving Picture Expert Group). MPEG-4 använder en teknik som liknar JPEG-tekniken som är vanlig för stillbilder (se kap. 23). Men vid rörliga bilder kan man dessutom utnyttja att innehållet ofta ändras mycket lite från en bildruta till nästa; genom att bara spara information om förändringen kan man få ner datamängden kraftigt. Detta gör att en MPEG-4 fil ofta innehåller en faktor 50 till 100 gånger mindre data än motsvarande okomprimerade videofil. Men kompressionsfaktorn beror på bildinnehållet, motivrörelser och hur hårt man vill komprimera. MPEG-4 är den kompressionsmetod som vanligen används vid sändning av digital-TV*. Precis som för jpeg-bilder använder man i videosammanhang oftast signaler som består av 8 bitar (se kap. 14.7), och med användning av gammakorrektion (kap. 14.6). Avancerade kameror tillåter ibland ett större bit-djup (t.ex. 10 bitar) i vissa inspelningsmoder, vilket möjliggör en större dynamisk vidd.

Som exempel på videoprestanda för en fullformat digitalkamera år 2019, kan nämnas att den typiskt kan spela in i format upp till 4K progressive, 30 bilder/s i MPEG-4 format. Alternativt kan man spela in bilder med färre pixlar men fler bilder per sekund (ofta upp till 120) för slow motion upptagningar. Även mobilkameror tillåter videoinspelning, ibland ända upp till 4K-format. I många kameror kan man välja mellan olika bildfrekvenser för en och samma bildupplösning (t.ex. mellan 24 och 120 bilder per sekund i 2K-upplösning). Genom att välja en lägre bildfrekvens kan man spara minnesutrymme. Att gå så lågt som 24 bilder/s kan tyckas tveksamt med tanke på vad som tidigare sagts om att man valde 50 eller 60 (del)bilder/s vid införandet av analog-TV för att undvika störande blinkningar[♥]. Men man får då komma ihåg att på den tiden använde TV-skärmarna katodstrålerör (CRT) som lyste upp bara en punkt i taget. Genom att svepa den lysande punkten över skärmen ritades en bild upp. I moderna skärmar (vanligen LCD) används oftast så kallad "active matrix technology" vilket innebär att alla skärmens pixlar lyser hela tiden, men uppgraderas med bildfrekvensen. Det innebär att skärmen inte blinkar på samma sätt som en CRT. Det som händer vid låga bildfrekvenser är i stället att snabba rörelser kan se ryckiga ut.

Både seriebildstagningar och HD-video kräver att stora mängder data kan överföras och behandlas med hög hastighet. Vid t.ex. 4K-video krävs att man 30 gånger per sekund kan överföra pixeldata från sensorn, reducera dessa till en 8 megapixel-bild inklusive färginterpolation, vitbalanskorrigering, gammakorrigering, MPEG4-komprimering mm. För att klara allt detta innehåller digitalkameror en särskild

* Eftersom ett flertal komprimerade TV-program sänds i samma kanal (multiplexing), innebär det att signalerna måste separeras och avkomprimeras i TV-mottagaren. Detta tar en viss tid (ofta i storleksordningen 10 sek), vilket gör att t.ex. klockor som visas i "direktsändning" kommer att visa flera sekunder fel (det gäller t.ex. klockan i Text-TV.)

♥ Den klassiska biografifilmen visades med 24 filmrutor/s, men man använde en roterande sektorbländare som gjorde att varje bildruta blinkade fram två gånger. Blinkningsfrekvensen blev därför 48 Hz.

bildprocessor skräddarsydd för syftet. För att öka snabbheten utnyttjar en sådan processor ofta flera parallella kanaler för databehandling. Men ibland sätter ändå processorkapaciteten en gräns för hur avancerade metoder som kan utnyttjas för reduktion av pixelantalet vid videoinspelning. Därför kan man t.ex. i vissa videomoder inte använda oversampling (se ovan) utan tillämpar den enklare binning-moden.

Efter bildprocessorn hamnar bilddata i ett buffertminne av typen DRAM (dynamic random access memory). Detta är samma typ av minne som RAM-minnet i en dator. DRAM-minnen är snabba och klarar av att ta emot dataströmmen från bildprocessorn både vid seriebilda- och videoinspelningar. Minneskapaciteten är dock ganska begränsad (ofta motsvarande 50-60 bilder i full upplösning), vilket begränsar längden på seriebildinspelningar. Minnesinnehållet måste sedan överföras till ett bildminneskort av typen "flash-minne" (samma typ som i USB-minnen) där det kan lagras permanent. Flash-minnen är betydligt långsammare än DRAM, men har fördelarna att kapaciteten är stor och att minnesinnehållet bevaras även utan spänningsmatning. Buffertminnet i kameran gör alltså att man, under en begränsad tid, kan samla på sig bilddata betydligt snabbare än vad bildminneskortet klarar av att ta hand om. Detta utnyttjar man vid seriebildstagning, där det efter varje tagning krävs en viss tid (kanske en minut) för att överföra data från buffertminnet till minneskortet. Men vid videoupptagningar räcker inte buffertminnet till, utan då måste minneskortet klara av att ta hand om data i samma takt som de genereras. Detta är möjligt på grund av det lägre pixelantalet och att man kan komprimera data kraftigt.

Minneskort till digitalkameror finns i en mängd olika storlekar vad gäller minneskapacitet, och med olika prestanda vad gäller skriv- och läshastighet. Ett typiskt kort för användning vid serietagningar och HD-video kan lagra 32 eller 64 gigabyte och klara hastigheter av 50 megabyte/s. Ett 64 gigabyte minneskort kan lagra över två timmar 2K-video och över en timme 4K-video. Detta torde räcka för de flesta praktiska syften. Det som i praktiken ofta sätter gränsen för hur långa kontinuerliga tagningar som kan göras är temperaturen på bildsensorn. När sensorn arbetar kontinuerligt drar den så mycket ström att temperaturen kontinuerligt ökar, och när den nått en viss övre gräns avslutar kameran helt enkelt inspelningen. Typiskt nås denna max-temperatur efter ca. 30-60 minuters kontinuerlig inspelning, men vid hög omgivningstemperatur förkortas tiden. När kameran varit avstängd så lång tid att temperaturen på sensorn sjunkit tillräckligt kan man fortsätta inspelningen.

15. Optiska avbildningsfel och diffraktion

15.1. Inledning

I kapitlet om optisk avbildning antog vi att en punkt i motivet avbildas som en perfekt punkt i den optiska bilden. Det räcker då att betrakta två från motivpunkten utgående strålar och se efter var dessa korsar varandra i bildrymden (Fig. 5.1). Alla andra från motivpunkten utgående strålar antogs även de passera samma punkt. Ack, om verkligheten vore så enkel! Då skulle man få en oändligt liten punktbild, och därmed oändlig ”skärpa” i bilden. I verkligheten är det två saker som begränsar bildskärpan och gör att en bild som förstoras alltmer kommer att se suddigare och suddigare ut. Den första begränsningen utgörs av avbildningsfelen, vilka helt enkelt är brytningsfel som uppstår när ljuset bryts i linsytorna. Den andra begränsningen utgörs av diffraktionen, vilken alltid uppträder när ljus färdas genom en öppning med ändliga dimensioner. Vi ska börja med att titta på några olika typer av avbildningsfel. Dels hur de yttrar sig, dels om det finns något ”knepp” att ta till för att minimera deras inverkan. Framställningen som görs är mycket förenklad på så sätt att bara de mest fundamentala felen tas med. Dessutom beskriver vi de olika felen i renodlad form, dvs. vi antar att bara ett fel i taget existerar. I verkligheten finns naturligtvis alla felen (mer eller mindre) närvarande samtidigt.

Avbildningsfelen brukar delas in i två huvudgrupper. a) Fel som uppträder redan i bildcentrum (men de finns med över hela bildytan). b) Fel som inte existerar i bildcentrum, men dyker upp när vi lämnar centrum och vanligtvis blir värre ju längre ut mot kanten vi kommer. Vi ska börja genomgången med den första kategorin, och tittar på avbildningen i en enkel lins. Två fel visar sig redan i bildcentrum, nämligen sfärisk aberration och kromatisk längdaberration (även kallad longitudinell kromatisk aberration).

15.2. Sfärisk aberration

Sfärisk aberration illustreras i Fig. 15.1. Vi ser att strålar som infaller på olika avstånd från optiska axeln kommer att brytas olika kraftigt. Resultatet blir att de inte skär varandra i en och samma punkt. Var sensorn än placeras får man en oskarp ljus fläck istället för en punktbild. Det bästa man kan göra är att hitta det läge längs optiska axeln där den oskarpa fläcken har minst utsträckning. Sfärisk aberration betyder ”sfärisk avvikelse”. Detta kan missuppfattas så att man tror att felet uppstår på grund av att man inte lyckats slipa linsytorna till perfekt sfärisk form. Så är inte fallet! Tvärtom, felet uppstår som en följd av att ytorna är just sfäriska. Man kan tillverka så kallade asfäriska linsytor och därmed kraftigt minska felet. Asfäriska linser var tidigare ovanliga och mycket dyra. Genom förbättrade tillverkningsmetoder (t.ex. precisionsgjutning) så har kostnaderna reducerats, och det är mycket vanligt att objektiven till digitalkameror innehåller en eller flera asfäriska linser. Asfäriska linser minskar inte bara sfärisk aberration, utan även andra avbildningsfel. Asfäriska linsytor förekommer också ofta i optik för projektorer, kikare och mikroskop.

Förutom med asfäriska linser, så minskar man avbildningsfelet genom att kombinera flera linselement. I sin enklaste form görs detta genom att kombinera en positiv och en negativ lins tillverkade av olika glassorter. Minskningen i sfärisk aberration som uppnås med en sådan kombination beror på att sfäriska aberrationen går åt motsatt håll för positiva och negativa linser. Även i mycket dyra objektiv, innehållande asfäriska linser och många linselement, finns emellertid en viss sfärisk aberration kvar. Därför är det intressant att veta om man kan göra något för att minska effekterna av felet. Svaret är mycket enkelt – blända ned, dvs minska bländaröppningen. Om man bara använder de centrala delarna av linsen, kommer man att helt skära bort strålar från linsens yttre zoner där ljuset bryts annorlunda. Resultatet blir en mer sammanhållen fokuspunkt, vilket ger bättre bildskärpa, se Fig. 15.1. Man får dock inte driva denna nedbländning för långt av två skäl. För det första blir bilden mörkare, vilket kräver längre exponeringstid. Detta kan ge problem med rörelseoskärpa. För det andra ökar diffraktionen ju mindre öppningen är. Detta ger i sin tur en oskarp bild, vilket ska beröras senare i detta kapitel. Som ledtråd kan man därför säga att bländartal runt 5.6 till 8 oftast resulterar i bäst skärpa. Digitalkameror med mycket små sensorer bör användas vid lägre bländartal; runt 4 kan ofta vara optimalt för dessa kameror, beroende på att diffraktionseffekter uppträder redan vid lägre bländartal om sensorn är liten.

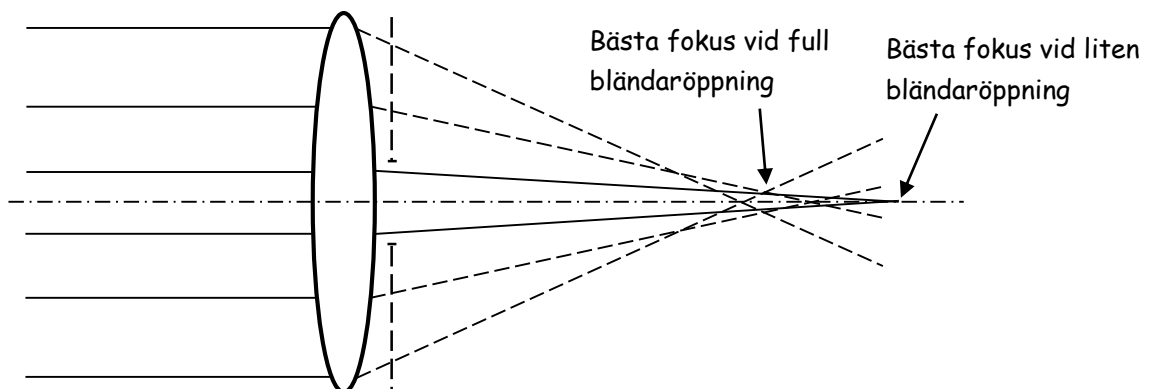


Fig. 15.1. Sfärisk aberration.

15.3. Kromatisk längdaberration

Låt oss nu titta på den andra aberrationen som dyker upp redan i bildcentrum, nämligen kromatisk längdaberration. Såsom namnet antyder handlar det om ett färgfel. Olika våglängder bryts helt enkelt olika kraftigt. Vi får därför olika brännvidder för olika våglängder. Detta gäller även för en tunn lins som används vid liten bländaröppning. I detta fall kan brännvidden beskrivas med den så kallade "lensmaker's formula"

$$\frac{1}{f} = (n-1) \left(\frac{1}{r_1} + \frac{1}{r_2} \right) \quad (15.1)$$

där n = glasets brytningsindex och r_1 och r_2 är linsytornas krökningsradier (vi har antagit bikonvex lins). Eftersom n i allmänhet är högre för kortare våglängder, så kommer f att minska med våglängden såsom visas i Fig. 15.2.

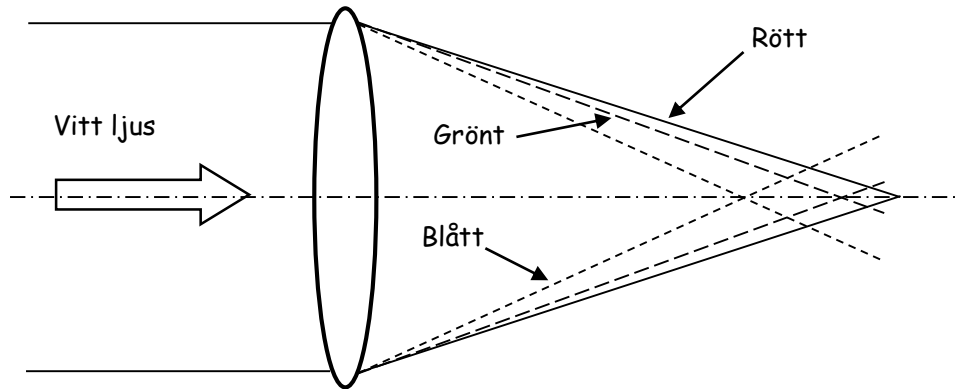


Fig. 15.2. Kromatisk längdaberration.

Precis som fallet var med sfärisk aberration, så kan den kromatiska aberrationen minskas genom att kombinera en positiv lins och en negativ lins där man använder olika glassorter i de bägge linserna. Att man därvid minskar felet, beror på att kromatiska aberrationen går åt olika håll för de bägge linserna. Det visar sig turligt nog att man med en och samma kombination av positiv och negativ lins kan reducera både sfärisk och kromatisk aberration. I denna enkla typ av "akromatobjektiv" väljs ofta krökningsradierna så att linselementen kan limmas ihop till en enhet, Fig. 15.3.

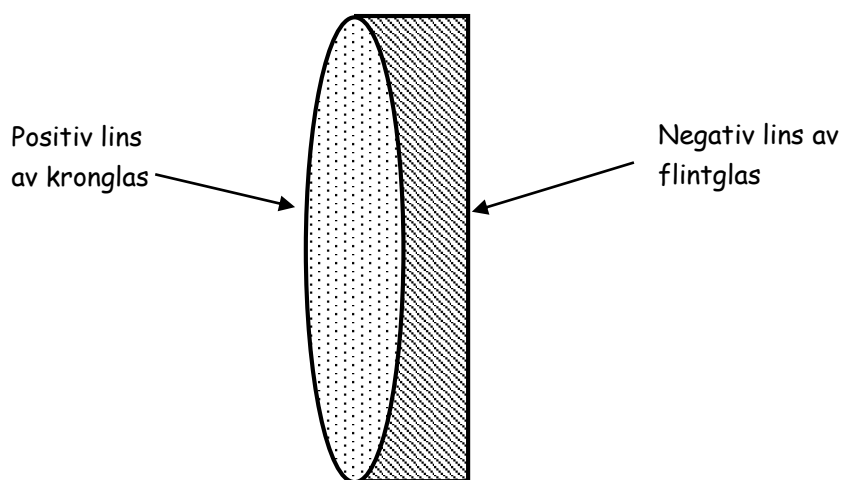


Fig. 15.3. Akromatobjektiv.

Akromatobjektiv brukar användas i kikare och teleskop, beroende på att de fungerar bra för små bildvinklar (någon enstaka grad). I fotografiska sammanhang är dock bildvinkeln oftast så stor att ett akromatobjektiv ger alltför dålig skärpa i bildkanten. Man tvingas därför använda fler linselement. Även mycket dyra objektiv uppvisar dock en viss kromatisk aberration, vilket resulterar i en suddighet i bilden (såvida inte monokromatiskt ljus används). Precis som för sfärisk aberration, så gäller att en mindre bländaröppning kommer att reducera inverkan av kromatisk aberration. Visserligen förändras inte brännvidden för olika våglängder av nedbländningen (jfr. lensmaker's formula), men den oskärpefläck vi får på sensorn kommer att minska när strålknipets diameter minskar. Det är precis samma effekt som gör att skärpedjupet ökar vid nedbländning. Som tidigare gäller dock att nedbländningen inte får bli så stor att diffraktion förstör skärpan.

När man förflyttar sig från bildcentrum ut mot periferin så tillkommer ytterligare ett antal aberrationer. Dessa kallas **lateral kromatisk aberration**, **bildfältskrökning**, **astigmatism**, **koma** och **distorsion**. Nedan följer korta beskrivningar av hur dessa fel yttrar sig.

15.4. Lateral kromatisk aberration

Lateral kromatisk aberration innebär att även om ljus av olika våglängder fokuseras i samma punkt, så behöver detta inte innebära att brännvidden är densamma för olika våglängder. Det visar sig nämligen att huvudplanens lägen kan vara olika för olika våglängder, se Fig. 15.4. Där ser vi ett exempel på ett objektiv (yttre linsytorna och huvudplanen utritade) där den longitudinella kromatiska aberrationen är korrigerad, medan den laterala kvarstår.

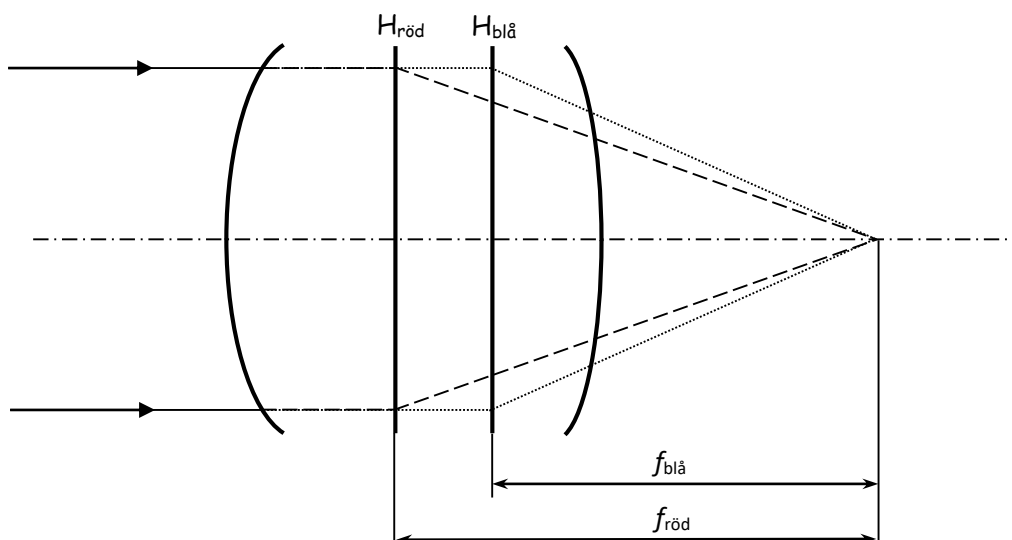


Fig. 15.4. Lateral kromatisk aberration. Ljus av olika våglängder fokuseras i samma punkt, men brännvidderna är olika. Detta innebär att avbildningsskalan kommer att variera med våglängden.

Lateral kromatisk aberration innebär alltså att brännvidden är olika för olika våglängder. Detta innebär att avbildningsskalan (ekv. 5.3) kommer att vara våglängdsberoende. I praktiken märks detta fel mest på att en skarp svart/vit övergång långt ute mot bildkanten kommer att bli färgad. Till skillnad från sfärisk och longitudinell kromatisk aberration, så minskar inte den laterala kromatiska aberrationen vid nedbländning. Det är alltså önskvärt att man genom objektivkonstruktionen lyckas göra felet tillräckligt litet. Detta görs genom en symmetrisk uppbyggnad, och lämpliga val av glassorter i de olika linselementen. När man arbetar med digitala bilder är det möjligt att i efterhand kompensera för den laterala kromatiska aberrationen hos optiken. Detta görs genom små justeringar av de röda, gröna och blå delbildernas storlek. När det i digitalkamerasammanhang talas om "kromatisk aberration" så är det oftast lateral kromatisk aberration som avses.

15.5. Bildfältskrökning

Bildfältskrökning innebär att bästa bildskärpa inte erhålls i ett plan vinkelrätt mot optiska axeln, utan längs en skålförmig yta såsom visas i Fig. 15.5.

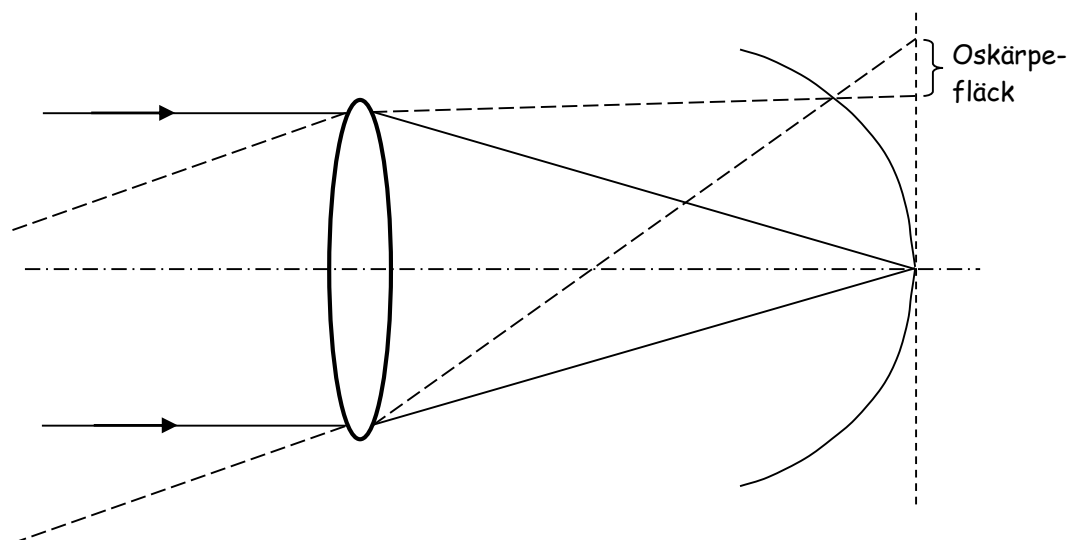


Fig. 15.5. Bildfältskrökning.

I praktiken yttrar sig felet i att man inte kan fokusera objektivet så att man får bästa möjliga skärpa i bildmitten och kanten samtidigt. Oskärpan kan dock minskas genom nedbländning. Genom lämplig kombination av glassorter och linselementens krökningsradier, kan man konstruera objektivet med liten bildfältskrökning.

15.6. Astigmatism

Astigmatism ger en oskärpefläck som påminner om en ellips, vars form, storlek och orientering ändras med fokuseringen. Detta illustreras i Fig. 15.6.

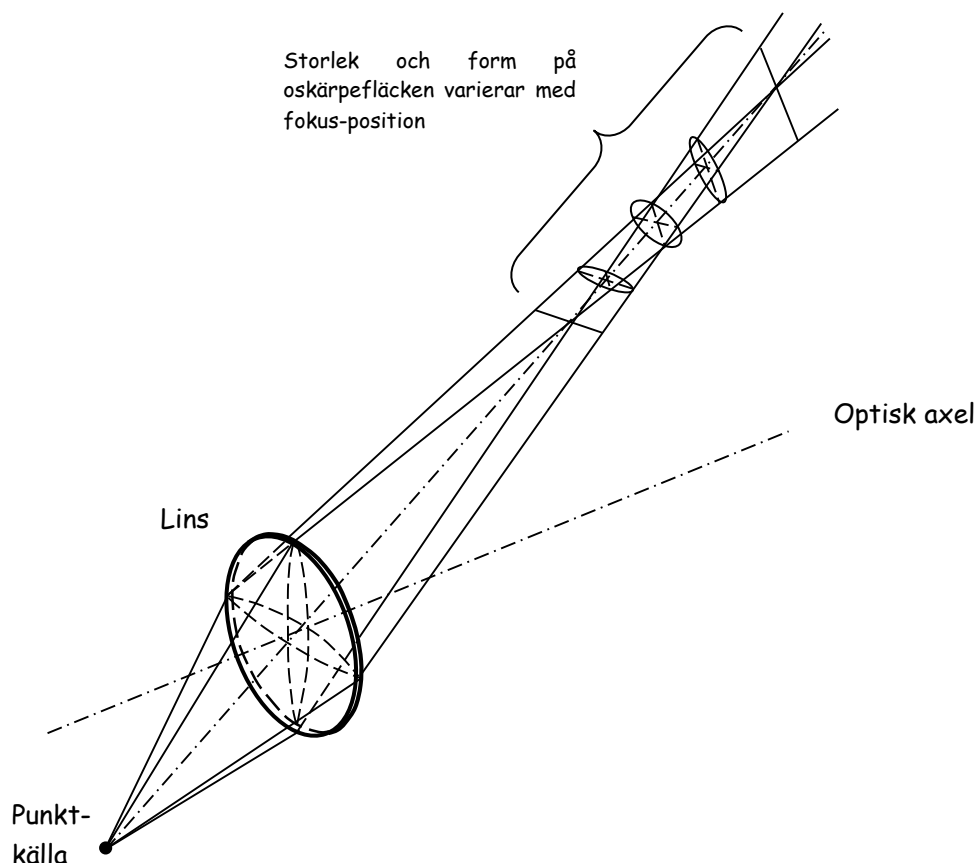


Fig. 15.6. Astigmatism.

Oskärpefläckens storlek ökar ju längre ut mot bildkanten vi kommer. Detta avbildningsfel är en starkt bidragande orsak till att det enkla akromatobjektiv som tidigare beskrevs bara är användbart för mycket små bildvinklar. Visserligen minskar astigmatismen vid nedbländning, men det behövs en mycket kraftig nedbländning för att få bukt med den i t.ex. en akromat. Detta resulterar i en mycket ljussvag bild, som dessutom blir suddig på grund av diffraktion. Därför är det av yttersta vikt att alla fotografiska objektiv (med undantag för extremt långa teleobjektiv) har god korrektion för astigmatism. Precis som för andra avbildningsfel, reduceras astigmatism genom att bygga objektiv med ett flertal linselement (både positiva och negativa) av olika glassorter. De enklaste objektiven där man påtagligt lyckats reducera astigmatismen består av tre linselement, så kallade tripletter, se exemplet i Fig. 15.7. Sådana objektiv användes tidigare i enklare kameror, förstöringsapparater och projektorer.

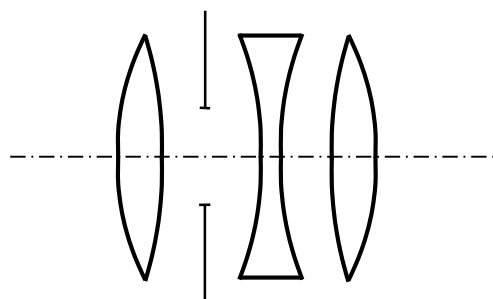


Fig. 15.7. Tripletobjektiv.

15.7. Koma

Koma har fått sitt namn efter oskärpefläckens utseende. Den liknar nämligen en liten komet med ett ljust "huvud" och en utbredd "svans" med avtagande ljusintensitet. Man kan säga att felet uppstår genom att, vid snett infall, olika zoner av linsen ger både olika brännvidd (precis som vid sfärisk aberration) och olika avbildningsskala. Detta illustreras i Fig. 15.8.

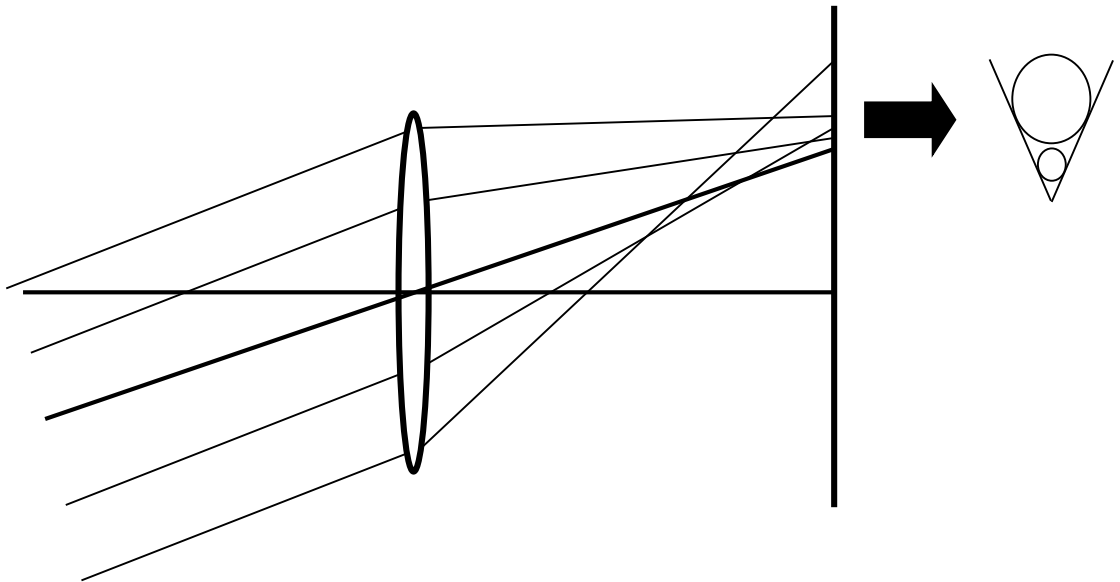


Fig. 15.8. Koma.

Eftersom koma yttrar sig som en oskärpefläck, så minskar den vid nedbländning. Vid objektivkonstruktion kan felet kraftigt reduceras genom en symmetrisk uppbyggnad med bländaren placerad i centrum.

15.8. Distorsion

Distorsion är ett avbildningsfel som skiljer sig från de flesta övriga felen genom att det inte påverkar bildens skärpa. Det ger istället en förvrängd geometri. Felet uppstår genom att avbildningsskalan varierar över bildplanet. Om skalan ökar när vi rör oss ut mot kanten, talar man om kuddformig distorsion. Om däremot skalan minskar mot kanten, har man tunnformig distorsion. Som framgår av Fig. 15.9, sammanhänger distorsionen intimt med bländarens placering. I ett helt symmetriskt system, med bländaren i mitten, kan distorsionen praktiskt taget elimineras (åtminstone för en viss avbildningsskala).

Eftersom distorsion inte ger oskärpa i bilden, så kommer inte felet att minska vid nedbländning. Felet kan emellertid korrigeras när bilderna datorbearbetas, precis som fallet var för lateral kromatisk aberration. Sådana korrekationer kan ske direkt i kamerans processor när bilderna tas ut i jpeg- eller tiff-format. När bilderna tas ut i

RAW-format, så kan motsvarande korrekationer göras i RAW-konvertern på en yttre dator. Bildfilerna brukar innehålla information om hur stora korrekationer som behöver göras.

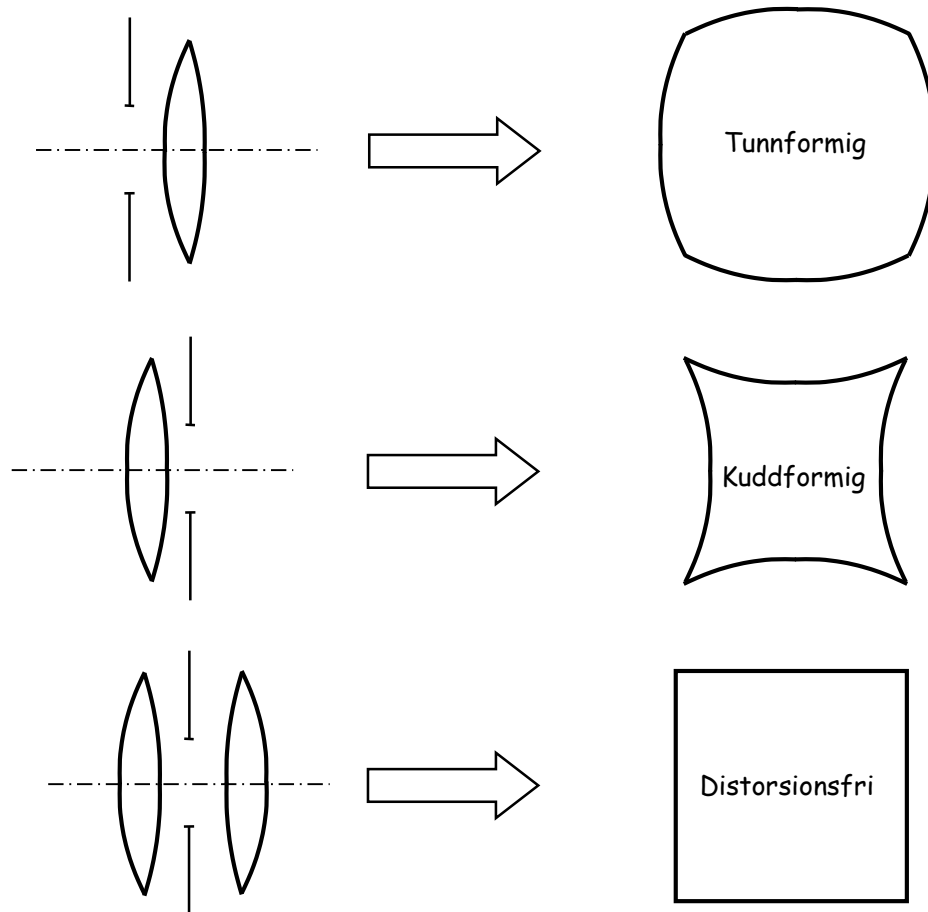


Fig. 15.9. Distorsion vid avbildning av ett kvadratisk motiv. Distorsionen påverkas starkt av bländarens placering.

15.9. Aberrationer, sammanfattning

Sammanfattningsvis kan sägas att de här beskrivna avbildningsfelen reduceras genom att man bygger objektiv bestående av flera olika linser, ofta även med asfäriska ytor. Man kombinerar både positiva och negativa linser, och använder ett flertal olika glassorter. De olika glassorterna skiljer sig åt vad gäller brytningsindex och dispersion (dvs hur brytningsindex varierar med våglängden). Ofta eftersträvar man ett symmetriskt objektiv med bländaren placerad i mitten. Detta byggsätt minskar distorsion, koma och lateral kromatisk aberration. Exempel på hur ett välkorrigerat objektiv kan vara uppbyggt visas i Fig. 15.10.

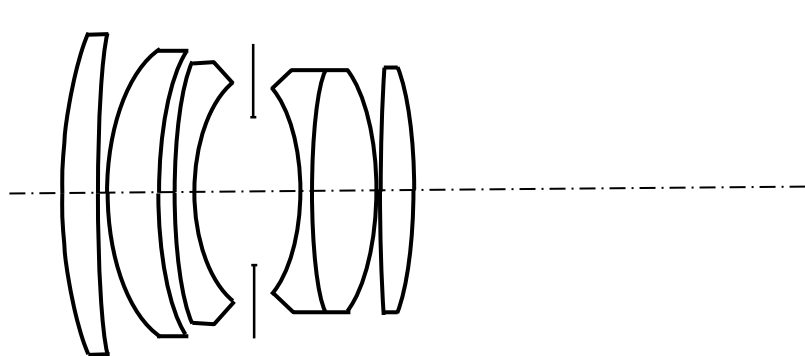


Fig. 15.10. Exempel på uppbyggnaden av ett välkorrigerat objektiv (brännvidd 55 mm, ljusstyrka 1.8).

Avbildningsfelen kan i princip reduceras hur långt som helst. Problemet är bara att kostnaderna ökar, samtidigt som nyttan av ytterligare korrekationer blir ganska marginell när man uppnått en viss gräns. Men låt oss anta att vi har oändliga resurser, och att vi därför kan göra avbildningsfelen oändligt små. Kommer vi då att få oändligt skarpa bilder? Nej, det visar sig att diffraktionseffekter kommer att begränsa skärpan även om vi helt lyckas bli av med alla avbildningsfel.

15.10. Diffraktion

Vi ska nu titta lite på prestanda hos ett diffraktionsbegränsat objektiv (dvs. ett objektiv med försumbara avbildningsfel). Låt oss anta att vi har en experimentuppställning enligt Fig. 15.11.

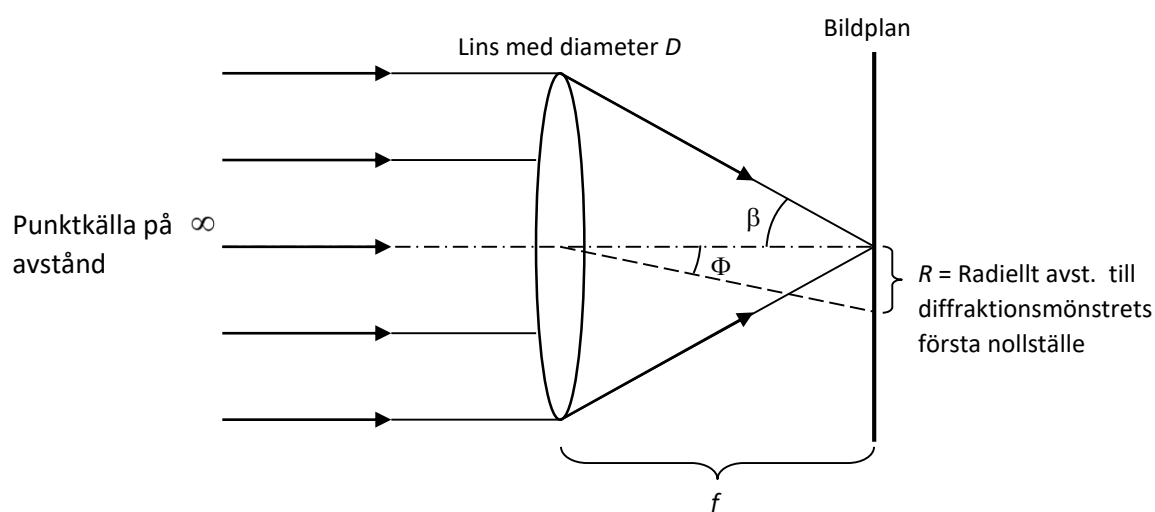


Fig. 15.11. En lins helt utan avbildningsfel skulle enligt geometrisk optik ge en perfekt punktbild. I verkligheten kommer diffraktion att förhindra detta. I stället erhåller man i bildplanet ett rotationssymmetrisk mönster enligt Fig. 15.12.

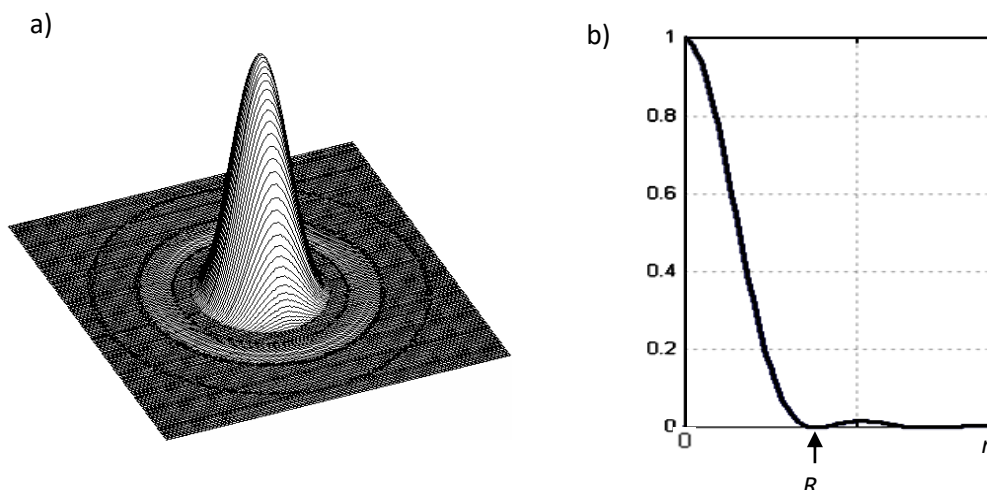


Fig. 15.12. a) Belysningsfördelning i bildplanet i Fig. 15.11. b) Mönstrets profil i radiell riktning. Utsträckningen, R , brukar anges till första nollstället.

Diffractionen ger en belysningsfördelning med ett maximum på optiska axeln, och ett antal sekundära maxima när man rör sig ut från axeln, Fig. 15.12. Mönstret är rotationssymmetriskt runt optiska axeln. Eftersom 84% av ljusflödet finns inom den centrala toppen med radie R (kallad Airy-fläcken), brukar ofta oskärpefläckens storlek definieras som utbredningen av detta centrala maximum. Från läroböcker i fysik får vi att

$$D \sin \Phi = 1.22 \lambda. \quad (15.2)$$

Vinkeln Φ , som definieras i Fig. 15.11, är vanligtvis mycket liten, vilket betyder att vi kan skriva $\sin \Phi \approx \frac{R}{f}$. Vi får då att $R \approx \frac{1.22 \lambda f}{D} = 1.22 \lambda F$, där F står för bländartalet. En liten oskärpefläck, och därmed skarp bild, får vi om vi använder en kort våglängd och lågt bländartal*. Vid bländartal mindre än eller ungefär lika med 8 så kan emellertid verkliga objektivet ofta inte anses diffraktionsbegränsade, utan avbildningsfelen kommer att bli klart märkbara. I kap. 16 ska vi titta på hur man kvantitativt kan beskriva bildskärpan med begreppet upplösningssförmåga.

* Enligt ekvationen verkar det inte finnas någon teoretisk gräns för hur liten oskärpefläcken kan bli. Genom att öka D och/eller minska f skulle man i princip kunna göra R godtyckligt liten. I verkligheten är detta inte möjligt. Ekvationen är inte giltig för låga värden på F . Den exakta ekvationen lyder $R = \frac{0.61 \lambda}{\sin \beta}$, där β är angiven i Fig. 15.11.

Bländartalet $F = \frac{f}{D}$ ska alltså bytas mot $\frac{1}{2 \sin \beta}$ vid låga bländartal, dvs $F \lesssim 1$ (gräns för ca. 10% fel). R kan inte bli mindre än drygt halva våglängden, eftersom $\sin \beta \leq 1$.

16. Bildkvalitet: Upplösningförmåga

16.1 Inledning

I detta och nästa kapitel ska vi titta på hur man kan definiera kvantitativa mått på bildskärpan. Det finns ett flertal faktorer som påverkar denna bildskärpa, t.ex. optiken, sensorn och vibrationer. Vi ska börja med att titta på optikens inverkan. Skärpan i den optiska bilden begränsas huvudsakligen av aberrationer och diffraktion (vi bortser från inverkan av ströljus och andra faktorer). Vi ska först titta på ett objektiv helt utan aberrationer, dvs. ett diffraktionsbegränsat objektiv. I föregående kapitel såg vi att en punktkälla då avbildas med en oskärpefläck vars radie ges av $R \approx 1.22\lambda F$, där F är bländartalet (formeln gäller ner till $F \approx 1$ om ca. 10% fel tolereras). I kapitlet om skärpedjup angavs den maximala oskärpefläckens diameter till ca. 30 μm för normal fotografering med fullformatsensor (24 mm x 36 mm). Om vi sätter in detta värde i ovanstående formel får vi att det maximalt användbara bländartalet är ca. 22. Av denna anledning brukar högsta bländartalet sällan överstiga 22 på fullformatskameror. Om man istället har en kompaktkamera kan sensorstorleken vara ca. 4 mm x 6 mm. I ett sådant fall kan man bara tillåta en sjättedel så stor oskärpefläck, vilket medför ett högsta bländartal av ca. 4. Att en mindre oskärpefläck krävs när sensorn är liten, beror på att bilden då behöver förstöras mer innan den hamnar på datorskärmen eller blir utskriven.

16.2. Upplösningförmåga enligt Rayleighkriteriet

I fysikaliska sammanhang brukar man ofta ange optiska systems prestanda enligt det så kallade Rayleighkriteriet. Man tittar då på hur optiken avbildar två lika starkt lysande punktkällor som ligger nära varandra. Om vi antar att vi inte har någon koherens (dvs bestämd fasrelation) mellan ljuset från de bägge källorna, så blir belysningen i bildplanet helt enkelt summan av två diffraktionsmönster enligt Fig. 15.12 som är något förskjutna i sidled relativt varandra, Fig. 16.1.

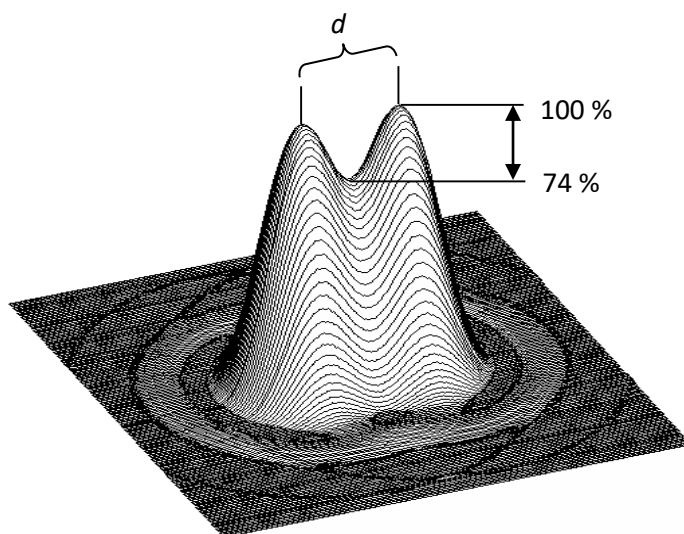


Fig. 16.1. Upplösninggräns enligt Rayleighkriteriet.

För att de två punktkällorna ska betraktas som ”upplösta” av optiken så kräver man att belysningen mellan de två topparna i belysningsprofilen ska ha sjunkit 26% eller mer från toppvärdet. Anledningen till valet av siffran 26% är att det är ungefär vad som krävs för att man med ögat ska kunna se att belysningen verkligen är lägre mellan topparna. Dessutom innebär detta val att separationen d blir precis lika stor som R i Fig. 15.11 och 15.12. I tabellen nedan ser vi några exempel på upplösningens förmåga, mätt i bildplanet, för olika bländartal (våglängden antas vara 550 nm). Observera att **låga** d -värden betyder **hög** upplösningens förmåga. Lite förenklat kan man säga att talet beskriver hur små detaljer som kan återges i bildplanet.

<u>Bländartal, F</u>	<u>Upplösning (diffraktionsbegr. optik), d</u>
45	30 μm
22	15 μm
11	7.4 μm
5.6	3.8 μm
2.8	1.9 μm
0.5*	0.3 μm (= 0.61 λ)

I verkligheten finns naturligtvis aberrationer i alla optiska system. Detta innebär att man för låga bländartal kommer att få upplösningens värden som är markant sämre än de värden som anges i tabellen. Detta beror på att aberrationerna ger en oskärpa som breddar punktbilden så att den blir större än diffraktionsmönstret. Man måste då öka på avståndet mellan punktbilderna för att förhindra att de flyter ihop, vilket ger ett större värde på d än i tabellen ovan. För ett typiskt objektiv kan d variera med bländartalet ungefär som i Fig. 16.2.

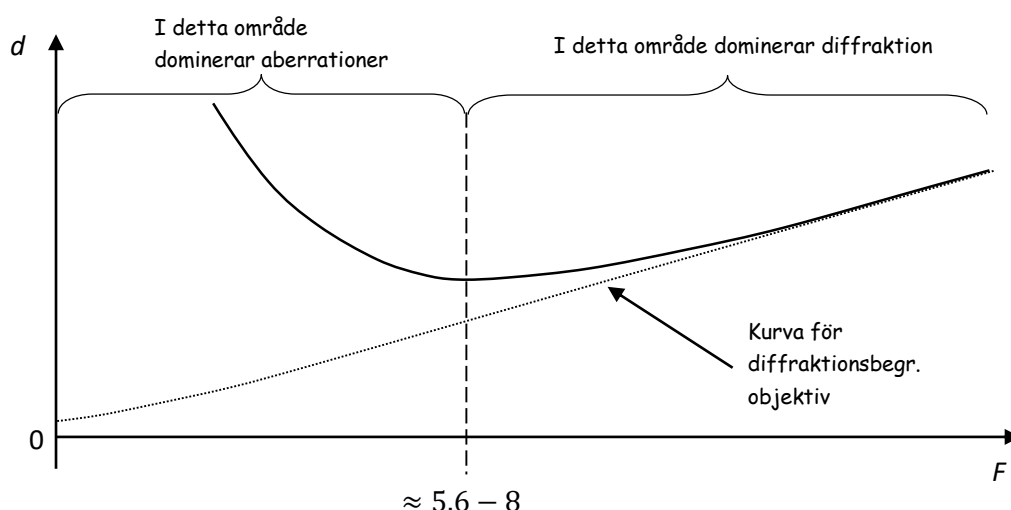


Fig. 16.2. Upplösningens förmågas beroende av bländartalet för ett typiskt objektiv.

* Bländartalet $F = \frac{f}{D}$ ska bytas mot $\frac{1}{2 \sin \beta}$ vid låga bländartal, dvs $F \lesssim 1$ (gräns för ca. 10% fel) se Fig. 15.11 och kapitel 15.10. Värdet i tabellens sista rad gäller för $\sin \beta = 1$ (motsvarande $F = 0.5$), vilket ger den teoretiskt högsta upplösningen (som i praktiken aldrig kan uppnås).

För bländartal större än 8 begränsas upplösningen huvudsakligen av diffraktion. Detta beror på att aberrationernas inverkan blir ganska försumbar vid så små bländaröppningar. Vid stora bländaröppningar däremot blir aberrationerna stora, och därför kommer upplösningen att bli sämre. Optimal upplösning erhålls ofta runt bländartal 5.6 – 8. Man har då bländat ner så pass mycket att aberrationernas inverkan kraftigt reducerats, men inte så mycket att diffraktionen ruinerar bildskärpan. Mycket högklassiga objektiv har bättre korrektion av aberrationerna, och kan därför förväntas ge optimal skärpa vid lägre bländartal, t.ex. 4 eller ännu lägre.

16.3. Bestämning av objektivets upplösningsförmåga

Nu vill naturligtvis den vetgirige fotoamatören genast störta iväg för att testa sina egna objektiv. Hur ska man då bära sig åt rent praktiskt? Det som behövs är två ∞ små, och lika starka, punktljuskällor, vilkas inbördes avstånd kan varieras noggrant. Dessa ska avbildas på stort avstånd av objektivet, och man ska med en kraftig lupp eller ett mikroskop betrakta belysningsfördelningen. Denna mätprocedur är inte så enkel i praktiken. Det man oftast gör är att istället för två punktljuskällor använda mönster av vita och svarta streck (typ övergångsställe, fast mindre). Sådana testmönster finns att köpa, men de kan också enkelt framställas med hjälp av dator och skrivare. Ett exempel på hur sådana testmönster kan se ut visas i Fig. 16.3.

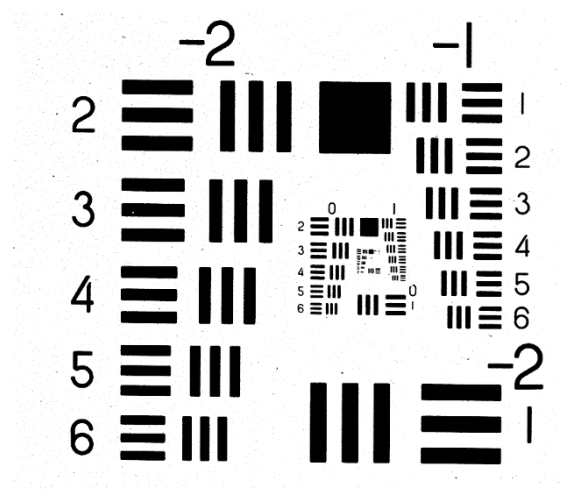


Fig. 16.3. En vanlig typ av upplösningssmönster för att testa optikens förmåga att återge små detaljer.

På grund av aberrationer och diffraktion i optiken kommer bilden av dessa streckmönster att suddas till. Om man illustrerar en belysningsprofil vinkelrätt mot strecken kan resultatet bli enligt Fig. 16.4.

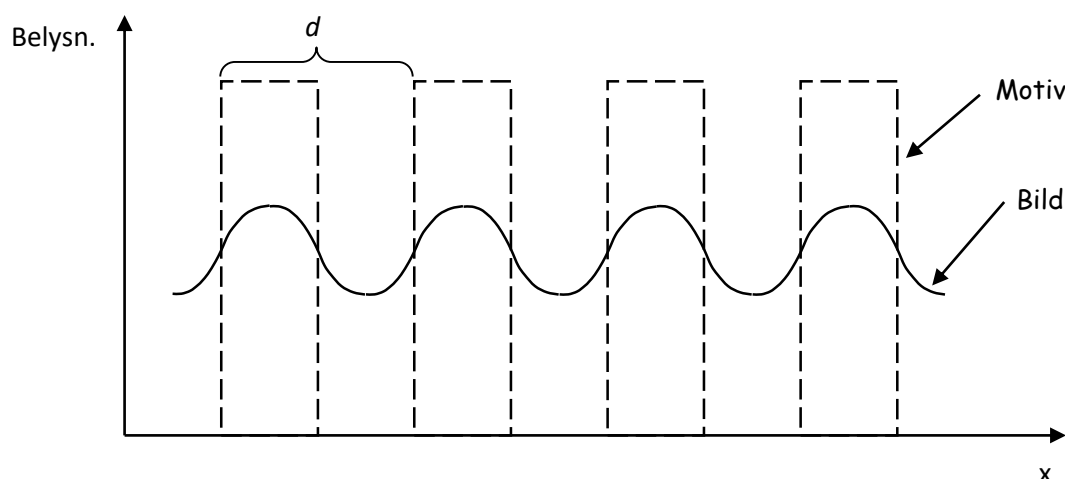


Fig. 16.4. På grund av avbildningsfel och diffraktion i optiken, kommer streckmönster att avbildas med lägre kontrast, och mjukare övergångar mellan ljusa och mörka partier, än vad som är fallet i motivet.

Precis som för de bägge punktljuskällorna, så krävs här att belysningen mellan topparna sjunker tillräckligt mycket för att vi ska uppfatta strecken och inte bara se en jämngrå yta. Det visar sig att det minsta periodavstånd som optiken klarar av att upplösa, d , ungefär motsvarar upplösningssgränsen enligt Rayleighkriteriet för två punktljuskällor. Vid praktiska tester brukar man avbilda testmönster både i bildcentrum och ute mot bildkanten. Vanligtvis blir upplösningen sämre ju längre ut mot bildkanten man kommer. Anledningen till att man har två vinkelräta linjeorienteringar är att upplösningssförmågan ute i bildkanten ofta är beroende på orienteringen. Detta beror på att astigmatism och koma ger oskärpefläckar som är asymmetriska.

Rent praktiskt görs upplösningsbestämningen visuellt, genom att man i hög förstoring betraktar den reella optiska bilden i bildplanet, Fig. 16.5. För kameror med film öppnas kamerans bakstycke och slutare, så att man kommer åt att betrakta den optiska bilden. För digitala systemkameror tas objektivet bort från kamerahuset och monteras separat i någon lämplig hållare. För digitalkameror med fast optik kan man inte göra någon test av enbart optiken.

Vid upplösningstester bör motivavståndet helst vara någon meter för att objektivet ska fungera optimalt (såvida det inte är ett makroobjektiv). Med ögat bestämmer man sedan vilket som är det tätaste linjemönster som (nätt och jämnt) syns. Det är härvid viktigt att luppens förstoring är tillräckligt hög, så att inte ögat begränsar upplösningen.

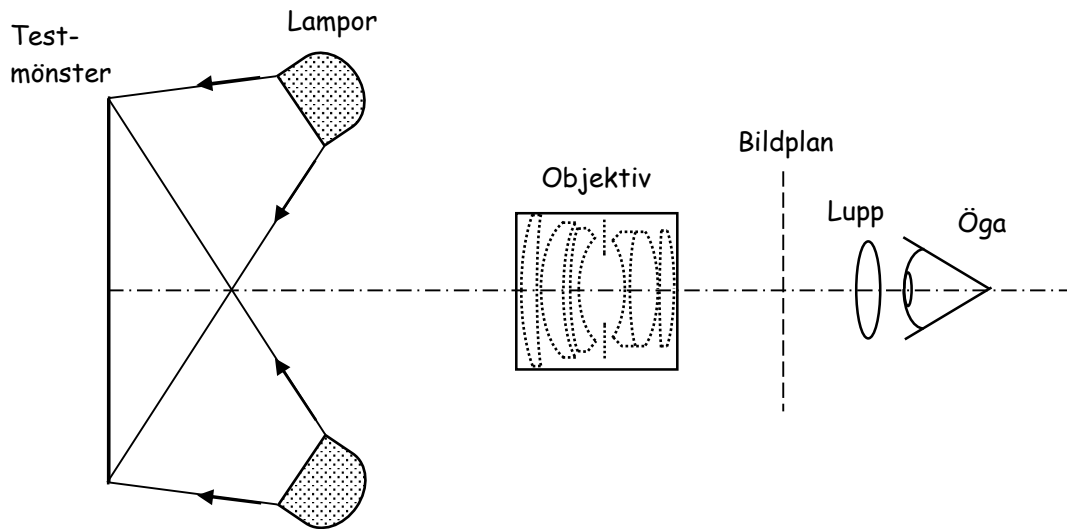


Fig. 16.5. Metod för att visuellt bestämma optikens upplösningsförmåga.

Istället för att uppge periodlängden, d , för det tätaste synliga mönstret som upplösningssgräns, så brukar man införa S = antalet perioder per längdenhet, dvs.

$$S = \frac{1}{d}.$$

Observera att det som avses är linjetätheten i bildplanet. Låt oss

exemplifiera! Antag att det tätaste motivmönster som kan avbildas skarpt består av 1 mm breda svarta streck med 1 mm breda vita mellanrum. Antag att motivavståndet är 5 m och objektivets brännvidd är 50 mm. Då gäller att antalet perioder är 0.5 mm^{-1} i motivet, och 50 mm^{-1} i bildplanet (avbildningsskalan är $1/100$). Upplösningssgränsen i detta fall ska alltså anges till $S = 50 \text{ mm}^{-1}$ (ofta anges sorten som "linjepar/mm"). Anledningen till att man anger upplösningssförmågan mätt i bildplanet, och inte i motivet, är att den då inte blir beroende av fotograferingsavstånd och brännvidd. Dessutom är det ju i allmänhet skärpan i bilden som är det intressanta.

16.4. Bestämning av sensorns upplösningssförmåga

Hittills har vi beskrivit hur upplösningssförmågan hos optiken kan testas. I fotografiska sammanhang kommer naturligtvis även sensorn att spela in (liksom tidigare filmen gjorde det). Även om optiken kunde ge en oändligt skarp bild, dvs. oskärpefläckens radie är lika med noll, så skulle man i bilden likafullt se en oskärpa. Detta beror dels på pixlarnas storlek, dels på att ljuset sprids i sensormaterialet. Dessutom kan man få en spridning i den laddning som fotonerna ger upphov till, Fig. 16.6.

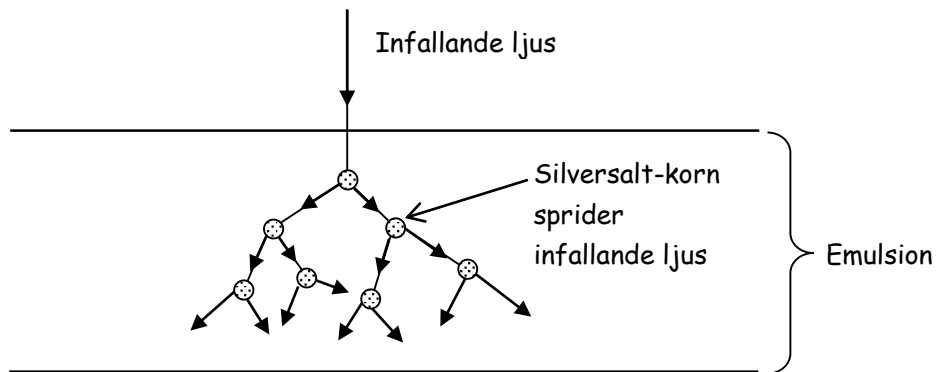


Fig. 16.6. Ljuset spridning mot kornen i en filmemulsion minskar skärpan i bilden. I en elektronisk sensor kan, förutom ljusspridning, även laddningsdiffusion sudda till bilden.

Fotografisk film kan man testa utan att få någon påverkan från objektivet. Detta kan göras genom att man kontaktkopierar mycket täta streckmönster direkt på filmen. Efter framkallning betraktas filmen i hög förstoring, och man bestämmer vilket som är det tätaste mönster som kan urskiljas. För en digitalkamera är det däremot svårt att separat testa sensors upplösningssmåga (i alla fall för en amatör). På många kameror kan man inte ta bort objektivet, och även på de kameror där det är möjligt så är det svårt att utföra "kontaktkopiering" av den typ som görs för film. Oftast är man därför tvungen att testa optik och sensor tillsammans genom att titta på den digitala bilden i kraftig förstoring. Men det är då svårt att veta vilken av komponenterna som är den svagaste länken. Om man kan ta bort objektivet och testa detta separat, kan man emellertid räkna ut sensors upplösningssmåga med hjälp av ekv. 16.1 nedan.

16.5. Systemets upplösningssmåga

Om man vet upplösningssgränsen för optiken och sensorn var för sig, kan man då uppskatta vad upplösningssgränsen blir för kombinationen optik plus sensor? Ja, det finns en approximativ formel för detta som lyder

$$\frac{1}{S_{bild}} \approx \frac{1}{S_{optik}} + \frac{1}{S_{sensor}} \quad (16.1)$$

där S_{bild} = upplösningssgränsen i bilden, S_{optik} = optikens upplösningssgräns och S_{sensor} = sensors upplösningssgräns. Denna formel kan också användas "baklänges", på så sätt att man kan räkna fram S_{sensor} om man har mätt upp S_{bild} och S_{optik} .

16.6. Sensorstorlekens inverkan

Som vi sett i detta kapitel så mäts upplösningen för kameror ofta som antal linjepar per mm i sensorplanet. När man jämför detta upplösningstal för olika kameror måste man emellertid ta hänsyn till sensorns storlek. Detta illustreras i nedanstående figur. Om man dubblar både bredden och höjden på en sensor, så innebär det att bara hälften så hög upplösning krävs för att få en slusbild med samma skärpa. I praktiken innebär detta praktiskt taget alltid att kameror med stora sensorer ger skarpare slusbilder. Naturligtvis krävs också att antalet pixlar i sensorn är tillräckligt högt för att kunna återge en bild med god skärpa. Bildens pixellering kommer att beröras närmare i kapitel 22.

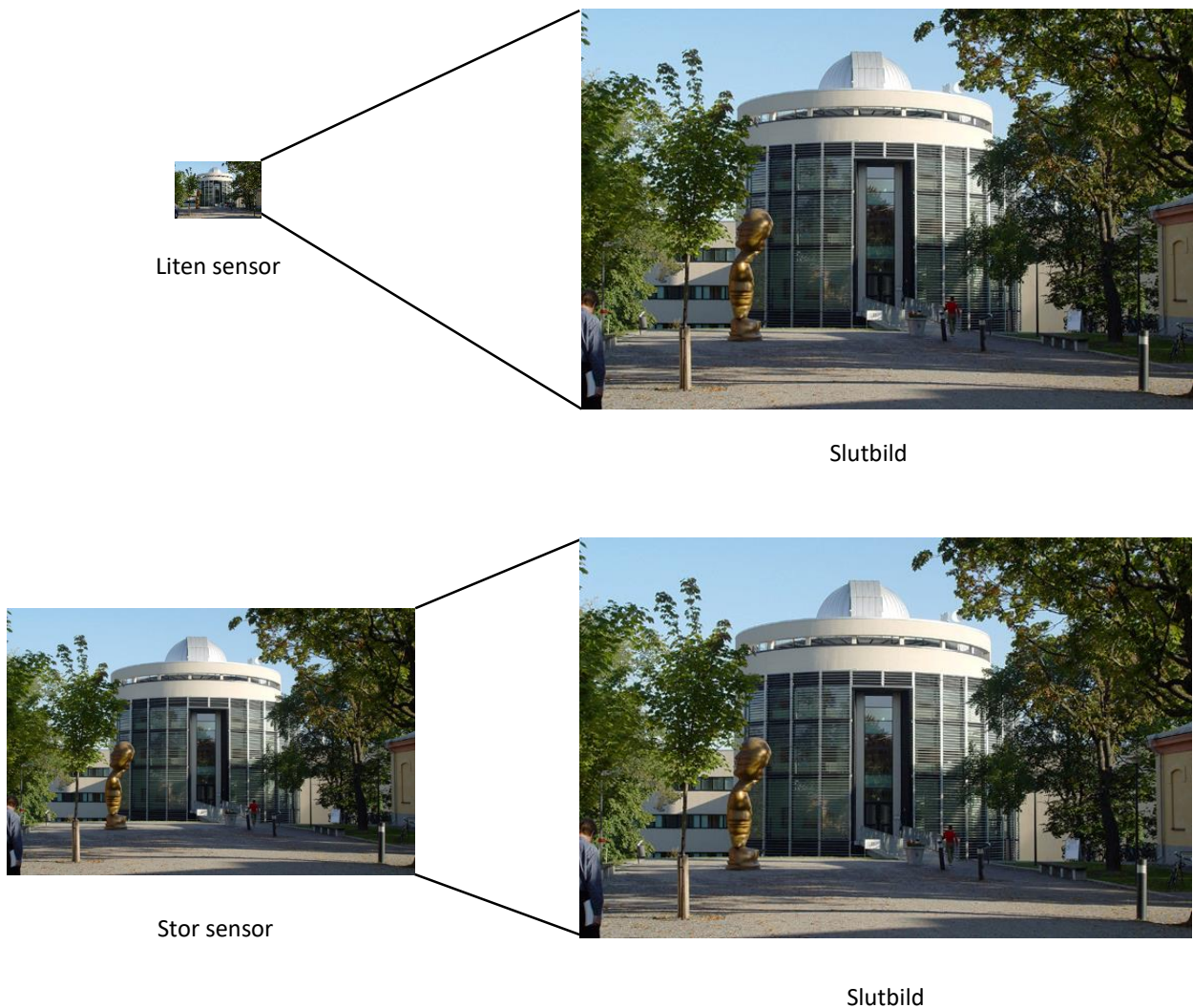


Fig. 16.7. Om en kamera har en liten sensor krävs en större förstoring för att få en slusbild i viss önskad storlek. Denna större förstoring medför att det ställs högre krav på skärpan i den ursprungliga bilden i sensorplanet.

16.7. Problem vid upplösningsmätningar

En nackdel med upplösningstester av det slag som här beskrivits, är att de är subjektiva. Olika personer kan komma fram till olika värden, och även en och samma person kan få olika resultat vid olika tillfällen. En annan nackdel är att upplösningförmågan säger långt ifrån allt om bildkvaliteten i ett system. Det är faktiskt tänkbart att ett system med en klart lägre upplösninggräns än ett annat system, ändå ger bilder som vi uppfattar som klart skarpare, Fig. 16.8. För att få ett riktigt bra kvalitetsmått behövs alltså mer information än vi får ur värdet på upplösningförmågan. I nästa kapitel ska vi se hur vi kan få ett sådant bättre kvalitetsmått.

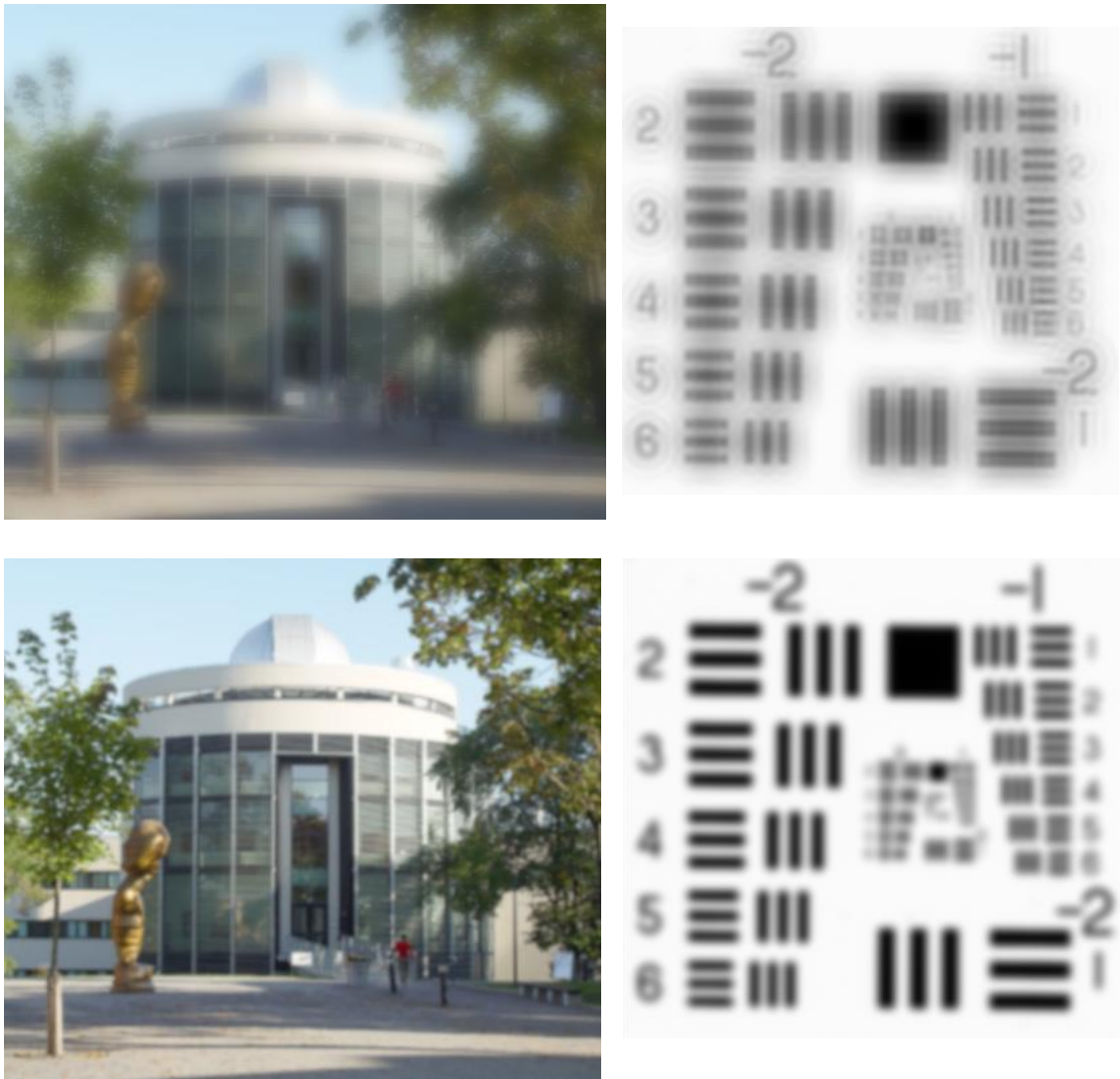


Fig. 16.8. Upplösningförmåga är inte alltid ett pålitligt mått på bildkvalitet. De två övre bilderna har framställts med ett avbildande system vars upplösningförmåga är 60% högre än för de nedre bilderna. Detta syns tydligt i bilderna på streckmönster. Trots den högre upplösningförmågan, upplever man de övre bilderna som suddigare. MTF-kurvor (se nästa kapitel) för de bägge systemen finns avbildade i Fig. 17.17.

17. Bildkvalitet: MTF

17.1. Inledning

När man arbetar med bildinformation är det ofta av stort intresse att kunna ange mått (helst objektiva och kvantitativa) på avbildningskvaliteten. Klassiskt har man använt begreppet upplösningsförmåga (se kap. 16), men denna är ofta svår att objektivt bestämma. Dessutom ger den inte alltid så bra information om bildens verkliga skärpa. Därför använder man idag ofta MTF (Modulation Transfer Function) för att beskriva ett systems avbildningskvalitet. I detta kapitel kommer en översiktlig presentation av detta ämne att göras för att förklara idén bakom det hela, samt vilken information man kan få ur MTF. (En utförligare presentation med den matematiska bakgrunden ges i Carlsson, K., Imaging physics, KTH 2016*.)

Låt oss först i mycket enkel bildmässig form förklara själva idén bakom MTF. I Fig. 17.1 visas ett tänkt motiv som ska fotograferas. Motivet består av ett linjemönster som blir tätare mot högerkanten. Under bilden av linjemönstret visas en luminansprofil för mönstret, dvs hur gråtonen varierar periodiskt. Maximala och minimala gråtonsvärdet är detsamma för alla linjer, dvs kontrasten är konstant. Vid fotografering av mönstret i Fig. 17.1 kommer de tätare linjerna att förlora mera i kontrast än de glesare linjerna såsom visas i Fig. 17.2. Tätare linjer kommer därför att synas otydligare, vilket vi upplever som oskärpa i bilden. Den streckade kurvan i Fig. 17.2 utgör kamerans MTF-kurva, vilken kommer att beskrivas i följande avsnitt. Den visar hur kontrasten i bilden minskar med ökande linjetäthet i motivet.

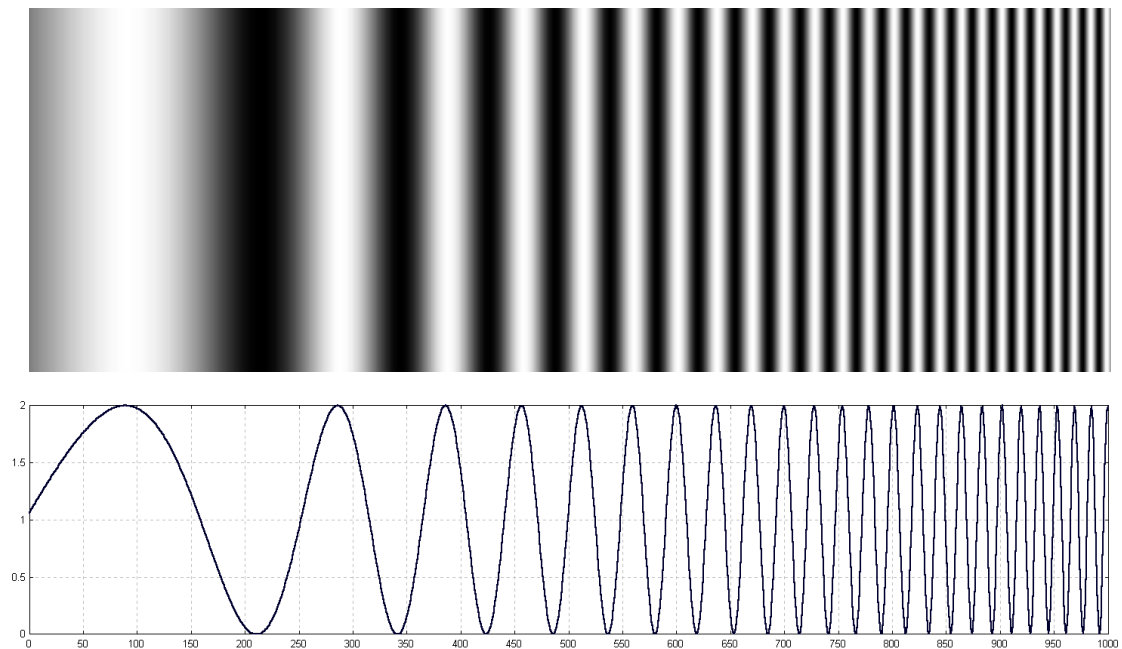


Fig. 17.1. Mönster med varierande linjetäthet, samt motsvarande luminansprofil.

*se <https://www.kth.se/social/files/57d11da0f276542d9216413a/Comp.ImagingPhysics.new.pdf>

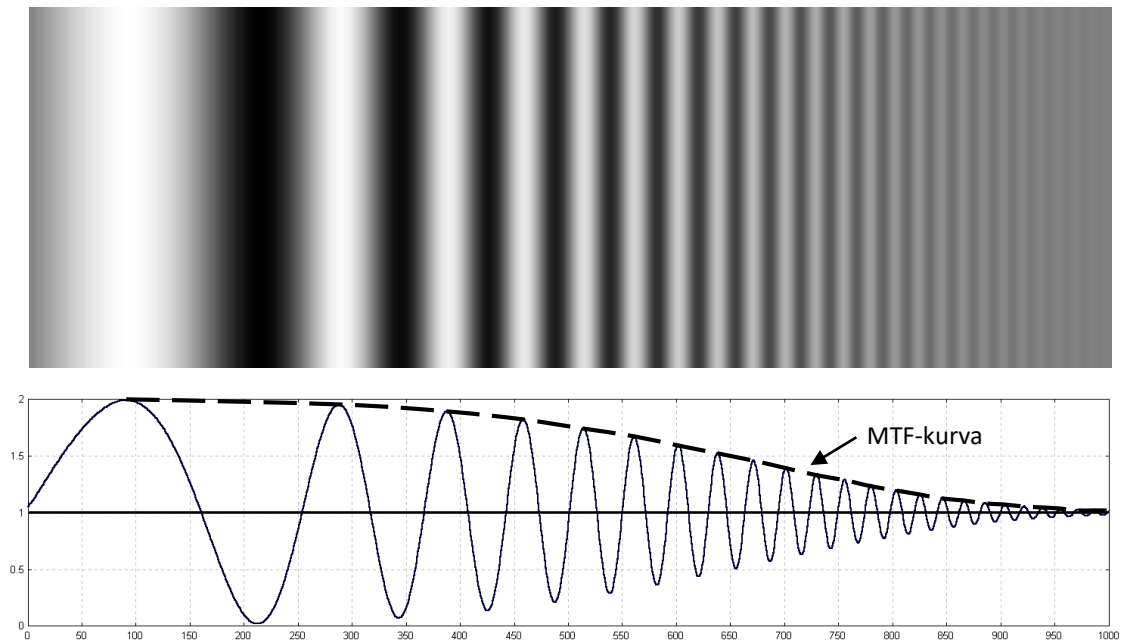


Fig. 17.2. Vid fotografering av mönstret i Fig. 17.1 kommer de tätare linjerna att förlora mer i kontrast än de glesa linjerna. Detta beskrivs av MTF-kurvan.

17.2. Punktspridningsfunktionen och MTF

Vi ska nu titta lite på den matematiska bakgrunden till MTF, och därvid införa begreppet punktspridningsfunktion. Låt oss betrakta ett generellt avbildande system enligt Fig. 17.3.

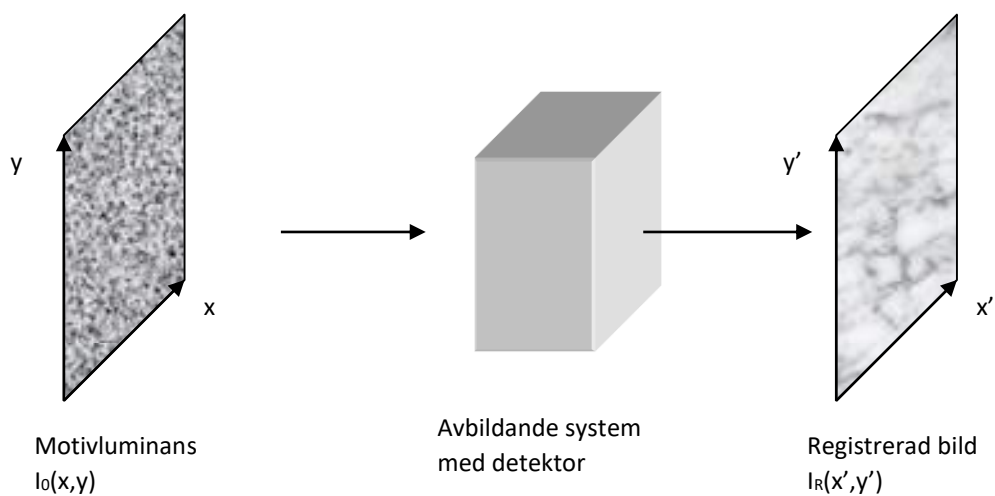


Fig. 17.3. Schematisk representation av en avbildningsprocess.

Vi antar att den optiska avbildningen är inkoherent*, så att ljusintensiteter kommer att adderas linjärt. Vidare antar vi att detektorns utsignal beror linjärt på exponeringen (vilket i allmänhet är fallet för CCD- och CMOS-sensorer). För enkelhets skull ska vi i fortsättningen betrakta endimensionella funktioner.

Vi börjar med att tänka oss ett mycket enkelt motiv, nämligen en enda punktkälla (oändligt liten lysande prick). Den registrerade bilden av detta motiv kallas systemets punktspridningsfunktion, $psf(x)$, vilken kan ses som den "oskärpefläck" vi får vid avbildning av punktkällan, Fig. 17.4.

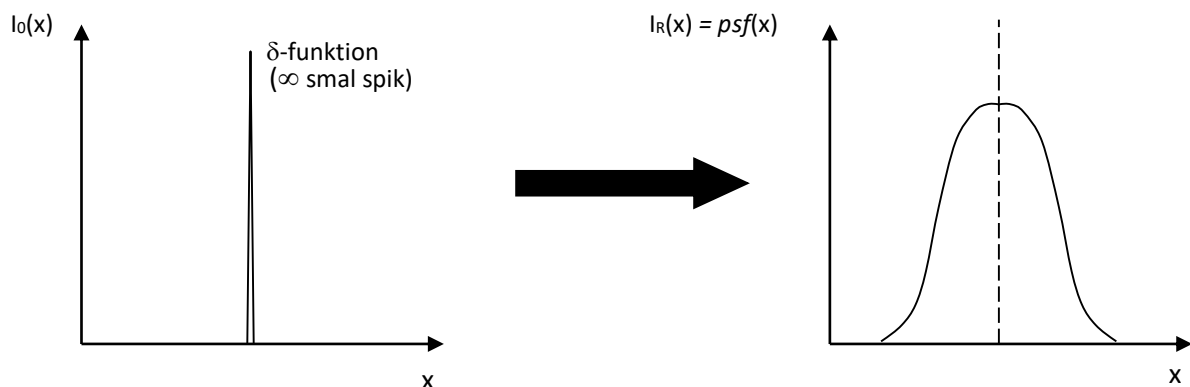


Fig. 17.4. Schematisk representation av hur en punktljuskälla avbildas.

Man kan visa att $psf(x)$ innehåller all information om systemets avbildningsegenskaper. Det betyder att om man känner psf så vet man precis hur bilden kommer att se ut för vilket godtyckligt motiv som helst.

Man kan i viss mån bedöma ett systems avbildningskvalitet genom att titta på dess psf . Naturligtvis ska ett bra system ha en så smal psf som möjligt. Idealet vore en psf som är en deltafunktion (∞ smal "spik"). Ett sådant system skulle kunna avbilda alla detaljer hos motivet, och man kan visa att motiv och bild då är identiska. I praktiken kan detta aldrig uppnås på grund av avbildningsfel och diffraktion i optiken, samt att sensorn också kommer att "sudda till" bildinformationen.

Istället för psf brukar man emellertid studera beloppet av dess Fouriertransform, vilket har fått namnet Modulation Transfer Function (MTF). Eftersom Fouriertransformen innehåller all information om originalfunktionen, förlorar man ingen information genom att transformera. Vinsten man gör är att informationen blir lättare att tolka om man använder MTF istället för psf . I detta kompendium ska vi inte fördjupa oss i Fouriertransformer, utan istället ska vi gå igenom vad MTF ger oss för intressant information. Rent praktiskt kan nämnas att MTF-data direkt kan mätas upp för t.ex. fotografiska objektiv med kommersiellt tillgänglig utrustning. Sådan information presenteras ofta i kurvform – ett typiskt exempel visas i Fig. 17.5.

* Inkoherent avbildning råder när ljuset som utsänds från en motivpunkt inte har någon bestämd fasrelation till ljus som sänds ut från andra motivpunkter. Olika punkter sänder alltså ut ljus oberoende av varandra. Detta gäller inte vid t.ex. laserbelysning av motivet.

17.3. Tolkning av MTF-kurvor, begreppet ortsfrekvens

MTF-kurvan i Fig. 17.5 visar ”frekvensgången” för systemet på ett sätt som är helt analogt med det man använder för elektriska kretsar, förstärkare etc. Skillnaden är att man här tittar på signaler som varierar sinusformigt i rummet och inte i tiden. För att åskådliggöra detta ska vi börja med att titta på begreppet ortsfrekvens. Antag att vi har ett fotografiskt objektiv med vilket vi avbildar ett regelbundet linjemönster. I bildplanet kan vi då se ett mönster enligt Fig. 17.6. Vi antar att belysningen varierar sinusformigt när vi rör oss i x- riktningen.

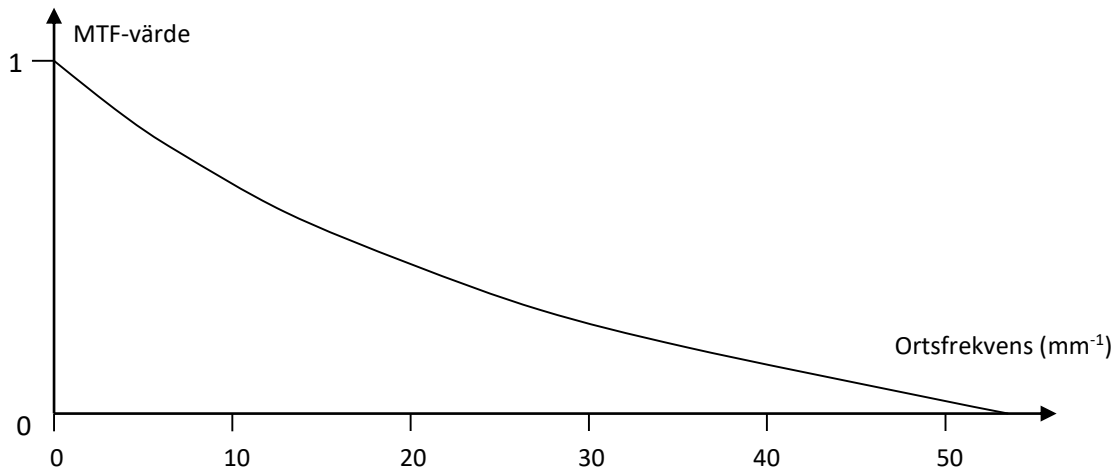


Fig. 17.5. En MTF-kurva visar MTF-värde som funktion av ortsfrekvens (dvs. den mönstertäthet som avbildas).

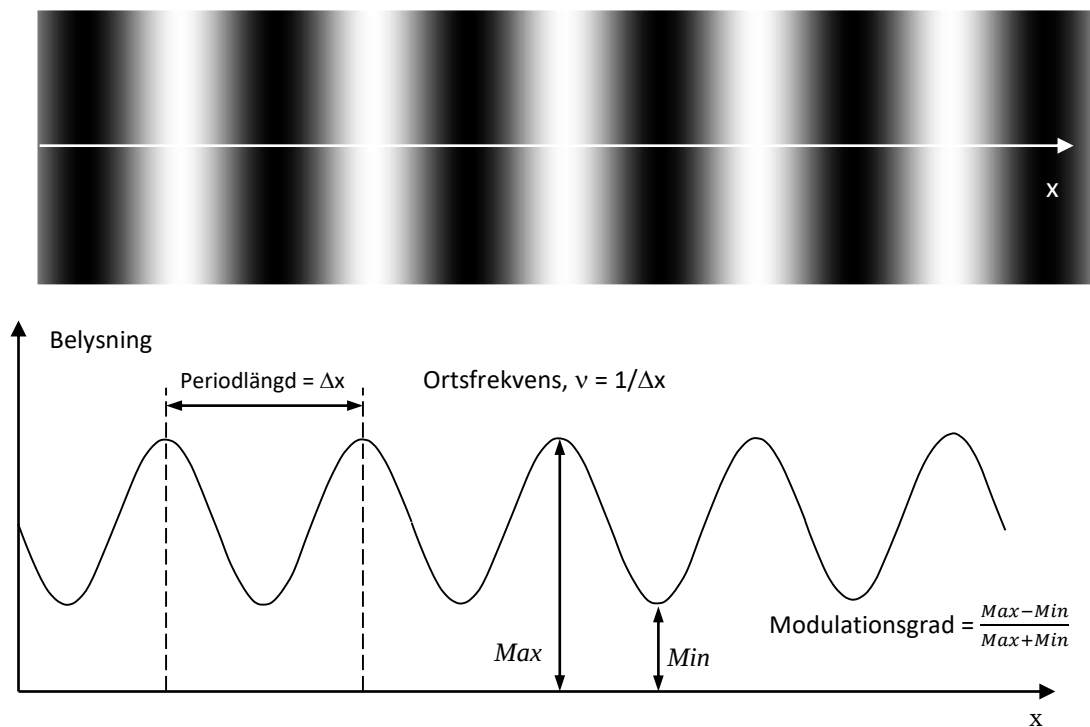


Fig. 17.6. Illustration av begreppen ortsfrekvens och modulationsgrad för ett linjemönster.

Ortsfrekvensen, som vi betecknar med ν (grek. ny), för detta mönster ges av $\nu = \frac{1}{\Delta x}$ = antalet svängningsperioder per längdenhet (precis som linjetätheten i

kap. 16). Sorten blir i SI-systemet m^{-1} , men för att inte få så stora mätetal brukar man ofta använda sorten mm^{-1} . Detta är alltså helt analogt med svängningar i tiden (och sorten Hertz) – den enda skillnaden är att vi nu jobbar med funktioner som varierar i rummet och inte i tiden. Att vi betraktar en sinusformig belysningsfördelning innebär ingen begränsning i resonemanget, eftersom en godtycklig funktion kan beskrivas som en summa av sinusfunktioner (Fourieranalys). Genom linearitetskravet vet vi också att en sinusformad motivluminans kommer att resultera i en sinusformig belysningsfördelning i bildplanet (ingen distorsion). Eftersom avbildningsskalan vid bildregistrering oftast är $\neq 1$ det viktigt att veta om ortsfrekvensangivelser gäller för motivet eller för bilden. Praxis är att man anger ortsfrekvenser i bildplanet. Om så ej är fallet brukar det tydligt anges.

Förutom ortsfrekvensen för det avbildade linjemönstret, så är även modulationsgraden, m , av intresse. Denna definieras som (se Fig. 17.6)

$$m = \frac{Max - Min}{Max + Min}. \quad (17.1)$$

Populärt uttryckt så anger modulationsgraden den kontrast linjemönstret har. Stort m anger hög kontrast, vilket vi uppfattar som en ”skarp” bild. Vi är nu redo att gå igenom hur MTF-värden kan tolkas.

MTF-värdet anger med vilken faktor modulationsgraden för ett sinusformat mönster av en viss ortsfrekvens förändras när vi går från motiv till bild, dvs $m_{bild} = m_{motiv} \times MTF$. Modulationsgraden kan aldrig öka i avbildningsprocessen, vilket medför att $MTF \leq 1$. För ortsfrekvensen noll är alltid $MTF = 1$. Inte helt oväntat är det så att MTF-värdet brukar avta mer eller mindre monotont när ortsfrekvensen ökar. Det innebär att kontrastförsämringen blir större ju tätare linjemönster (finare detaljer) vi avbildar. Mot denna bakgrund kan vi säga att en perfekt MTF-kurva skulle ha värdet = 1 för alla ortsfrekvenser. Det skulle innebära att vi kan avbilda hur täta linjemönster som helst utan att förlora någon kontrast (skärpa). För ett sådant system kan man visa att bildfunktionen blir en perfekt kopia av motivet, dvs. ingen information går förlorad. Detta svarar som tidigare nämnts mot en *psf* som är en deltafunktion. Inget verkligt avbildande system kan leva upp till detta. Optikens MTF kommer t.ex. att begränsas av diffraktion och avbildningsfel. I sensorn (CCD, CMOS eller fotografisk film) kommer ljusspridning och andra effekter att ”sudda till” bilden, vilket också ger en icke-ideal MTF. Fig. 17.7 illustrerar vilken information man kan utläsa ur en MTF-kurva.

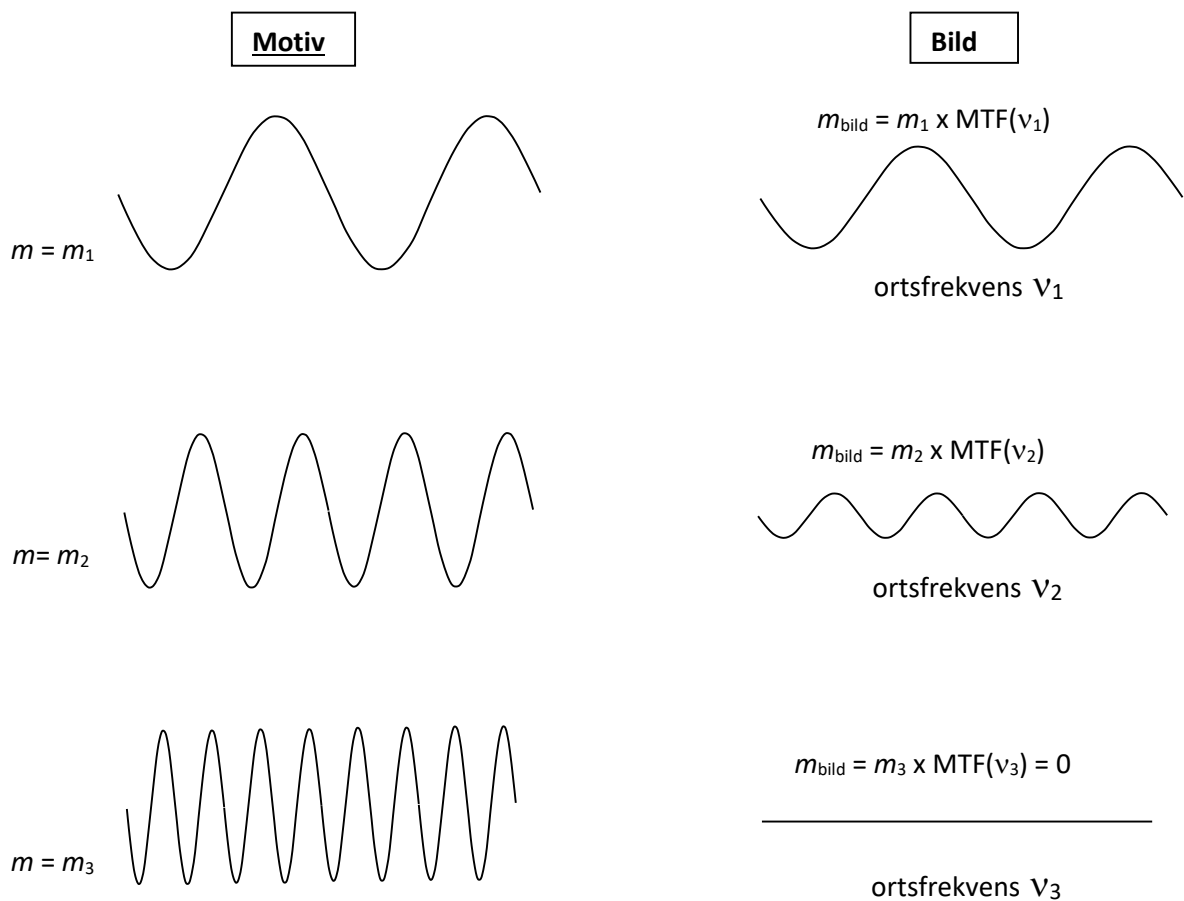
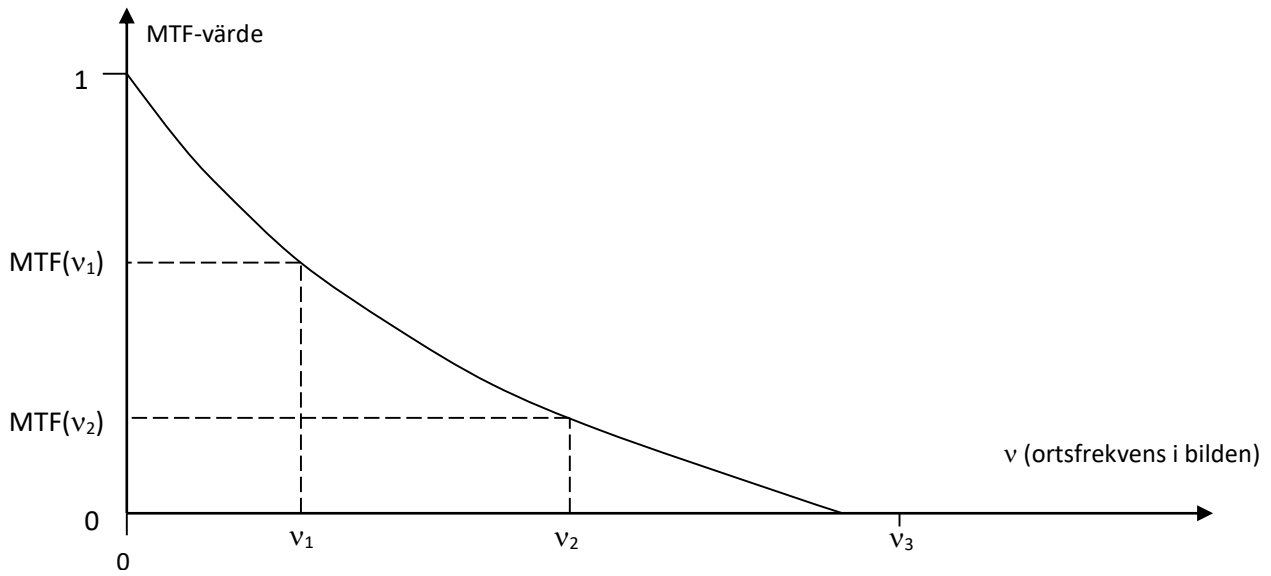


Fig. 17.7. MTF-värdet talar om med vilken faktor ett mönsters modulationsgrad (= kontrasten) minskar när vi går från motiv till bild. Kontrastförlusten blir i allmänhet större ju tätare mönster som avbildas.

17.4. Objektivets MTF

För all optik gäller att MTF är identiskt noll över en viss ortsfrekvens, den så kallade gränshfrekvensen. Om optiken helt saknade avbildningsfel skulle denna gränshfrekvens vara

$$v_{gr} = \frac{1}{\lambda \cdot F} \text{ (se fotnot)*} \quad (17.2)$$

där λ = våglängden och F är bländartalet. För ett diffraktionsbegränsat objektiv avtar MTF monotont upp till v_{gr} , se Fig. 17.8. Ett verkligt objektiv med avbildningsfel har en MTF-kurva som ligger lägre än den diffraktionsbegränsade. Vid höga bländartal (högre än ca. 5.6 eller 8) brukar fotografiska objektiva vara nästan diffraktionsbegränsade (åtminstone i bildcentrum). Både aberrationer och diffraktion är våglängdsberoende. Vid MTF-mätningar (och datorsimuleringar) på optik brukar man använda vitt ljus med ett väldefinierat spektralinnehåll.

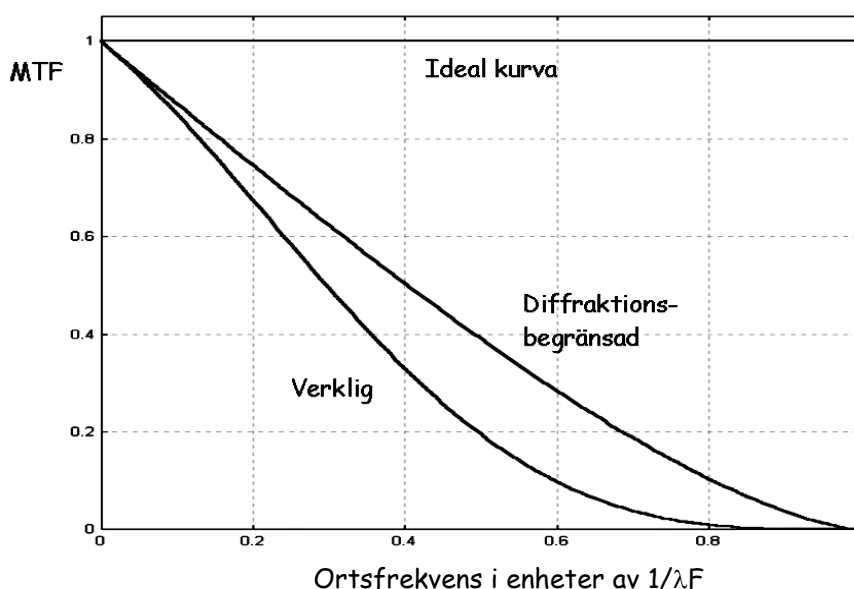


Fig. 17.8. MTF-kurvor för idealfallet (inga avbildningsfel eller diffraktion), ett diffraktionsbegränsat objektiv (inga avbildningsfel), samt ett verkligt objektiv (behäftat med både avbildningsfel och diffraktion). Eftersom diffraktionen aldrig kan upphävas, är den diffraktionsbegränsade kurvan det bästa man (nästan) kan åstadkomma för ett objektiv. Det matematiska uttrycket för MTF för ett diffraktionsbegränsat objektiv ges av ekv. 17.3.

* Enligt denna formel skulle v_{gr} gå mot oändligheten då bländartalet går mot noll (det finns ju inget rent teoretiskt som hindrar att bländartalet blir hur litet som helst). Formeln är emellertid bara en approximation till den exakta formeln $v_{gr} = \frac{2 \sin \beta}{\lambda}$, där β är vinkeln i Fig. 15.11. Den teoretiskt

högsta ortsfrekvensen som kan återges i bildplanet är alltså $2/\lambda$. Den approximativa formeln ovan gäller ned till ca. bländartal 1.0 om man tolererar 10% fel.

I Fig. 17.9 visas uppmätta MTF-kurvor för ett fotografiskt objektiv vid olika bländartal. Sämst (dvs. lägst) MTF-kurva får man vid bländartal 2. Detta beror på att avbildningsfelen är klart märkbara vid denna bländaröppning. Detta ger en suddighet i bilden, vilket drar ned MTF-värdena. Vid successiv nedbländning till 2.8, 4 och 5.6 sker en kontinuerlig förbättring av MTF-kurvan. Detta beror på att effekten av avbildningsfelen kommer att minska vid nedbländning. Bilden blir skarpare, och MTF högre. När man sedan bländar ned från 5.6 till 8 får man ingen förändring i MTF. Fortsätter man nedbländningen så kommer MTF successivt att försämrans. Detta beror inte på att avbildningsfelen återigen ökar. Avbildningsfelens inverkan kommer hela tiden att minska ju mer man bländar ned. Det som drar ned MTF-kurvan vid höga bländartal är istället diffraktionen. Bländaröppningen har helt enkelt blivit så liten att den oskärpefläck vi får på grund av diffraktion blir klart märkbar (ju mindre öppning, desto mer diffraktion). Det är ganska vanligt att man får optimal MTF-kurva runt bländartal 5.6-8. Man har då fått en optimal balans mellan avbildningsfel och diffraktion (jfr. Fig. 16.2).

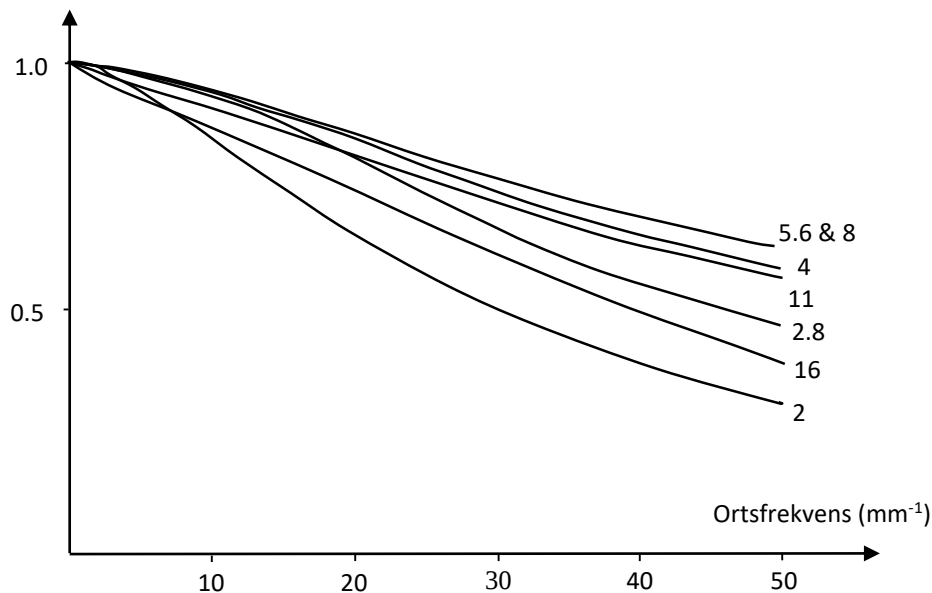


Fig. 17.9. MTF-kurvor vid olika bländartal för ett Leica 50 mm $f:2$ objektiv. Mätningarna är gjorda i bildcentrum.

Som vi såg i kap. 15 så kommer i allmänhet avbildningsfelen att bli större när man lämnar bildcentrum och rör sig ut mot kanten, vilket gör att man kan förvänta sig en större oskärpefläck. Likaså kan man förvänta sig att oskärpefläcken blir osymmetrisk på grund av astigmatism och koma. Det gör att MTF kommer att vara beroende av var i bildfältet man gör mätningen, samt orienteringen av de linjemönster som avbildas. För att fullständigt beskriva ett objektivs prestanda skulle det därför behövas ett mycket stort antal MTF-kurvor. Av praktiska skäl måste man naturligtvis begränsa sig, och det är vanligt att man presenterar MTF-kurvor för bildcentrum och en bit ut mot bildkanten (t.ex. 2/3 ut mot hörnet). I sista fallet presenterar man kurvor för två vinkelräta mönsterorienteringar, radiellt och tangentiellt, se Fig. 17.10. Endast en kurva behövs däremot för bildcentrum, eftersom oskärpefläcken där är rotationssymmetrisk även när man har avbildningsfel (koma och astigmatism finns ju inte i centrum).

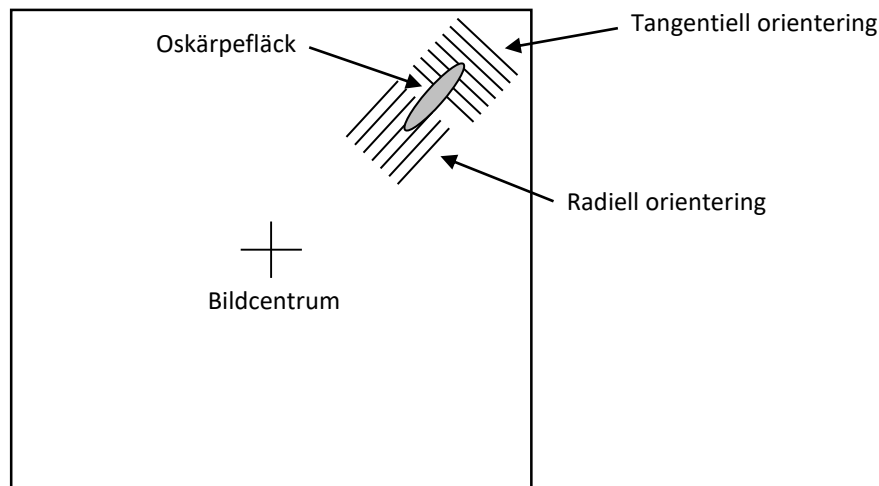


Fig. 17.10. Vid mätning av MTF utanför bildcentrum brukar man registrera två kurvor, en för radiell mönsterorientering och en för tangentiell mönsterorientering. Den tangentiella kurvan brukar bli sämre än den radiella beroende på att oskärpefläcken oftast är mer utbredd i radiell riktning.

Exempel på hur MTF-kurvor i bildcentrum och bildkanten kan se ut illustreras i Fig. 17.11. Eftersom MTF beror på bländartalet, så presenteras en sådan skala av MTF-kurvor oftast för två eller tre olika bländartal.

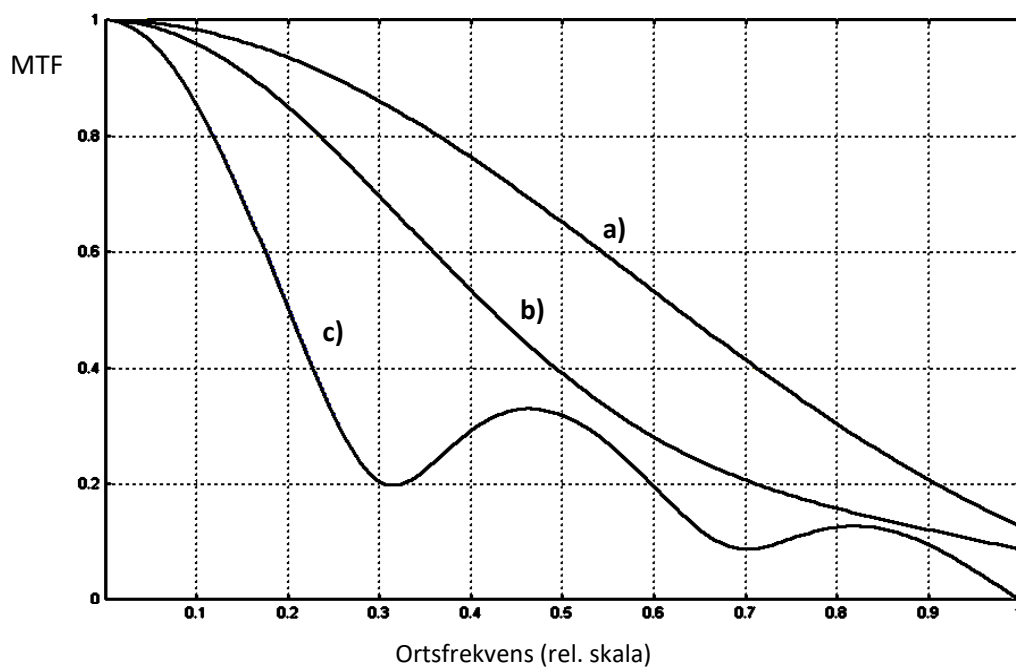


Fig. 17.11. Exempel på hur MTF-kurvor kan se ut i bildcentrum respektive bildkanten. a) Bildcentrum. b) Bildkanten, radiell mönsterorientering. c) Bildkanten, tangentiell mönsterorientering. Skillnaden mellan kurvorna b) och c) beror på att oskärpefläcken inte är rotationssymmetrisk, jfr. Fig. 17.10.

Hittills har objektivets MTF-värde ritats upp som funktion av ortsfrekvensen. Ett annat sätt att illustrera MTF, är att rita upp MTF-värdet vid en viss ortsfrekvens som funktion av avståndet från bildcentrum. Sådana kurvor används ofta när objektivtester publiceras t.ex. i fototidskrifter. I Fig. 17.12 visas exempel på sådana kurvor. Ofta visas i samma diagram fyra kurvor som representerar MTF för två olika bländartal och för radiell och tangentiell mönsterorientering. Om radiella och tangentiella kurvorna ligger nära varandra innebär det att oskärpefläcken är nära rotationssymmetrisk, och höga MTF-värden innebär att den har liten utsträckning. Fördelen med denna typ av kurvor är att de ger en bra uppfattning om hur bildskärpan varierar över bildplanet. En nackdel är att de bara visar MTF-värdet vid en enda ortsfrekvens. Man behöver egentligen veta MTF-värdet vid flera frekvenser för att kunna bedöma bildkvalitén. Ibland visas därför hur MTF-värdet vid ett par olika ortsfrekvenser varierar över bildplanet, men om man samtidigt ska ha med olika bländartal och mönsterorienteringar så blir det många kurvor att hålla rätt på.

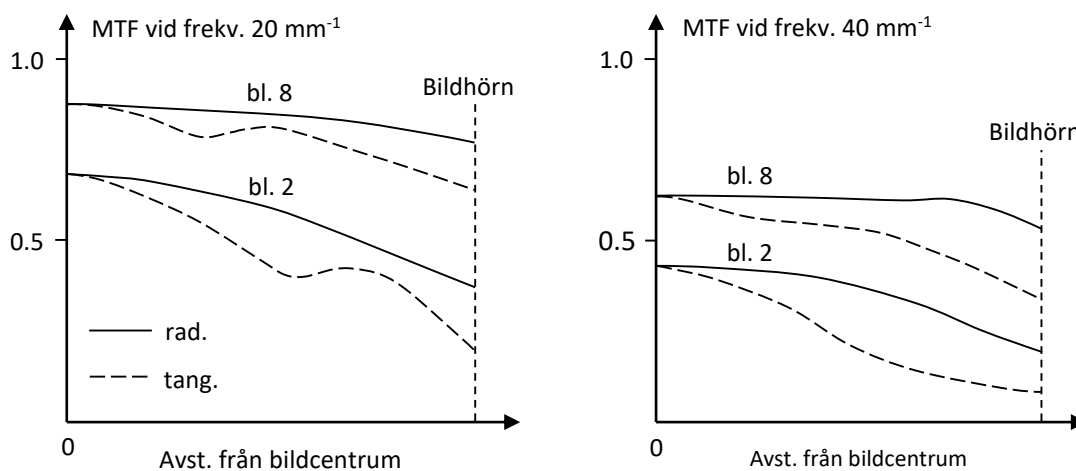


Fig. 17.12. Det är ofta av intresse att veta hur MTF varierar över bilytan. Då används diagram som visar hur MTF vid två eller tre ortsfrekvenser (i detta fall 20 och 40 mm^{-1}) varierar som funktion av avståndet från bildcentrum. Ofta visas kurvor för ett par olika bländartal, och man brukar ta med både radiell och tangentiell mönsterorientering eftersom detta ger värdefull information om bildkvalitén.

Objektivets MTF begränsas inte bara av diffraktion och avbildningsfel, utan även av ströljus. Inverkan av ströljus illustreras i Fig. 17.13. Ströljus kommer att sänka MTF procentuellt ungefär lika mycket vid alla ortsfrekvenser, utom för de allra lägsta (alla MTF-kurvor börjar ju på värdet 1 vid ortsfrekvensen noll). Ju högre ströljusnivå desto mer kommer MTF-kurvan att sänkas. Ju mer utbrett ströljuset är över bilytan, desto snabbare sjunker MTF-kurvan när ortsfrekvensen ökar från värdet noll. På grund av mätapparens utformning kommer dock ströljusets inverkan oftast inte med vid MTF-mätningar på optik. Istället mäter man ibland upp ströljusnivån separat (se kap. 9.15).

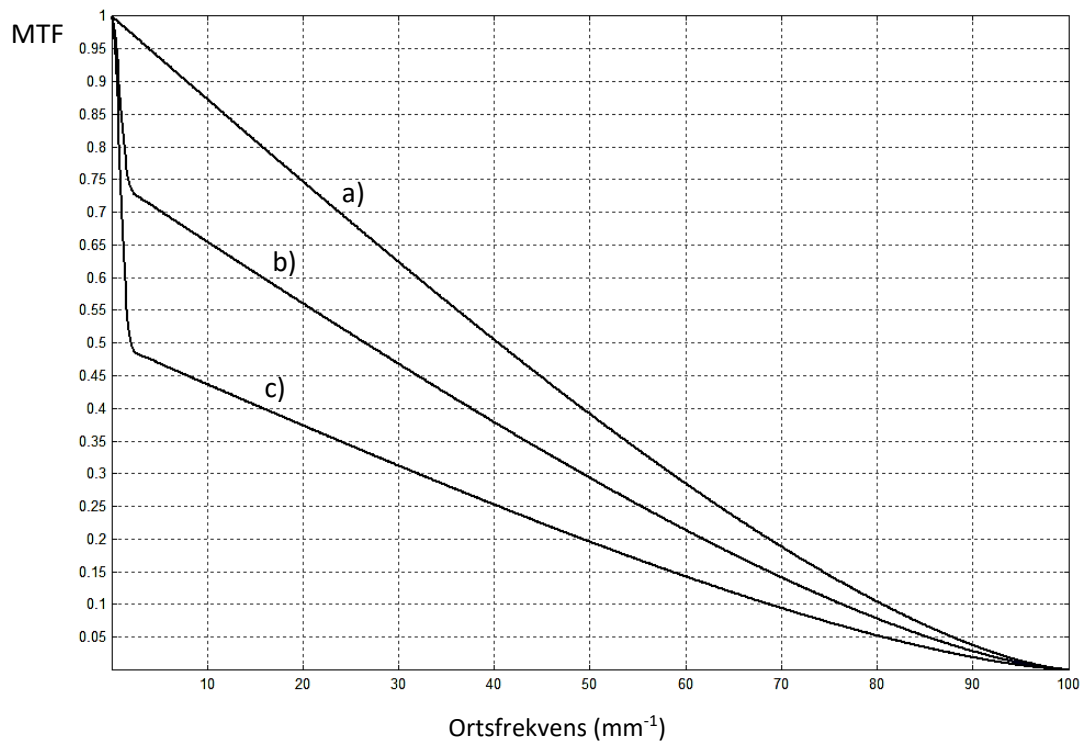


Fig. 17.13. Exempel på hur MTF-kurvan påverkas av ströljus i objektivet. a) Utan ströljus. b) Med ströljus. c) Med ännu mera ströljus. I normala fall är ströljuskivån så låg att effekten inte blir så dramatisk som i figuren.

17.5. MTF för hela avbildningskedjan

Ofta har man i en avbildningskedja flera faktorer som påverkar avbildningskvalitén, t.ex. optik, sensor, rörelseoskärpa mm. Om inverkan från var och en av dessa faktorer kan beskrivas med en MTF-kurva, gäller att hela kedjans MTF ges av $MTF_{total} = MTF_{optik} \cdot MTF_{detektor} \cdot MTF_{rörelseoskärpa} \cdot \dots$. Detta kallas för multiplikationsregeln för MTF, vilken gör det möjligt att beskriva hela kedjans avbildningsegenskaper om man känner de enskilda komponenternas. Ett exempel på detta visas i Fig. 17.16.

17.6. MTF för CCD- och CMOS-sensorer

MTF för en ideal matrissensor med element som har en fönsterbredd (pixelbredd) a kan approximeras med $MTF(\nu) = \left| \frac{\sin(a\pi\nu)}{a\pi\nu} \right|$ (den som vill veta varför hänvisas till Carlsson, K., Imaging Physics, KTH 2016). Exempelvis ger en fönsterbredd av $10 \mu\text{m}$ en MTF enligt Fig. 17.14. Som jämförelse ges en MTF-kurva för en fotografisk film i Fig. 17.15.

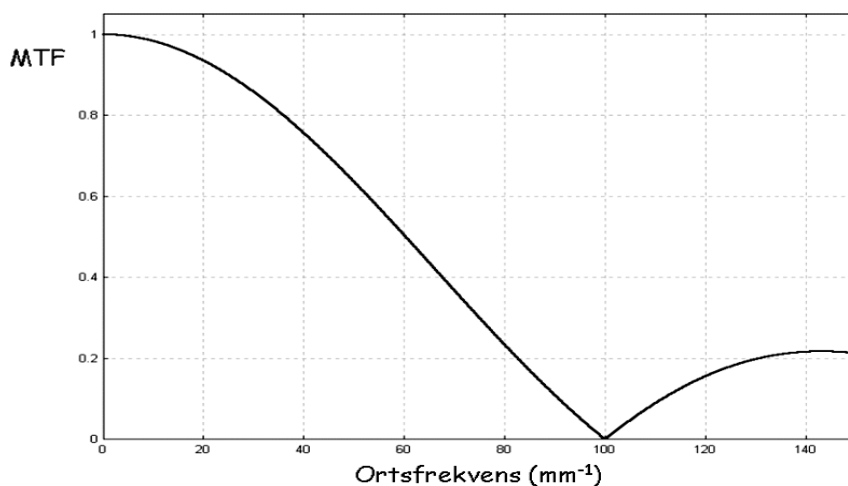


Fig. 17.14. En ideal sensor (ingen laddningsdiffusion etc.) med en fönsterbredd av $10\ \mu\text{m}$ har en MTF-kurva med ovanstående utseende. MTF för andra pixelbredder än $10\ \mu\text{m}$ fås enkelt genom att skala om ortsfrekvensaxeln. Pixelbredden $5\ \mu\text{m}$ kommer sålunda att ge ett nollställe vid $200\ \text{mm}^{-1}$.

Ortsfrekvenser över första nollstället kommer att återges 180° fasförskjutna, vilket innebär att detta frekvensområde är mindre användbart. På grund av icke-ideala egenskaper blir MTF för en verklig sensor sämre än vad som visas i ovanstående kurva.

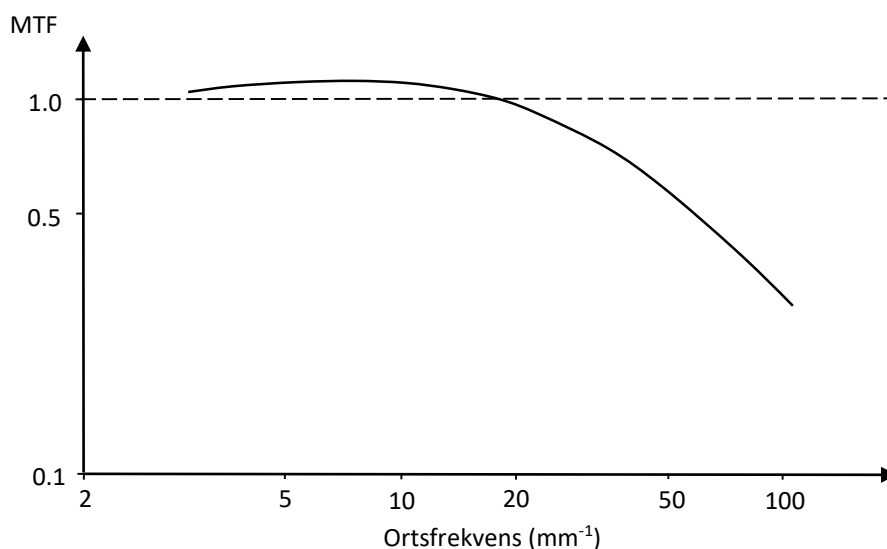


Fig. 17.15. Exempel på MTF-kurva för svartvit negativfilm (Kodak Tri-X). Observera att bägge axlarna har logaritmisk skalning, vilket är vanligt för film-MTF. Värden > 1 beror på felaktigheter orsakade av framkallningseffekter.

En verklig matrissensor kommer att ha en MTF som ligger lägre än vad som visas

i Fig. 17.14. En mer realistisk modell ges av funktionen $MTF_{\text{sensor}} \approx \left| \frac{\sin(a\pi\nu)}{a\pi\nu} \right|^2$.

Modellen bygger på erfarenhet från tester på digitalkamerabilder. (Resultatet stämmer väl med Fig. 22.7 i detta kompendium.)

Total MTF för en digitalkamera fås genom att multiplicera objektivets MTF med sensorns MTF, Fig. 17.16.

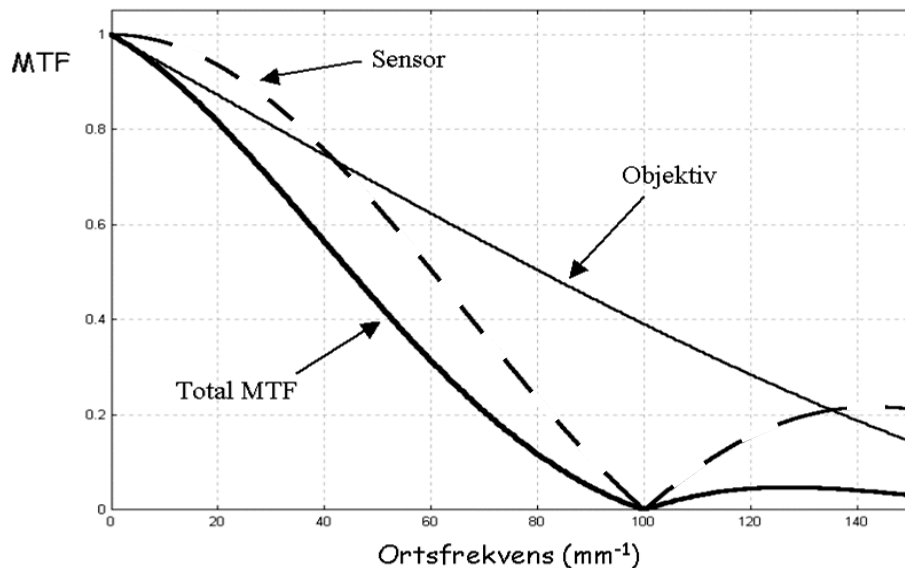


Fig. 17.16. Total MTF för en digitalkamera fås som produkten av objektivets och sensorns MTF. Kurvorna i figuren är idealiserade; mera realistiska kurvor visas i Fig. 17.24.

17.7. Tolkning av MTF-kurvor

I Fig. 17.8 är det inte svårt att direkt se vilken av kurvorna som är bäst och vilken som är sämst. En MTF-kurva ska ju ligga så högt som möjligt, och sträcka sig till så höga ortsfrekvenser som möjligt. Men i många fall korsar MTF-kurvor varandra, så att en kurva är högst vid låga ortsfrekvenser och en annan är högst vid höga ortsfrekvenser. Vi ser fall av detta både i Fig. 17.9 och 17.16. Hur ska man i sådana fall bedöma vilken kurva som är bäst? Låt oss titta på ett ännu mer extremt fall än de som illustreras i Fig. 17.9 och 17.16. I Fig. 17.17 ser vi MTF-kurvorna för de två system som användes för att framställa bilderna i Fig. 16.8 i föregående kapitel.

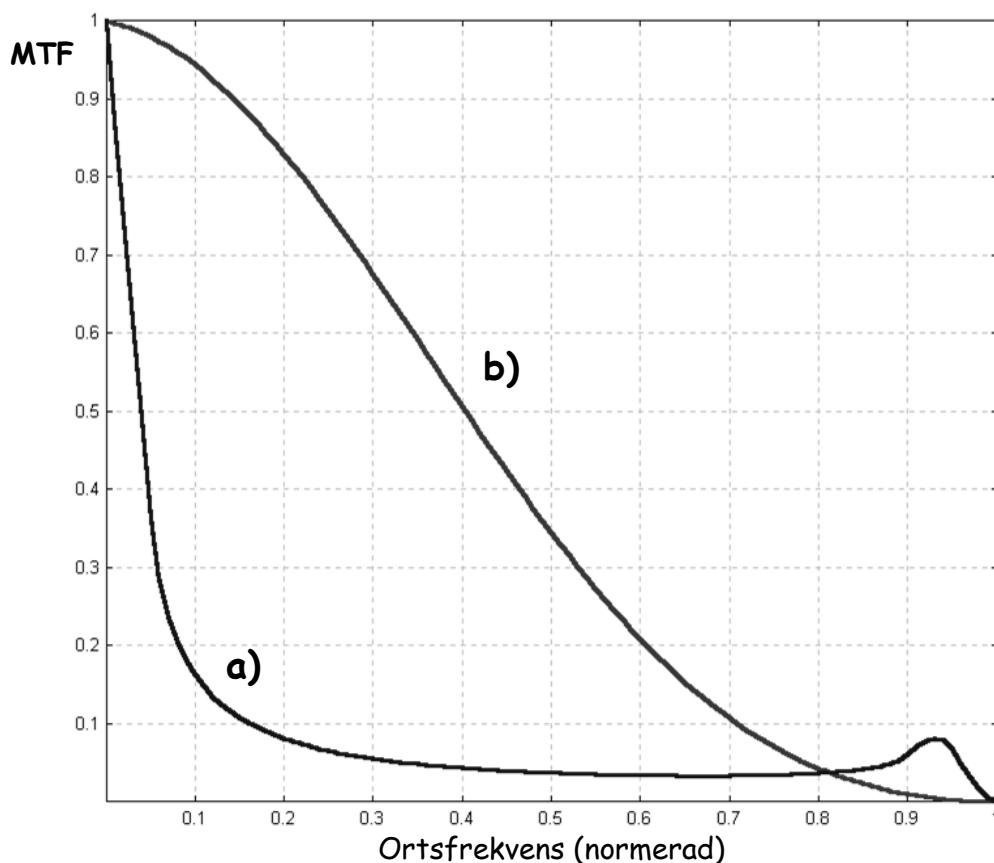


Fig. 17.17. MTF-kurvor för de avbildande system som användes för att framställa bilderna i Fig. 16.8. a) MTF för systemet med den högre upplösningförmågan. b) MTF för systemet med den lägre upplösningförmågan.

När man tittar på bilderna i Fig. 16.8 ser man att b-systemet ger en betydligt bättre upplevd bildkvalitet än a-systemet. Detta trots att a-systemet har klart bättre upplösningförmåga. Denna bättre upplösningförmåga för a-systemet beror på en högre MTF vid ortsfrekvenser nära gränshänsynsfrekvensen. Men det som är förödande för a-systemet är att MTF faller ner till så låga värden redan vid låga och medelhöga ortsfrekvenser. Det är detta som ger ett allmänt suddigt intryck i bilderna. Sens moral blir alltså att det viktigaste är att MTF-kurvan har så höga värden som möjligt vid låga till medelhöga ortsfrekvenser. Det är inte lika viktigt vad som händer vid de allra högsta frekvenserna.

17. 8. Jämförelser av MTF för kameror med olika sensorformat

För digitalkameror finns ett stort antal olika sensorformat (se kap. 14). I digitala kompaktkameror finns ofta små sensorer, t.ex. 4 mm x 6 mm, medan det i mer avancerade kameror såsom systemkameror ofta sitter sensorer som är i så kallat fullformat (24 mm x 36 mm). Även större sensorer förekommer i avancerade professionella kameror. I kapitel 16 (Upplösningförmåga) såg vi att man måste ta hänsyn till sensorstorleken när man jämför upplösningvärden mellan olika kameror. På motsvarande sätt måste man ta hänsyn till sensorstorleken när man jämför MTF-kurvor för olika kameror. I annat fall blir inte jämförelsen rättvis. För att illustrera detta kan vi titta på Fig. 17.18.

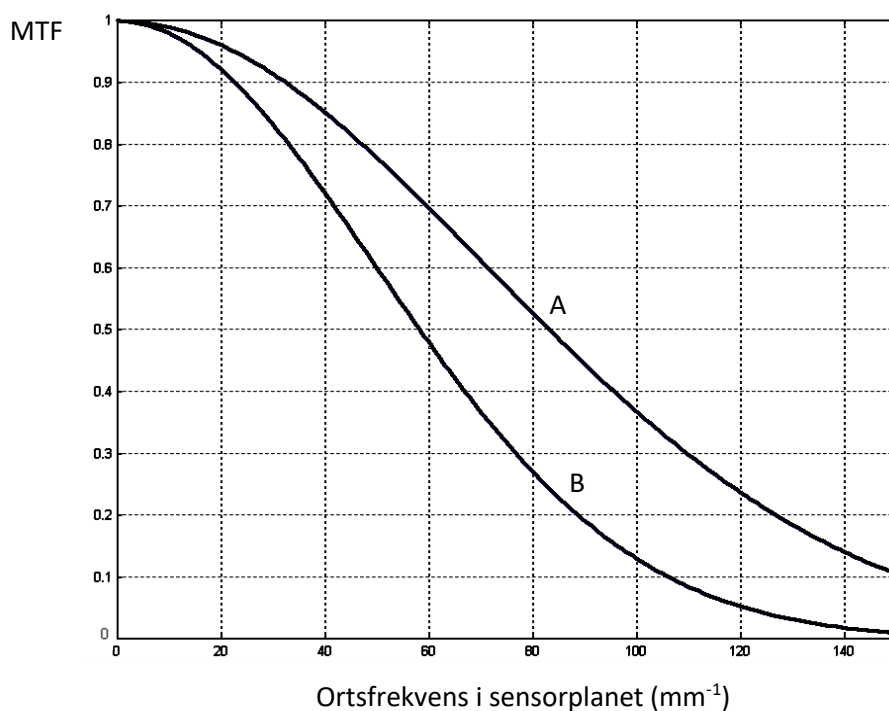


Fig. 17.18. Kurvor som visar total MTF för två kameror. Kamera A har en sensorstorlek av 5.0 mm x 7.5 mm, och kamera B har en sensorstorlek av 10 mm x 15 mm.

I Fig. 17.18 ligger kurva A över kurva B för samtliga ortsfrekvenser, så därför borde det inte vara någon tvekan om att kamera A ger skarpare bilder än kamera B. Kan man tycka. Om vi emellertid studerar data i bildtexten, ser vi att kamera B har en sensor som är dubbelt så bred och dubbelt så hög som sensorn i kamera A. Detta innebär att om man fotograferar samma motiv med bägge kamerorna, och så att lika mycket av motivet kommer med på bägge bilderna, så kommer alla motivets ortsfrekvenser att återges dubbelt så höga i kameran A som i kameran B. För att motivets detaljer ska återges lika skarpt i bägge bilderna, så krävs alltså att kamera A's MTF-värden är lika höga som kamera B's värden då man i A-fallet avläser vid en dubbelt så hög ortsfrekvens som i B-fallet. Ur figuren ser man att så inte är fallet. Exempelvis har B-kurvan för 40 mm⁻¹ ett MTF-värde av 0.72, medan A-kurvan vid 80 mm⁻¹ bara har ett värde av 0.53. Kamera B är alltså klart bättre än kamera A. Man kan lätt övertyga sig om att detta också stämmer vid andra ortsfrekvenser. Sensorstorlekens betydelse vid tolkning av MTF diskuteras ytterligare i slutet av detta kapitel.

17.9. Praktisk mätning av MTF för digitalkameror

Att i praktiken mäta upp MTF-kurvor är inte så enkelt. Det finns till exempel många felkällor som måste undvikas (Är testmönstrets kontraster noggrant kända? Finns det icke-lineariteter i systemet?). Det kan också vara jobbigt att mäta upp ett tillräckligt stort antal värden så att man får en välbestämd kurva. Sedan många år

finns professionell mätutrustning för kameraobjektiv som möjliggör snabb och noggrann mätning. Men sådan utrustning är mycket dyr, och dessutom mäter den inte upp sensorns egenskaper. Man får därför inte med hela kamerans MTF i mätningarna. För att mäta kamerans totala MTF har det utvecklats en standardiserad metod som kommer att kortfattat beskrivas i detta avsnitt. Metoden finns fastlagd i två ISO-dokument (nr. 12233 och 14524). Den kallas ”slanted edge method”, alltså ”snedkants-metoden”. Ett antal praktiska mätresultat på digitalkameror kommer också att presenteras och kommenteras.

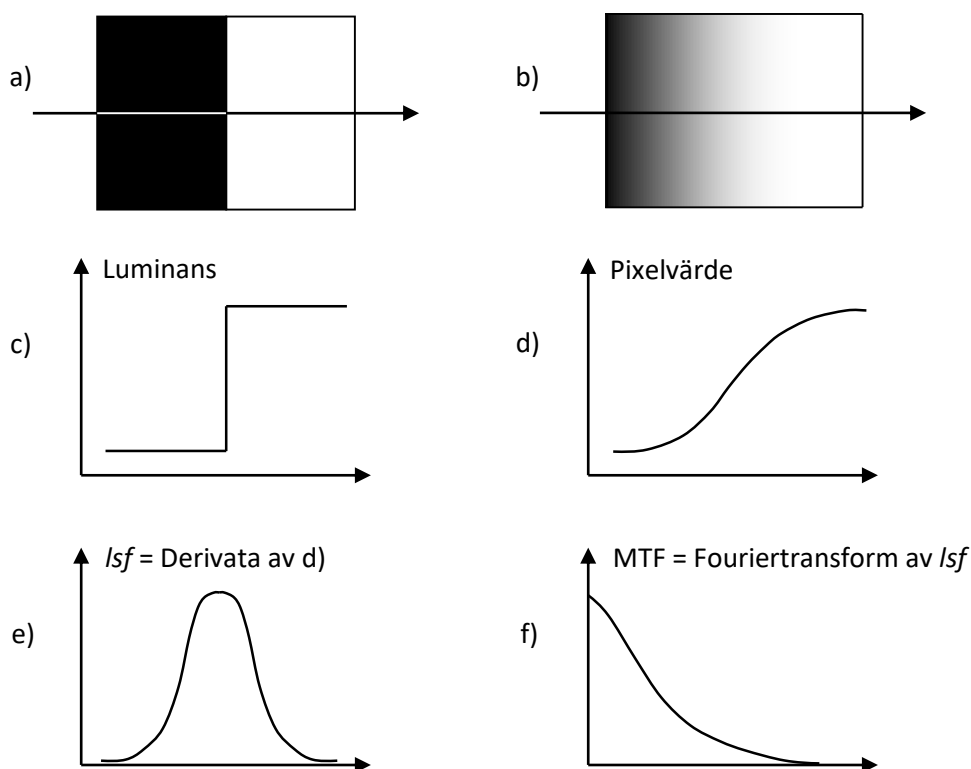


Fig. 17.19. Schematisk beskrivning av hur mätning av en digitalkameras MTF görs enligt ISO-normen. a) Motivets innehåller en skarp svartvit kant. b) Bilden av kanten kommer inte att återges helt skarpt pga. optikens och sensorns begränsningar. c) Luminansprofil för motivets skarpa kant. d) Pixelprofil uppmätt i bilden (stegsvaret). e) Pixelprofilen deriveras för att ge linjespridningsfunktionen (lsf). f) Fouriertransformen av lsf ger MTF.

Principen för MTF-mätningar på digitalkameror är mycket enkel och billig. Den är därför lämplig också för fotoamatörer. Metoden bygger på att vi mäter upp kamerans linjespridningsfunktion, lsf . Detta är motsvarigheten till punktspridningsfunktionen, psf (Fig. 17.4), för det fall att motivet är en oändligt tunn linje istället för en punkt. Av praktiska skäl är det lättare att avbilda en svartvit kant än en linjeljuskälla, och ur kantsvaret (stegsvaret) kan man räkna fram lsf såsom framgår av Fig. 17.19. För att minska pixellerings effekter i stegsvaret (Fig. 17.19 d) så låter man den svartvita kanten i motivet avvika några grader från vertikal riktning. Detta gör att man ur den digitala bilden kan erhålla en tätare sampling av

stegsvaret än vad som ges av avståndet mellan pixlarna i sensorn (sampling finns närmare beskrivet i kap. 22). Principen för denna "supersampling" illustreras i Fig. 17.20, och ett exempel på uppmätt stegsvar visas i Fig. 17.21

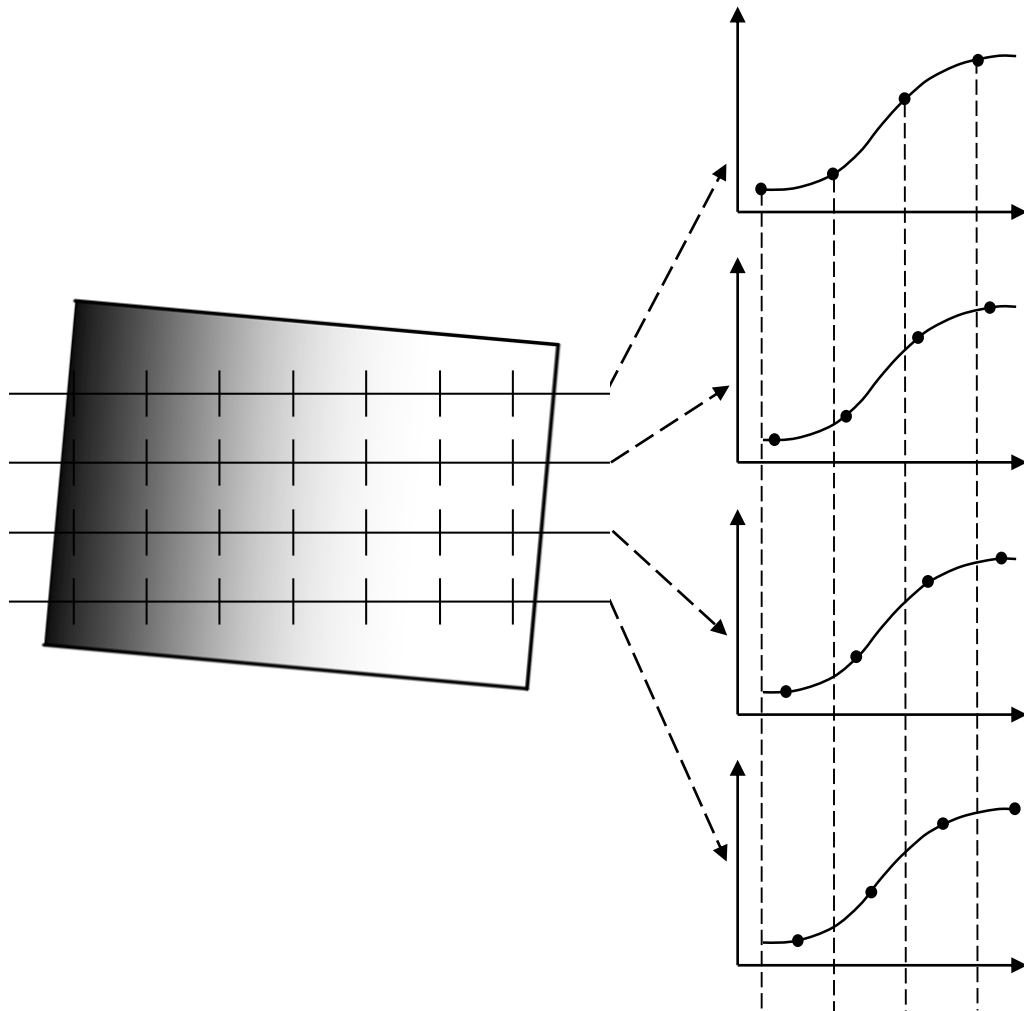


Fig. 17.20. Vid fotograferingen låter man motivets svartvita kant vara vriden några grader från vertikalriktningen. Inne i kameran innebär detta att de olika pixelraderna i sensorn kommer att ligga något förskjutna i sidled relativt bilden av kanten. Genom att kombinera värden från olika pixelrader kan man på detta sätt få en tätare sampling, och därmed mindre hackighet i det registrerade stegsvaret, än om kanten hade varit vertikal.

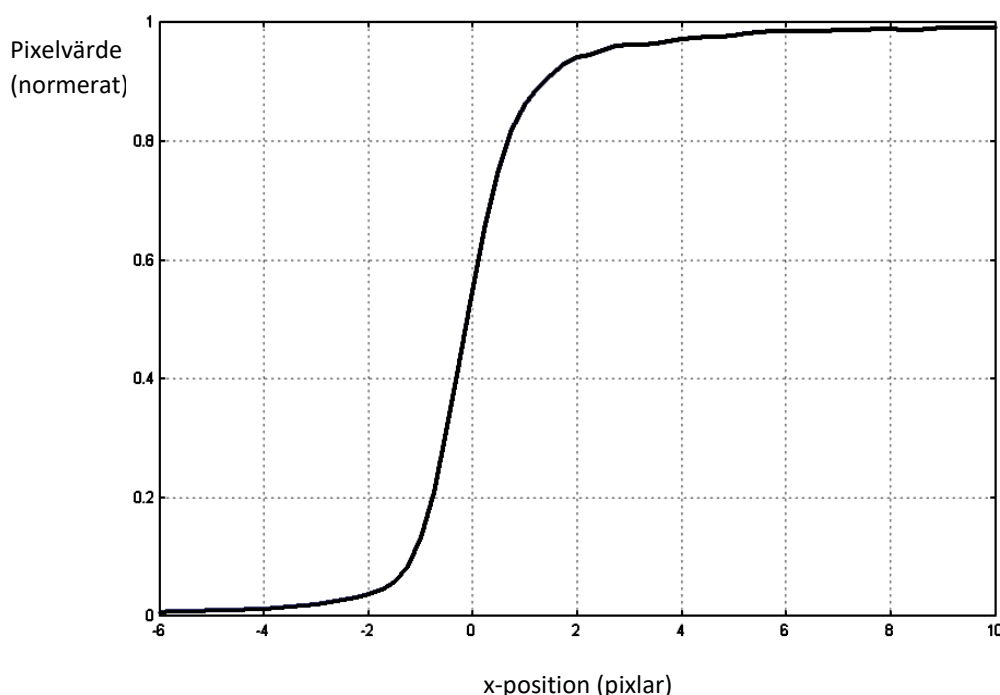


Fig. 17.21. Exempel på ett experimentellt uppmätt stegsvar från en digitalkamera (SONY DSC-R1: Sensorstorlek 14.4 mm x 21.5 mm, 10 Mpixlar). Genom att "supersampling" utförts enligt beskrivningen i Fig. 17.20, erhålls en mycket mjukare kurva med fler värden än om man bara hade gjort avläsningar vid hela pixelvärden.

Med hjälp av programvara* kan man göra MTF-beräkningar för digitalkameror enligt ISO-normerna. Förutom MTF kan man också erhålla *lsf*, mått på vissa avbildningsfel mm. I Fig. 17.22 visas MTF-kurvor för en digitalkamera (SONY A7, sensorstorlek 24 mm x 36 mm, 24 Mpixlar) vid några olika bländartal. Precis som i Fig. 17.9 ser man att bästa MTF erhålls ungefär vid bländartal 5.6, eftersom det ger en god balans mellan avbildningsfel och diffraktion. Ortsfrekvensaxeln i Fig. 17.22 är graderad i perioder/pixel (på engelska = cycles/pixel), vilket är mycket vanligt i MTF-kurvor för digitalkameror. Med kännedom om centrum-till-centrum avståndet mellan pixlarna i kameran (i detta fall 6.0 μm), kan man räkna ut att 1 period/pixel motsvarar ortsfrekvensen $\frac{1}{6.0 \times 10^{-6}} = 1.67 \times 10^5$ perioder/m ≈ 170

perioder/mm i bildplanet. En anledning till att man ofta återger ortsfrekvensaxeln graderad i perioder/pixel, är att det ger god information om hur allvarliga problem man kan förvänta sig med moiré-effekter i bilderna (kap. 22.4). Idealt vill vi ha en MTF-kurva med så höga värden som möjligt upp till 0.5 perioder/pixel. Över denna gräns ska däremot MTF-värdena vara så låga som möjligt för att minimera moiré-effekter.

* Det finns både kommersiell mjukvara och freeware som utför denna typ av beräkningar, se <https://www.kth.se/social/files/5c86765456be5bdfb084187e/MTFm%C3%A4tningar.pdf>.

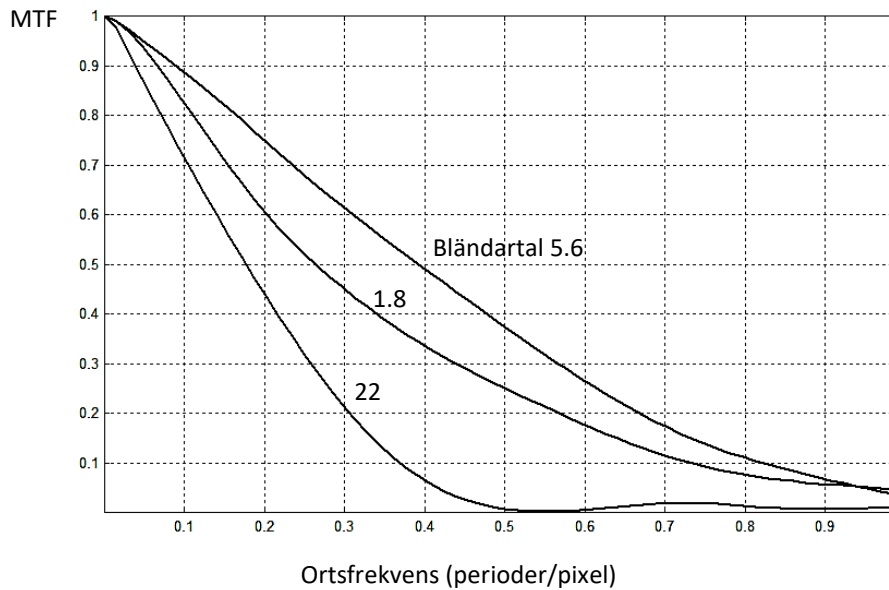


Fig. 17.22. MTF-kurvor vid olika bländartal för digitalkameran SONY A7 utrustad med ett 55 mm objektiv. Bästa resultat erhålls vid bländartal 5.6. När bländarinställningen ändras från 1.8 till 5.6 reduceras avbildningsfelen märkbart genom att bländaröppningen reduceras. Ytterligare mindre bländaröppning (högre bländartal) reducerar avbildningsfelen ytterligare, men istället kommer diffraktion att sudda till bilden. Det finns alltså ett bländartal som utgör den bästa kompromissen vad gäller avbildningsfel och diffraktion.

I Fig. 17.23 visas ett exempel på MTF-kurvor för bildcentrum och bildkant, i bägge fallen för bländartal 5.6. Att MTF-kurvan är lägre i bildkanten beror på att optiken har större avbildningsfel där (se kap. 15).

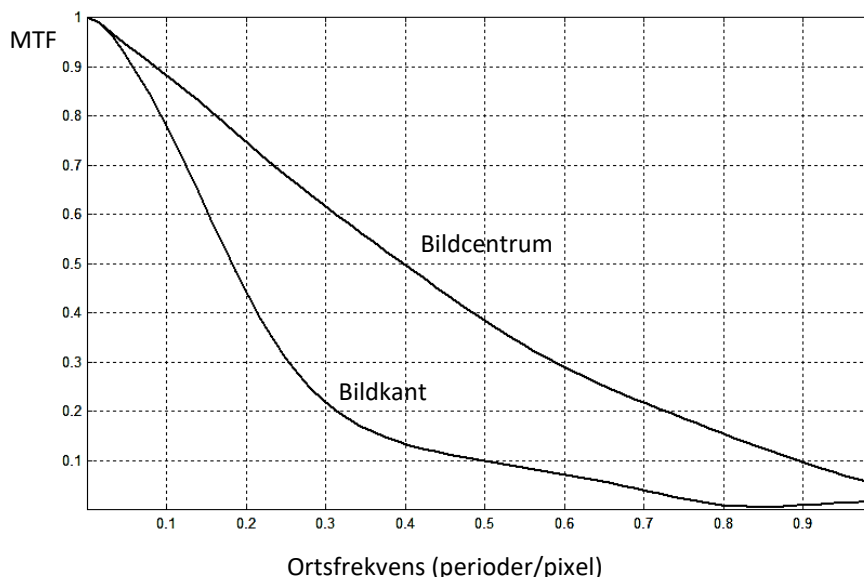


Fig. 17.23. MTF-kurvor för bildcentrum samt bildkanten. Bilden är suddigare i kanten på grund av avbildningsfel som astigmatism och koma, och därför kommer MTF-kurvan att vara lägre där. Kamera: SONY A7 med zoom inställd på 50 mm brännvidd. Bländartal 5.6.

MTF-kurvorna i Fig. 17.22 och 17.23 återger kamerans totala MTF, dvs de är påverkade av både optik och sensor. Ibland kan det emellertid vara intressant att studera dessa komponenter var för sig. På så sätt kan man avgöra hur mycket var och en begränsar bildkvaliteten. Om kameran har löstagbar optik (vilket är fallet för systemkameror) kan denna testas separat. Detta kräver dock som nämnts avancerad mätutrustning, och är knappast något som amatörer ger sig på. Samma sak gäller MTF-mätningar på sensorn. För många kameror gäller att optiken inte går att avlägsna från kamerahuset, vilket gör att optik och sensor helt enkelt inte kan testas separat. Det finns dock en möjlighet att uppskatta sensorns MTF ur kamerans totala MTF. Enligt kap. 17.5 så gäller att $MTF_{total} = MTF_{optik} \cdot MTF_{sensor}$ (om fotograferingen görs med stativ, och i övrigt under ordnade former, kan troligen skakningsoskärpa mm. försummas). Om vi använder ett bländartal av ca. 8 eller högre, kan vi förvänta oss att ett bra objektiv är nästan diffraktionsbegränsat (kap. 17.4), åtminstone bör detta gälla i bildcentrum. För ett diffraktionsbegränsat objektiv gäller att

$$MTF(x) = \frac{2}{\pi} \left[\arccos(x) - x\sqrt{1-x^2} \right] \quad (17.3)$$

I ekv. 17.3 är $x = \frac{\nu}{\nu_{gr}}$, där ν är verklig ortsfrekvens och

$\nu_{gr} = \text{gränsfrekvensen} = \frac{1}{\lambda F}$. Denna funktion finns uppritad i Fig. 17.8. Om

bländartal F och våglängd λ är givna, vet vi alltså vilken MTF vi kan förvänta oss för objektivet. Ekvation 17.3 gäller exakt för monokromatiskt ljus, men vi kan förvänta oss att den också gäller någorlunda bra för ljus med en något utbredd spektralfördelning, om vi som λ stoppar in ett medelvärde. Detta har utnyttjats i Fig 17.24 som visar tre MTF-kurvor: a) Total MTF, b) teoretisk diffraktionsbegränsad MTF för optiken och c) MTF_{sensor} erhållen genom beräkning ur formeln

$MTF_{sensor} = \frac{MTF_{total}}{MTF_{optik}}$. Kameran är en SONY DSC-R1. Vid fotograferingen

användes ett gröntfilter ($\lambda \approx 540$ nm) på kameran för att få en mer väldefinierad ljusvåglängd. Jämfört med teoretisk MTF_{sensor} (kap. 17.6) så ligger kurvan i Fig. 17.24 betydligt lägre. Detta kan delvis bero på att sensorn inte har ideala egenskaper, men det torde också bero på att sensorn är utrustad med ett optiskt lågpasfilter (se kap. 22.5) vilket reducerar MTF.

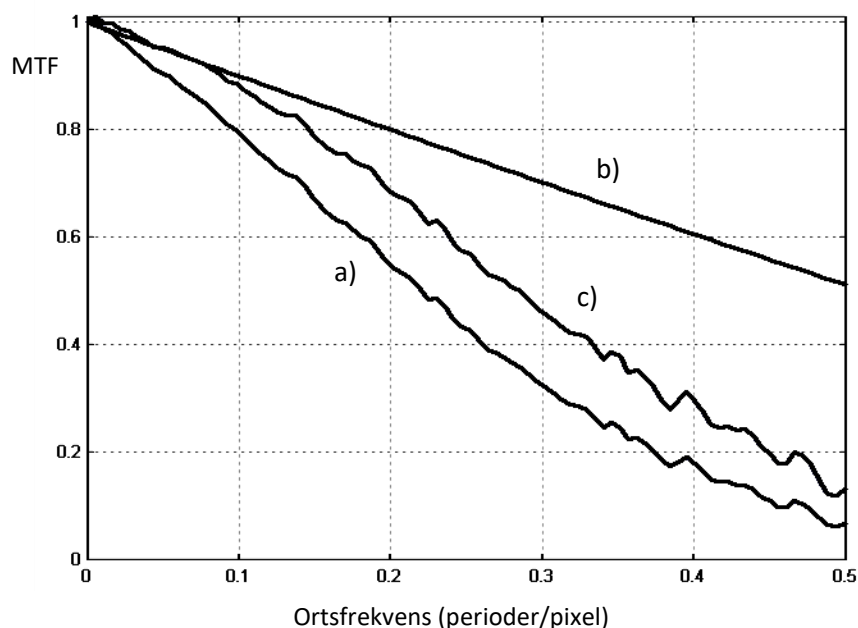


Fig. 17.24. En uppskattning av MTF_{sensor} kan erhållas genom att anta att optiken är diffraktionsbegränsad. Detta är i allmänhet ett hyggligt antagande för bländartal runt 8 och högre. a) Kamerans totala MTF. b) MTF för diffraktionsbegränsad optik vid bländartal 8 och 540 nm ljusvåglängd. c) Uppskattad MTF_{sensor} erhållen genom att dividera total MTF med optikens MTF. Kamera: SONY DSC-R1. Ojämnheten i a- och c-kurvorna orsakas av brus i mätningarna.

För att testa stabiliteten i resultatet i Fig 17.24 upprepades proceduren för bländartal 16. MTF_{sensor} överensstämde mycket bra för de två fallen, utom vid höga ortsfrekvenser då kurvan blev mycket brusig för fallet med bländartal 16. Detta berodde på att MTF_{total} blev mycket låg vid höga ortsfrekvenser på grund av diffraction i objektivet. När sedan MTF_{sensor} räknades fram ur dessa låga MTF_{total} värden blev bruset kraftigt förstärkt.

Vid MTF-mätningar på digitalkameror, kommer resultatet att kraftigt påverkas av om kameran eller den efterföljande bildbehandlingen gör en mjukvarumässig "sharpening" (se kap. 17.14) för att få bilderna att se skarpare ut. Olika grader av sharpening kan ofta väljas i kamerans meny. Sådan sharpening görs nästan alltid på bildfiler som sparas i jpeg-format. Enda sättet att helt undvika sharpening är att spara bilderna i så kallat RAW-format, men detta är ofta inte möjligt i enklare kameror. **Om MTF-mätningar görs på bilder som utsatts för sharpening blir resultatet i stort sett värdelöst.** I och med att sharpening har utförts har man infört en kraftig icke-linearitet, vilket raserar hela den matematiska grunden för MTF. Resultatet går inte att tolka på samma sätt som en korrekt MTF-kurva, och kan inte jämföras med en korrekt kurva. En ytterligare viktig synpunkt är att MTF ska visa verkliga prestanda hos kamerans komponenter, dvs. optik och sensor, och inte inbegripa bildbehandling vare sig denna sker i kameran eller vid efterföljande datorbearbetning. Som exempel på hur förvridna resultaten kan bli av sharpening, visas i Fig. 17.25 stegsvar och MTF-kurvor för tre olika grader av sharpening. Kameran är en SONY DSC-R1. Kurvan utan sharpening har erhållits ur en RAW-fil. De övriga kurvorna har erhållits från jpeg-filer med två olika kamerainställningar på sharpening (max. och min.). Man kan observera att även när

kamerans sharpening ställs till minimum så kan den ha en klart märkbar effekt på MTF (men detta kan variera för olika kamerafabrikat och modeller).

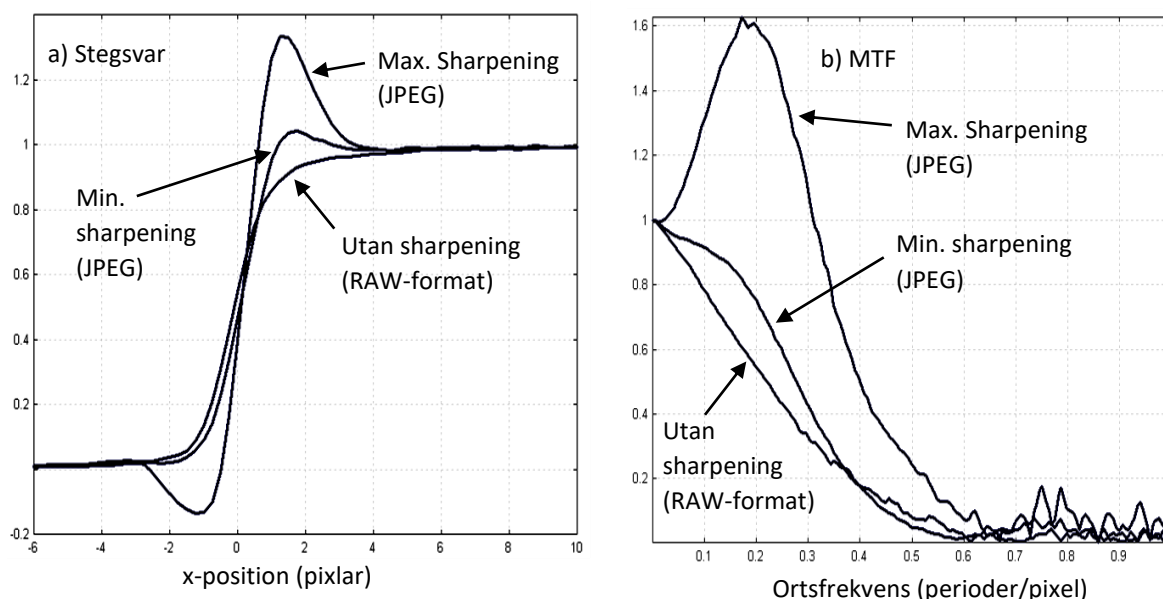


Fig. 17.25. Ofta görs redan i kameran en bearbetning av digitalbilden för att den ska se skarp ut. Sådan "sharpening" ger kraftigt utslag i uppmätta MTF-kurvor, och gör det omöjligt att bedöma kamerans prestanda vad gäller optik och sensor. För att undvika sådana sharpening-effekter helt måste man i allmänhet använda bilder i RAW-format. Alla kameror har dock inte denna möjlighet. a) Stegsvär för en och samma kamera vid olika nivåer av sharpening. b) Motsvarande MTF-kurvor. Vid maximal sharpening går MTF-värdet upp ända till 1.6 i en kraftig puckel. Detta är naturligtvis rent nonsens. Ett MTF-värde är alltid ≤ 1 .

För att bedöma om sharpening har utförts, och i så fall uppskatta om den är kraftig eller måttlig, kan man titta på kamerans stegsvar. Om detta varken visar underslängar vid låga pixelvärden eller överslängar vid höga pixelvärden kan man sluta sig till att ingen, eller mycket måttlig, sharpening har gjorts. I sådana fall kan man tro på att MTF-kurvan blir riktig. Vissa försök har gjorts att räkna om stegsvaret från kameror med sharpening till en "standardiserad" sharpening. Detta för att man ska kunna jämföra MTF för kameror som har olika grad av sharpening i bilderna. Detta är en tveksam metod av flera skäl. Dels innebär sådan extra manipulation att man inte riktigt vet hur resultatet påverkas. Dessutom faller hela den matematiska grunden för MTF-teorin samman, vilket medför att man inte riktigt vet vad kurvorna egentligen visar. Slutligen vill man att MTF ska visa kameraoptikens och sensors prestanda, oberoende av bildbehandling, eftersom man kanske vill jämföra med MTF för andra bildskapande tekniker.

När man vill jämföra MTF-kurvor för kameror med olika sensorstorlekar och/eller olika antal pixlar, måste man ge akt på hur ortsfrekvensaxeln är graderad. Ofta är det nämligen nödvändigt att räkna om ortsfrekvensvärdena innan man kan jämföra

MTF-värden. Detta har redan berörts i avsnitt 17.8. Sålunda gäller att om ortsfrekvensaxeln är graderad i mm^{-1} (perioder per millimeter), så måste man räkna om värdena utgående från sensorns totala mått. Låt oss ta ett exempel:

Antag att vi jämför två kameror, a och b , med sensorstorlekar 15 mm x 22.5 mm respektive 10 mm x 15 mm. Då ska man jämföra kamera a 's MTF-värden vid en viss ortsfrekvens (t.ex. 100 mm^{-1}) med kamera b 's MTF-värden vid en 1.5 gånger högre frekvens (150 mm^{-1}). Orsaken är att kamera a ger en 1.5 gånger så stor bild på sensorn som kamera b . Detta medför i sin tur att motivets alla ortsfrekvenser återges en faktor $\frac{1}{1.5}$ lägre i kamera a än i kamera b .

I MTF för digitalkameror anger man ofta ortsfrekvensaxeln i enheten perioder/pixel. Om man inte gör några omräkningar framstår då en kamera med få pixlar i sensorn som bättre än den egentligen är, medan det motsatta gäller för en kamera med många pixlar. Låt oss ta ett exempel på detta också:

Antag att vi jämför två kameror, a och b , med 2000 x 3000 pixlar respektive 3000 x 4500 pixlar (i detta fall spelar sensorns geometriska storlek ingen roll). Ortsfrekvensen 0.5 perioder/pixel svarar i fallet a mot 1000 perioder längs sensorn kortsida, medan den i fallet b svarar mot 1500 perioder längs sensorns kortsida. I fallet b motsvarar alltså 0.5 perioder/pixel ett tätare mönster i motivet. Botemedlet mot detta är att jämföra MTF-värdet för sensor a vid 0.5 perioder/pixel med b -sensorns MTF-värde vid ortsfrekvensen

$$0.5 \times \frac{2000}{3000} = 0.33 \text{ perioder/pixel.}$$

För att slippa omräkningar vid jämförelse av sensorer med olika storlek och/eller olika antal pixlar, så har man infört ytterligare ett sätt på vilket man kan ange ortsfrekvensen i MTF-kurvor. Genom att ange ortsfrekvensen i perioder per sensorhöjd kan man direkt jämföra MTF-kurvor för kameror med olika sensorstorlekar och antal pixlar. För att omvandla en ortsfrekvens från perioder/mm till perioder/sensorhöjd, så multiplicerar man helt enkelt med sensorns höjd i mm. För att omvandla från perioder/pixel, så multiplicerar man med antalet pixlar på sensorns höjd. Man skulle naturligtvis lika gärna kunna välja att ange perioder/sensorbredd istället, men det vanliga är att man använder sensorhöjden. I fotolitteraturen anges ofta ortsfrekvensen i "line widths per picture height". Denna enhet ger precis dubbelt så höga tal som enheten perioder/sensorhöjd, eftersom man räknar både svarta och vita linjer separat istället för att titta på hela perioder av ett linjemönster.

Att uttrycka ortsfrekvensen i perioder/sensorhöjd (eller line widths per picture height) är att föredra när man ska jämföra olika kameror. Oberoende av antalet pixlar i sensorn, eller dess fysiska dimensioner, kan man då direkt jämföra MTF-kurvorna med varandra utan att göra några omräkningar. Att detta är möjligt beror på att ett linjemönster i motivet kommer att i bilden få precis samma ortsfrekvens mått i perioder/sensorhöjd i alla kameror, så länge en lika stor motivarea avbildas i

samtliga fall*. Exempel på MTF-kurvor med ortsfrekvensenheten line widths per picture height, samt en förklaring till enheten ges i Fig. 17.26.

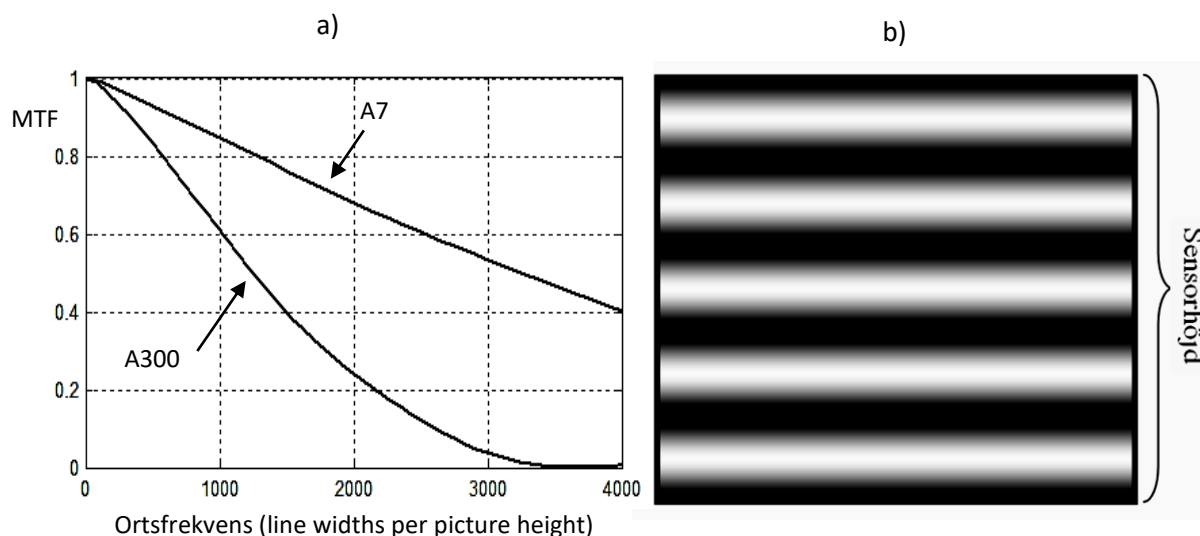


Fig. 17.26. a) MTF-kurvor för två digitalkameror, SONY A7 (24 Mpixlar, sensorstorlek 24 mm x 36 mm) och SONY A300 (10 Mpixlar, sensorstorlek 15.8 mm x 23.6 mm). Genom att ortsfrekvensskalan är graderad i enheten "line widths per picture height" kan dessa kurvor direkt jämföras utan att man behöver räkna om ortsfrekvensskalan. b) Illustration av begreppet "line widths per picture height". I detta fall uppgår denna storhet till 10, dvs 5 svarta och 5 vita linjer (= 5 perioder av mönstret). Detta sätt att ange ortsfrekvens är frikopplat både från sensorstorlek och antal pixlar.

Vad betyder då t.ex. 1000 line widths per picture height? Jo, det betyder helt enkelt att linjemönstret som avbildas på sensorn har sådan täthet att 1000 linjer (500 svarta och 500 vita) får plats längs sensorns kortsida. Eftersom man räknar både svarta och vita linjer, dvs halvperioder i mönstret, så svarar detta mot 500 perioder längs sensorns kortsida (och naturligtvis lika många längs den digitala bildens kortsida). Genom att ange ortsfrekvensen på detta sätt, uppnås alltså den stora fördelen av att direkt kunna jämföra MTF-kurvor oberoende av sensorstorlek och antal pixlar.

Alla de beskrivna metoderna för att ange ortsfrekvensen används i den tekniska litteraturen. Det är därför viktigt att kunna räkna om ortsfrekvenser som anges i mm^{-1} , perioder/pixel och perioder/sensorhöjd (eller line widths per picture height) på ett sätt som gör att man rättvist kan bedöma MTF-kurvor. Såsom nämnts, bör man också vara observant på om en MTF-kurva är influerad av sharpening i kameran (eller i efterföljande led). Om så är fallet är den i stort sett värdelös. Tecken på att sharpening har förekommit är om man ser en tendens till puckel i MTF-kurvan för relativt låga ortsfrekvenser.

* Sorten perioder/sensorhöjd är ekvivalent med sorten perioder/motivhöjd, förutsatt att man med "motivhöjd" avser den sträcka av motivet som avbildas längs sensorns höjd. Mera om detta i nästa delkapitel.

17.10. Sensorstorlekens inverkan på kamerans MTF

Som vi såg i slutet av kap. 17.9 så kan man enkelt jämföra kameror med olika sensorstorlek och antal pixlar, om man graderar MTF-kurvans orsfrekvensaxel i perioder/sensorhöjd. Den kamera som har högst MTF kommer då att ge bästa skärpa och detaljrikedom i bilden om man fotograferar lika stor motivarea med alla kamerorna. Som framgår av Fig. 17.26 a) så är kameran A7 klart överlägsen A300. Detta beror till stor del på den större sensorn i A7, och man kan fråga sig om vi kommer att få en ännu bättre MTF om vi väljer en kamera med ännu större sensor. Finns det någon yttersta gräns för hur bra MTF kan bli?

Kameror med större sensor än fullformat (24 mm x 36 mm) är mycket dyra och sällsynta, men genom en kontakt med Kungl. Biblioteket i Stockholm erhöll författaren mätdata från en Sinar-kamera med sensorstorlek 36 mm x 48 mm. MTF-kurvan för denna kamera visas i Fig. 17.27 tillsammans med kurvor för A7 och A300. Man ser genast att kameran med störst sensor har bäst MTF, och den med minst sensor har sämst MTF.

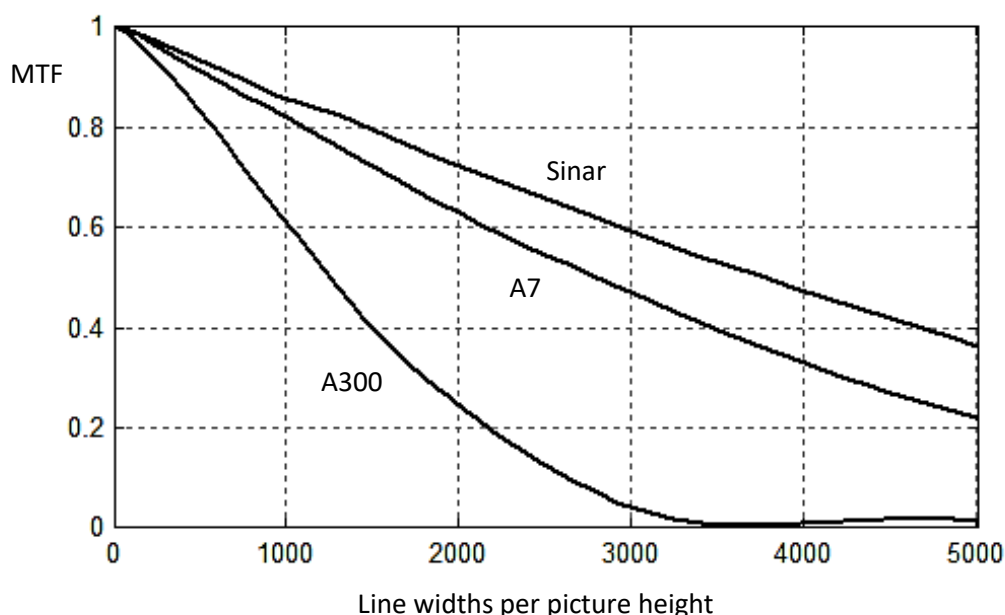


Fig. 17.27. MTF-kurvor för kameror med tre olika sensorstorlekar. Sinar (36 mm x 48 mm), Sony A7 (24 mm x 36 mm) och Sony A300 (15.8 mm x 23.6 mm). I samtliga fall var bländartalet 8 och brännvidden "normalbrännvidd" för respektive sensorstorlek.

Naturligtvis spelar även optikens kvalitet en viktig roll (den var nog lite sämre i fallet A300 ovan), men det visar sig generellt att kameror med större sensor så gott som alltid har en bättre MTF. Naturligtvis finns det praktiska (och ekonomiska) begränsningar för hur stor sensor en kamera kan ha, men man kan också undra om fysikens lagar någonstans sätter en gräns för hur detaljrika bilderna kan bli. Låt oss titta lite teoretiskt (och praktiskt) på detta.

Låt oss ta som utgångspunkt att vi vill fotografera ett motiv som innehåller små detaljer som vi vill avbilda med god kontrast. Samtidigt vill vi ha med en viss total

motivarea, och inte bara en närbild av en mindre del. Vi ska också anta att antalet pixlar är tillräckligt stort för att undvika störande pixelleringseffekter (dessa kommer att beröras i kap. 22), samt att sensorns MTF är så bra att den inte begränsar kamerans prestanda (MTF för sensorn kan teoretiskt göras godtyckligt hög, t.ex. genom ett extra avbildningssteg med förstoring). Kort sagt, det är kameraobjektivet som utgör den yttersta gränsen för vad som är teoretiskt möjligt att åstadkomma vad gäller MTF.

Den yttersta gränsen för optiken sätts av diffraktionen i bländaröppningen. Ett objektiv vars prestanda endast begränsas av diffraktion kallas diffraktionsbegränsat, och man har då antagit att avbildningsfelen (kap. 15) är försumbara. Bra objektiv är ofta nära diffraktionsbegränsade för bländartal F från omkring 5.6 och uppåt. Låt oss anta att vi har ett objektiv som används vid så högt bländartal att vi helt kan bortse från avbildningsfel (välj högre F än 5.6 om nödvändigt, detta är ingen begränsning i resonemanget). OK, låt oss börja undersökningen utgående från dessa förutsättningar. Den som vill skipa detaljerna kan hoppa fram några sidor där en sammanfattning ges.

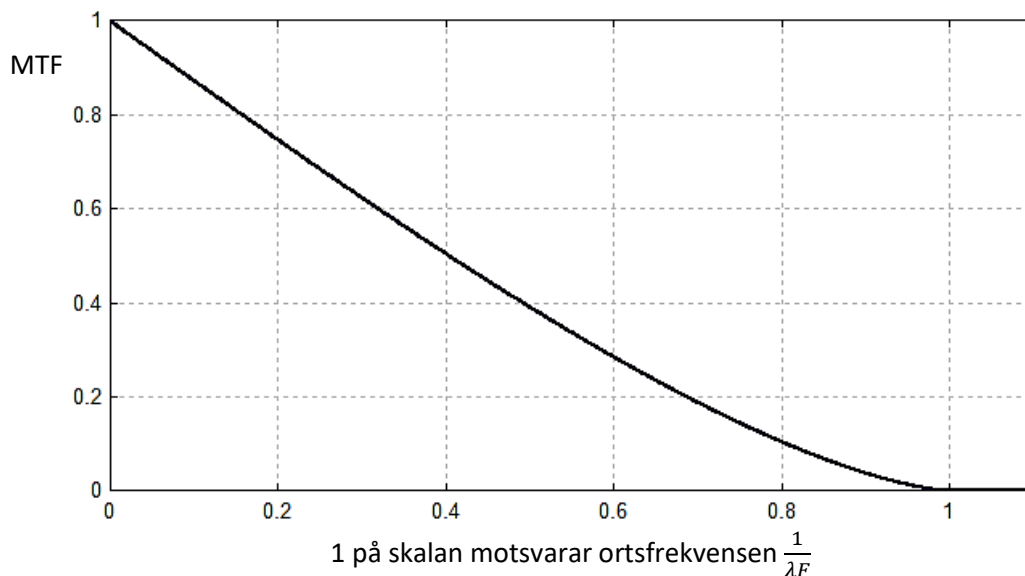


Fig. 17.28. MTF för ett diffraktionsbegränsat objektiv. λ är ljusvåglängden och F är bländartalet.

MTF för ett diffraktionsbegränsat objektiv ges av ekv. 17.3 och finns utritad i Fig. 17.28. Kurvans form är densamma för alla diffraktionsbegränsade objektiv, det enda som skiljer sig åt är den så kallade gränsfrekvensen, $\frac{1}{\lambda F}$ (enhet: perioder/längdenhet), där kurvan går ner till noll. Gränsfrekvensen kan därför tas som ett mått på hur bra MTF är, högre gränsfrekvens innebär att MTF-värdena ligger högre vid alla orsfrekvenser. För att jämföra kameror med olika sensorstorlekar och pixelantal, så kan man som vi sett använda orsfrekvensenheten perioder/sensorhöjd eller line widths/picture height (den sistnämnda ger ett dubbelt så högt tal, för övrigt är de ekvivalenta). Låt oss kalla sensorns utsträckning i höjdlängd för h (t.ex. 24 mm för en fullformatsensor). Vi får då att gränsfrekvensen för MTF blir $\frac{h}{\lambda F}$ med enheten "perioder per sensorhöjd". λ är givet om vi jobbar

med synligt ljus, och F hade vi ställt in så att objektivet är diffraktionsbegränsat. λF är alltså en konstant. Vi ser då att gränshöjden växer linjärt med sensorhöjden h . Sens moral: Ju större sensor, desto bättre.

Men om vi bara ökar sensorstorleken, och låter allt annat vara oförändrat, så får vi ju med mera av motivet på bilden. Om vi ska få med samma motivarea som tidigare, måste vi antingen förlänga brännvidden eller förflytta oss närmare motivet. Vilket av dessa alternativ man väljer beror mycket på omständigheterna. Är det lätt att flytta sig relativt motivet (svårt vid astrofotografering, t.ex.), och vilka brännvidder finns tillgängliga? Men vilket alternativ vi än väljer, så gäller att gränshöjden ges av $\frac{h}{\lambda F}$ så länge motivavståndet är mycket större än brännvidden (dvs vi fotograferar på "stort" avstånd, typ landskapsbilder eller i extremfallet astrofotografering). Så länge detta gäller, så ökar gränshöjden linjärt med sensorstorleken, och bilderna blir hela tiden detaljrikare och bättre. Gränsen sätts då bara av hur stor kamera som det är praktiskt (och ekonomiskt) möjligt att arbeta med. Vad gäller objektivet måste man emellertid se upp med att det verkligen är konstruerat för att ge god kvalitet inom hela den area som sensorn täcker. Det finns t.ex. objektiv som är konstruerade för att användas med APS-sensor (ca. 16 mm x 24 mm). Om ett sådant objektiv används tillsammans med en fullformatsensor (24 mm x 36 mm) kommer bilden att bli suddig utåt kanterna. Ännu större krav på objektivets bildfält ställs naturligtvis för sensorer som är större än fullformat.

I fallet att motivavståndet inte är mycket större än brännvidden, så gäller lite andra förhållanden. (Och där hamnar vi ju slutligen även i ovanstående fall om vi bara ökar brännvidden eller går närmare motivet, i varje fall om vi betraktar jordiska tillämpningar.) Bländartalet, $F = \frac{f}{D}$, måste då ersättas med $\frac{b}{D}$, där b är bildavståndet (Fig. 5.1). Eftersom b ökar ju närmare motivet vi kommer, så kommer detta att sänka gränshöjden jämfört med situationen vid stort motivavstånd. Vi kan nu skriva gränshöjden som $\frac{h}{\lambda b} = \frac{hD}{\lambda b} = \frac{hf}{\lambda Fb}$. Vi kan utnyttja att $\frac{f}{b} = \frac{1}{1+\frac{h}{A}}$, vilket gör att gränshöjden kan skrivas som $\frac{1}{\lambda F(\frac{1}{h}+\frac{1}{A})}$, där h = sensorhöjd och A = motivhöjd (dvs A är den sträcka av motivet som avbildas längs sensorns höjd h). Detta är ett generellt uttryck som gäller för alla motivavstånd. Det tidigare behandlade fallet "stort motivavstånd" betyder att $A \gg h$, vilket innebär att vårt generella uttryck ger en gränshöjd nära $\frac{h}{\lambda F}$, vilket (som sig bör) överensstämmer med vårt tidigare resultat för stora motivavstånd. Men om vi håller på med makrofotografering så måste vi använda det generella uttrycket, vilket innebär att gränshöjden beror både på sensorns storlek och motivets storlek. I Fig. 17.29 visas hur gränshöjden beror på kvoten $\frac{h}{A}$, dvs förhållandet mellan sensorns höjd och motivets höjd. För låga värden på $\frac{h}{A}$, vilket svarar mot stora motivavstånd, så ökar gränshöjden ungefär linjärt med sensorhöjden precis som vi tidigare sett. Men för $\frac{h}{A}$ -värden större än ungefär 1 uppträder en kraftig icke-linearitet, och kurvan planar ut asymptotiskt mot värdet $\frac{A}{\lambda F}$ (sort: perioder/sensorhöjd = perioder/motivhöjd). Detta är alltså den absolut högsta gränshöjden vi kan få, oberoende av hur stor sensor vi har. Högsta ortsfrekvensen i motivet som vi

teoretiskt kan avbilda blir då $\frac{1}{\lambda F}$ (perioder/längdenhet). Detta innebär alltså att fysiken sätter en gräns för hur små detaljer i motivet vi kan avbilda oavsett hur stor sensor vi använder. För $\lambda = 550 \text{ nm}$, och $F = 5.6$ blir den teoretiskt högsta frekvensen i motivet som kan avbildas 325 perioder/mm (periodlängd ca. $3 \text{ }\mu\text{m}$). Men detta förutsätter en oändligt stor sensor. Mer realistiskt vore en sensor lika stor som motivet, vilket skulle ge hälften så hög frekvens, 160 mm^{-1} , respektive $6 \text{ }\mu\text{m}$ periodlängd.

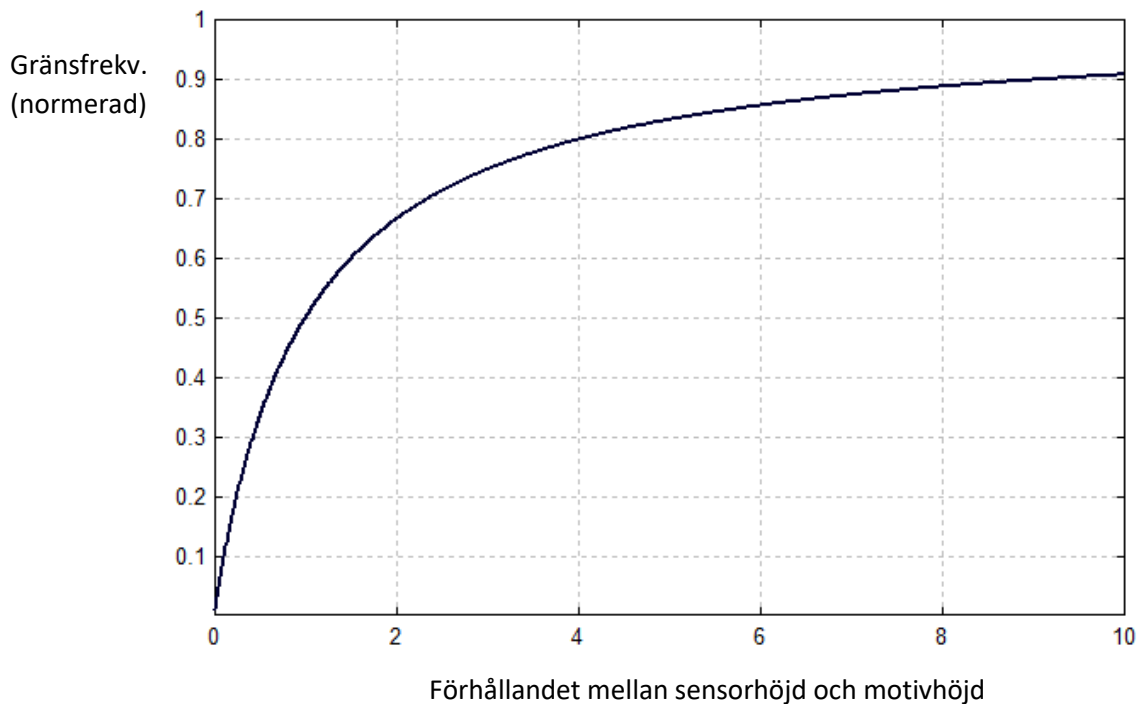


Fig. 17.29. Gränsfrekvensen för MTF i enheter av $\frac{A}{\lambda F}$ som funktion av förhållandet mellan sensorhöjd och motivhöjd, $\frac{h}{A}$.

Slutsatsen blir alltså att vid liten avbildningsskala (stort motivavstånd) lönar det sig bra att öka sensorstorleken. Bilderna blir då allt detaljrikare. Men när man ökar avbildningsskalan till omkring 1 och större, lönar det sig allt sämre att satsa på en större sensor. Låt oss ta ett exempel. Säg att vi vill avbilda ett motiv med höjden 10 mm. Använder vi då en kamera med APS-sensor (höjd 15.6 mm) blir kvoten $\frac{h}{A}$ lika med 1.56. Om vi istället använder en fullformatsensor (höjd 24 mm) blir $\frac{h}{A}$ lika med 2.4. Enligt kurvan i Fig. 17.29 medför denna ökning från 1.56 till 2.4 en ökning i gränsfrekvens med ca. 16%, medan sensorytan ökat med över 130%. Ganska dålig utdelning alltså. För mindre motiv än 10 mm lönar det sig ännu sämre att öka sensorstorleken.

Eftersom gränsfrekvensen är omvänt proportionell mot λF , innebär det att vi ska använda kortast möjliga våglängd och lägsta möjliga bländartal för att se så små detaljer som möjligt. Om vi jobbar med synligt ljus är det inte så mycket vi kan göra åt våglängden, och minskar vi F väsentligt under 5.6 kommer aberrationer vanligen att begränsa skärpan. Som vi sett ovan innebär det att vi inte kan urskilja

motivdetaljer mindre än några mikrometer. Vill vi se detaljer mindre än så, är alltså inte fotografering med vanliga kameror och objektiv lämpliga. Då bör vi istället använda mikroskop. I mikroskop nöjer man sig med att avbilda en mycket liten motivarea. Detta medför att motivhöjden A är liten, vilket betyder att man kan använda höga $\frac{h}{A}$ -värden utan att använda en särskilt stor sensor. På så sätt hamnar man långt till höger på kurvan i Fig. 17.29, och kommer på så sätt närmare den teoretiskt maximala gränsen. Men viktigare än detta är att mikroskopobjektiv är konstruerade på lite annat sätt än vanliga kameraobjektiv. Mikroskopobjektiv är optimerade för en enda avbildningsskala, vanligen mellan 10 och 100 (objektivförstoringen). Dessutom utnyttjas en mycket liten bildvinkel. Dessa begränsningar gör att man kan få nära diffraktionsbegränsade prestanda vid mycket lägre bländartal än 5.6. Faktum är att mikroskopobjektiv arbetar vid så låga bländartal att de ekvationer som vi hittills använt inte gäller längre. Formeln för teoretiskt högsta gränsfrekvensen, $\frac{A}{\lambda_F}$ måste ersättas med $\frac{2A \sin \alpha}{\lambda}$, där α är definierad i Fig. 9.13 ($\frac{A}{\lambda_F}$ är en approximation som gäller för ”små” α -värden). $\sin \alpha$ kallas objektivets numeriska apertur, NA , och finns angivet på alla mikroskopobjektiv (för immersionsobjektiv är $NA = n \sin \alpha$, där n är immersionsmediets, t.ex. olja, brytningsindex. Den högsta ortsfrekvensen i motivet (preparatet) som kan avbildas ligger för ljusmikroskopi runt $5 \mu\text{m}^{-1}$ (periodlängd $0.2 \mu\text{m}$), vilket är typiskt för ett objektiv med $NA \approx 1.4$ (oljeimmersion) och förstoring 100 gånger. Mikroskopobjektiv är som nämnts optimerade för ett ganska litet bildfält (klarar alltså bara små värden på A), mindre ju högre NA . Vanligen är produkten $A \times NA$ runt 0.6 mm eller lägre, och förstoringen är sådan att maximala motivhöjden A ger en bildhöjd av ca. 20 mm. Detta gör att det i praktiken knappast lönar sig att använda en större sensor än ca 20 mm x 20 mm, och med ca. 20 megapixlar. Utnyttjar man ett större bildfält än så får man dålig bildkvalité i kanterna. För synligt ljus (säg $\lambda \approx 550 \text{ nm}$) och $A \times NA$ runt 0.6 mm så hamnar gränsfrekvensen på lite drygt 2000 perioder per sensorhöjd vid fotografering i mikroskop.

Sammanfattning av resultaten:

- Fotografering på ”stora” avstånd. Fysiken sätter ingen gräns. Gränsen sätts av utrustningens storlek och pris. Så stor sensor som möjligt är bäst. Ett exempel på detta är astronomiska teleskop, som inte är annat än gigantiska digitalkameror.
- Mikrofotografering, typ mikroskop. Fysiken sätter gränser för hur små detaljer i motivet vi kan avbilda. Högsta ortsfrekvensen som kan avbildas i preparatet ges av $\frac{2NA}{\lambda}$, där NA = objektivets numeriska apertur. På grund av det begränsade bildfältet hos mikroskopobjektiv lönar det sig i praktiken inte att använda större sensorer än ca. 20 mm x 20 mm.
- Makrofotografering, typ avbildningsskala 1:1. Precis som i fallet med mikroskopi, sätter fysiken en gräns för hur små detaljer i motivet som kan avbildas (och den är betydligt lägre än i mikroskopi). Å andra sidan avbildas i allmänhet en mycket större motivyta. Detta gör att det ofta lönar sig med en större sensor än i mikroskopifallet, men över en viss sensorstorlek blir

förbättringen marginell. För detaljer får man gå till den mer detaljerade redogörelsen ovan.

Avslutningsvis ska vi titta på några MTF-kurvor för den vanligaste situationen, nämligen fotografering på "stora" avstånd (första fallet ovan). Vi antar som tidigare att objektivet är diffraktionsbegränsat, och vi antar att bländartalet är 5.6. Detta val av bländartal är lite godtyckligt, men bra objektiv är ofta nära diffraktionsbegränsade vid $F = 5.6$ och detta är ett bländartal som ofta används i praktiken. Som våglängd sätter vi 550 nm, som är ett medelvärde för synliga våglängder.

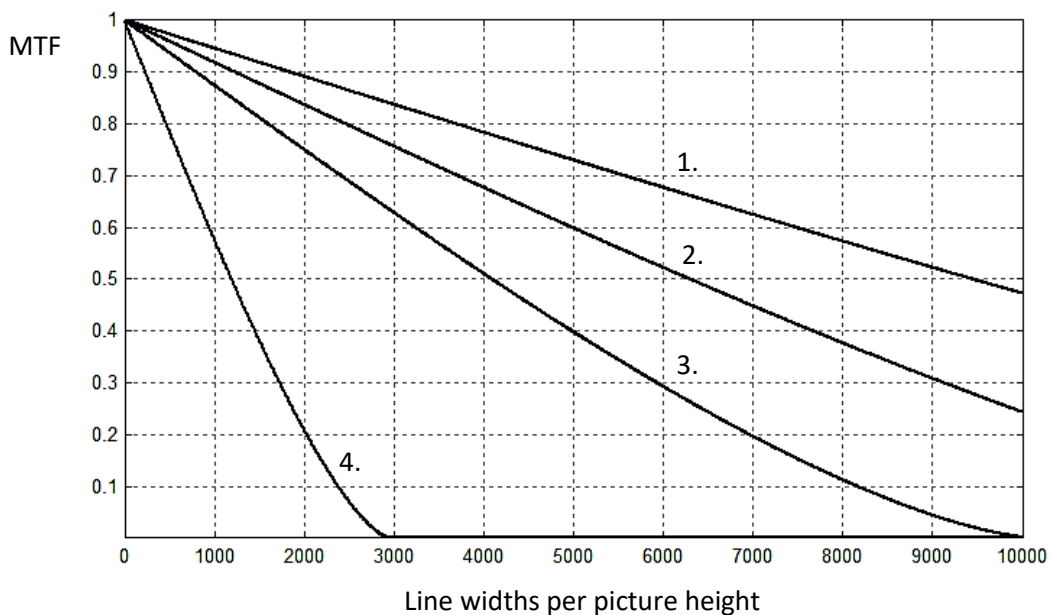


Fig. 17.30. MTF för kamera med diffraktionsbegränsat objektiv vid bländartal 5.6, och för olika sensorstorlekar (sensorn antas oändligt bra). Sensorstorlekarna i mm är: 1) 36 x 48, 2) 24 x 36, 3) 15.6 x 23.5, 4) 4.55 x 6.17.

Sensorstorlekarna 2, 3 och 4 i Fig 17.30 är ofta förekommande i praktiken, och benämns "fullformat", "APS" och " $1/2.3$ tum". De första två formaten förekommer i lite mer avancerade kameror, medan det sistnämnda är mycket vanligt i små kompaktkameror. Sensorstorlek 1 används bara i mycket dyra proffskameror. Bilder tagna i mikroskop kommer att ge en MTF som hamnar i närheten av kurva 4 i Fig. 17.30 (vi såg att förväntad gränsfrekvens var i bästa fall ca. 2000 perioder/sensorhöjd, dvs ca. 4000 line widths per picture height). Detta betyder inte att mikroskop har dåliga prestanda, utan det innebär att man har optimerat för att kunna se så små detaljer som möjligt i ett preparat (på bekostnad av bildfältet). Mikroskopfotografier, och fotografier tagna med kompaktkameror, innehåller alltså mindre total information i varje enskild bild jämfört med t.ex. Sinar- och A7-kamerorna. Med de sistnämnda kamerorna kan man göra väldigt stora utskrifter som både kan betraktas på stort avstånd för att ge en överblick, och som också tål att betraktas på kort avstånd för att se små detaljer. Så stora utskrifter lönar sig alltså inte att göra för mikroskopi- eller kompaktkamerabilder, eftersom dessa innehåller betydligt mindre total information.

Kurvorna i Fig. 17.30, liksom det övriga resonemanget i detta delkapitel, förutsätter att sensors MTF är oändligt bra, vilket naturligtvis inte är fallet i praktiken. Som redan nämnts så är det inget som rent teoretiskt hindrar att sensors MTF blir hur bra som helst, medan en sådan gräns finns för optiken. Men man kan ändå undra över hur bra MTF kan vara för en realistisk sensor, och hur mycket den skulle påverka kamerans totala MTF. Gränsen för sensors MTF sätts av den geometriska storleken på en individuell pixel – ju mindre pixelstorlek, desto bättre MTF. Å andra sidan kommer en liten pixel att samla in väldigt få fotoner, vilket innebär sämre fotometrisk kvalitet (signal/brus och dynamik, se kap. 14.5). Pixelstorlekar neråt 1 μm är vanligt förekommande i kompaktkameror, så vi kan ta det som exempel på en liten, men realiserbar, pixelstorlek. I kap. 17.6 beskrivs MTF för sensorer med olika pixelbredder, och där kan vi se att en sensor med bredden 1 μm skulle ha en MTF enligt Fig. 17.14, men med ett nollställe vid ortsfrekvensen 1000 mm^{-1} . Som vi sett har diffraktionsbegränsad optik en gränshfrekvens av $\frac{1}{\lambda F}$ (Fig. 17.28), vilken för $\lambda = 550 \text{ nm}$ och $F = 5.6$ blir lite drygt 300 mm^{-1} . Sensors MTF är alltså mycket bättre än optikens i detta fall. I verkligheten är sensors MTF lite sämre än vad som visas i Fig. 17.14, men det är ändå realistiskt att utgå från att sensorer kan göras MTF-mässigt klart bättre än optiken. Men att ha en pixelstorlek på ca. 1 μm i t.ex. en fullformatsensor (24 mm x 36 mm) skulle bli problematiskt på grund av det stora antalet pixlar (många hundra miljoner) i de digitala bilderna. Dessutom kommer, som redan nämnts, den fotometriska kvalitén att bli lidande av så små pixlar. Men även med betydligt större pixelstorlekar än 1 μm , så kvarstår det faktum att det i nästan alla fotografiska situationer är en kvalitetsmässig fördel att använda en kamera utrustad med så stor sensor som möjligt. Det krävs dock ett stort antal pixlar i en stor sensor, både för att kunna återge den information som optiken ger, och för att undvika störande pixelleringseffekter typ moiré (kap. 22).

17.11. Contrast Transfer Function (CTF)

Ibland publiceras kurvor som liknar MTF-kurvor, men som betecknas med CTF (Contrast Transfer Function). Precis som för MTF, så visar dessa kurvor hur mycket modulationsgraden för linjemönster minskar när man går från motiv till bild (se kap. 17.3). Skillnaden jämfört med MTF-kurvor, är att CTF-värdena gäller för linjemönster med fyrkantprofil (homogent svarta och homogent vita linjer) och inte för mönster med sinusprofil (vilket är fallet för MTF, se Fig. 17.6)*.

Anledningen till att man införde CTF var att det är mycket lättare att framställa testmönster i form av linjer med fyrkantprofil än med sinusprofil (detta utnyttjas t.ex. i laboration 2 i kursen SK1140 på KTH). I och med att MTF-mätningar för digitalkameror numera oftast utförs genom avbildning av en svartvit kant (som är lätt att framställa), se kap. 17.9, har mycket av den ursprungliga motiveringen för att införa CTF försvunnit. Det finns dock fortfarande fall där CTF kan vara användbart av olika skäl, t.ex. i tillämpningar där motivet utgörs av linjemönster med fyrkantprofil (avsyning av kretskort mm). Det är därför inte helt ovanligt att CTF-kurvor förekommer i datablad för t.ex. optik. För att kunna tolka CTF-kurvor kan det vara bra att veta lite om hur CTF förhåller sig till MTF.

* Tyvärr råder det ibland en viss oklarhet angående vad som avses med CTF. Speciellt inom elektronmikroskopi-kretsar använder man ofta beteckningen CTF när man i själva verket talar om MTF.

Ett linjemönster med fyrkantprofil kan med hjälp av Fourieranalys beskrivas som en oändlig summa av harmoniska funktioner (sinus eller cosinus) med olika frekvenser (grundton plus övertoner), Fig. 17.31. Vid avbildningen minskar kontrasten för var och en av dessa harmoniska svängningar med en faktor som ges av MTF-värdet vid aktuell frekvens. Om man adderar dessa kontrastminskade harmoniska svängningar, får man bilden av det fyrkantmodulerade linjemönstret. Ett exempel på detta visas i Fig. 17.32. Modulationsgraden i bilden dividerad med modulationsgraden i motivet är CTF-värdet. Vid CTF-beräkningar används mätvärden tagna mitt på de positiva och negativa halvperioderna av kurvprofilen i bilden (motsvarande x-värden 0.25 och 0.75 i Fig. 17.32).

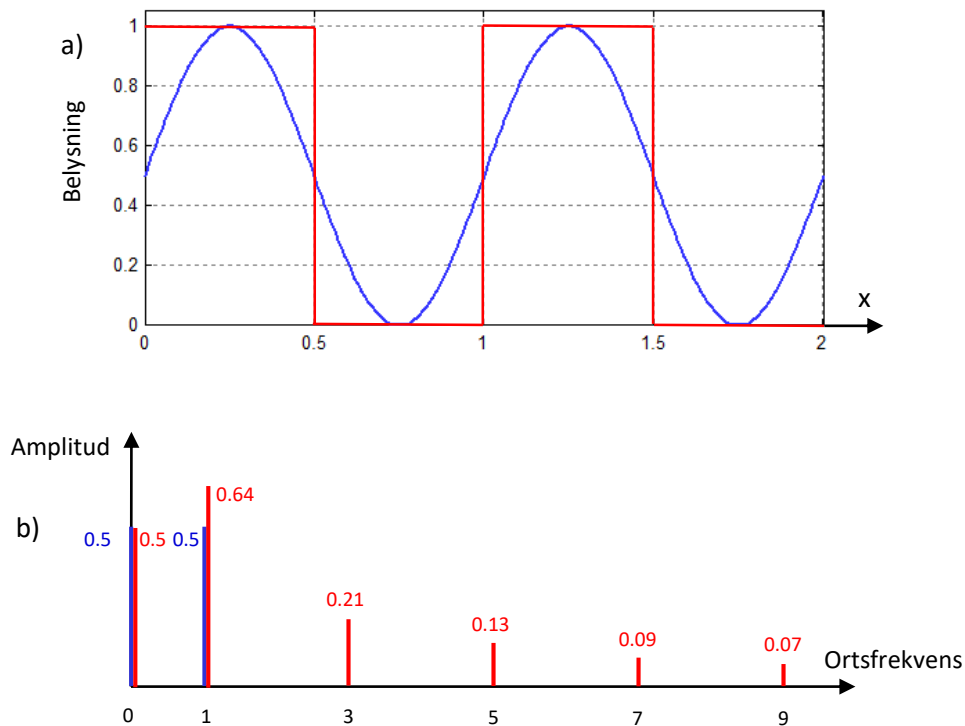


Fig. 17.31. a) Fyrkantprofil och sinusprofil med samma medelvärden, amplituder och ortsfrekvenser (frekvensen har satts till 1). b) Frekvensinnehåll i de bägge funktionerna. I sinusfallet är medelvärdet 0.5 och amplituden 0.5 vid frekvens 1 (blå siffror och staplar). I fyrkantfallet (röda siffror och staplar) är medelvärdet också 0.5, men amplituden för frekvensen 1 (grundton) är $\frac{2}{\pi}$ och dessutom har ett oändligt antal udda övertoner dykt upp med amplituder $\frac{2}{n\pi}$, där n är ett udda heltal (i figuren visas bara n upp till 9).

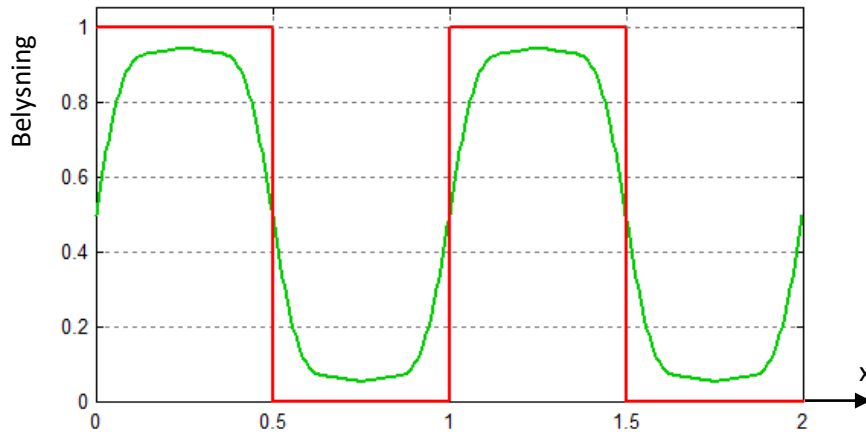


Fig. 17.32. Exempel på hur belysningsprofilen hos ett fyrkantmodulerat linjemönster (röd kurva) förändras vid avbildning i ett diffraktionsbegränsat objektiv (grön kurva). I exemplet har antagits att objektivets gränshfrekvens är 7 i den normerade skala som ges i Fig. 17.31 b (dvs. bara grundtonen och de två första övertonerna avbildas). Eftersom objektivet dämpar övertonerna mer än grundtonen, kommer skarpa svart/vit-övergångar att smetas ut så att belysningsprofilen får en mera avrundad form.

Sambandet mellan CTF och MTF kan beskrivas genom följande formel:

$$CTF(v) = \frac{4}{\pi} \left[MTF(v) - \frac{MTF(3v)}{3} + \frac{MTF(5v)}{5} - \frac{MTF(7v)}{7} + \dots \right] \quad (17.4)$$

där v är ortsfrekvensen (perioder per längdenhet) för linjemönstret med fyrkantprofil. Precis som för MTF, gäller att CTF-kurvan börjar på höjden ett vid $v = 0$ och sedan sjunker med ökande ortsfrekvens.

Som ett praktiskt exempel kan vi jämföra CTF och MTF för ett diffraktionsbegränsat objektiv, vars MTF ges av ekv. 17.3. Vi ser där att gränshfrekvensen där MTF går ner till noll (och sedan stannar på noll) är $v_{gr} = \frac{1}{\lambda F}$, där λ är ljusvåglängden och F är bländartalet. För $v \geq \frac{v_{gr}}{3}$ ger ekv. 17.4 att $CTF(v) = \frac{4MTF(v)}{\pi} = 1.27 \times MTF(v)$, dvs bara första termen i summan är skild från noll. För de övre två tredjedelarna av ortsfrekvensaxeln kommer alltså CTF-kurvan att ha precis samma form som MTF-kurvan, men den ligger på 27% högre nivå. Bilden av linjemönstret kommer där att få en ren sinusprofil, eftersom samtliga övertoner är bortfiltrerade. För $v < \frac{v_{gr}}{3}$ måste man ta med flera termer i summan i ekv. 17.4, och kurvformen för CTF kommer då att avvika från den för MTF. Resultatet visas i Fig. 17.33.

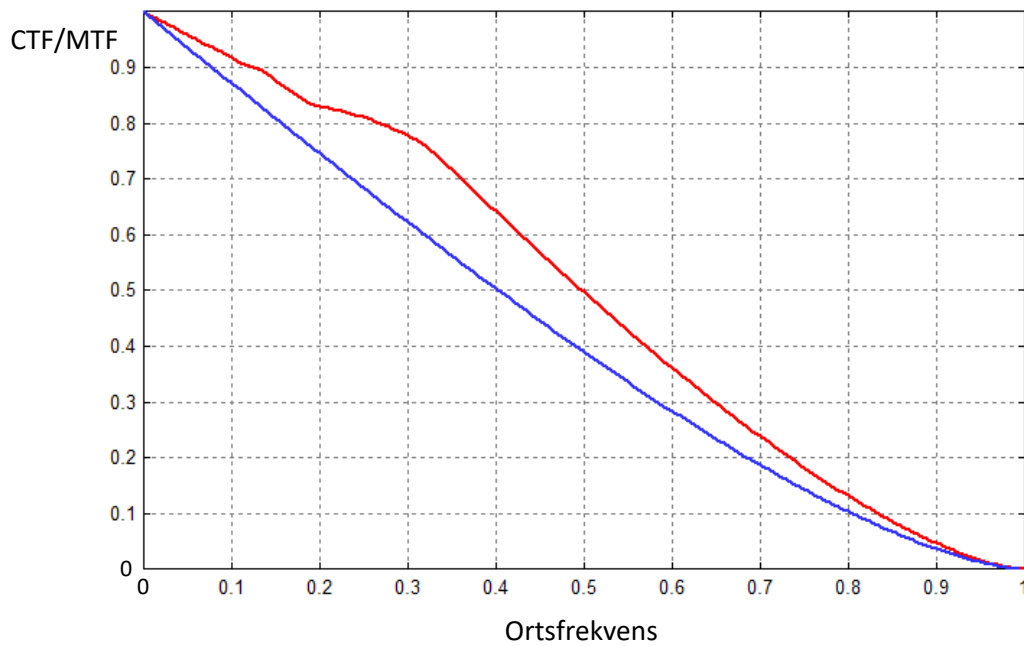


Fig. 17.33. CTF-kurvan för ett diffraktionsbegränsat objektiv (röd) jämförd med motsvarande MTF-kurva (blå). CTF-kurvan ligger generellt högre än MTF-kurvan. Detta innebär att linjemönster med fyrkantprofil återges med högre kontrast, och upplevs därför skarpare, än mönster med sinusprofil. Objektivet gränshfrekvens är normerad till 1 (värdena multipliceras med $\frac{1}{\lambda F}$ för att få verklig orsfrekvens).

CTF-kurvan i Fig. 17.33 uppvisar lite svängningar för frekvenser under $\frac{1}{3}$. Detta beror på att successivt allt fler termer i ekv. 17.4 är skilda från noll, och de har omväxlande positivt och negativt tecken. Svängningarna avtar dock när orsfrekvensen går mot noll eftersom högre ordningens termer i ekv. 17.4 är mycket små.

I Fig. 17.33 visas CTF-kurvan för ett diffraktionsbegränsat objektiv. Förutom objektiv, så innehåller en kamera naturligtvis en sensor vars prestanda också kan beskrivas med CTF. När vi betraktade MTF såg vi att kamerans totala MTF kunde erhållas som produkten av optikens och sensorns MTF (kap. 17.5). Denna multiplikationsregel gäller inte för CTF. Detta är en nackdel, eftersom det betyder att systemets totala CTF inte på ett enkelt sätt kan erhållas ur de enskilda komponenternas CTF.

17.12. Färginterpolationens inverkan på MTF

I en sensor med Bayer färgmosaik, Fig. 17.34, detekterar varje pixel endast en färgkomponent (rött, grönt eller blått). De övriga färgkomponenterna erhålls (gissas) genom interpolationsräkningar där man använder grannpixlarnas värden, se kap. 14.9. Det finns en mängd olika interpolationsalgoritmer av olika komplexitet. När man tar ut bilderna i RAW-format kan man göra interpolationen

på en yttre kraftfullare dator än kamerans inbyggda processor, och därigenom få ett bättre resultat.

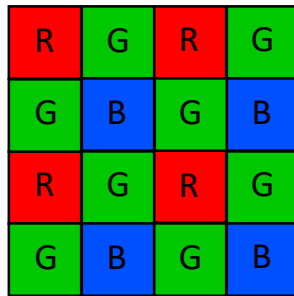


Fig. 17.34. Bayer färgfiltermask. I kap. 14.8 beskrivs filtermasker för sensorer mer utförligt.

Oavsett vilken metod som används, så innebär alltid färginterpolation att det finns risk för felaktiga färger på små bilddetaljer och skarpa kanter, Fig. 14.26. Dessutom kommer interpolationen att påverka skärpan i bilderna, och det är den aspekten vi ska titta lite på här.

Interpolationen utförs i två dimensioner, och ser olika ut för gröna respektive röda och blå element, Fig. 14.25. Denna skillnad beror på att 50% av pixlarna är gröna medan 25% är blå och 25% är röda. Bilinjär interpolation är en enkel metod där man helt enkelt använder medelvärdet av grannpixlarnas värden i en 3x3 omgivning. För att uppnå bättre resultat används ofta så kallad bikubisk interpolation eller ännu mer avancerade metoder. Detta har beskrivits i kap. 14.9. Vi ska här titta lite på hur användandet av bilinjär interpolation kan tänkas påverka sensorns MTF.

Som beskrivits i kap. 17.2, kan man erhålla MTF som beloppet av Fouriertransformen av punktspridningsfunktionen, *psf*. För att detta ska gälla måste *psf* vara skift-invariant, dvs den får inte variera från pixel till pixel i en bild. Tyvärr är detta just fallet när man utför färginterpolation. Som ett exempel för att illustrera problemet kan vi välja den gröna färgkomponenten i en bild, samt begränsa oss till att titta på linjemönster som bara varierar i x-riktning (dvs. vertikala streckmönster).

Vi utgår från en ideal sensor enligt Fig. 17.34 där 100% av sensorytan täcks av pixlar, och varje pixel har en jämn ljuskänslighet över hela sin yta. För en pixel täckt med grönsfilter blir *psf* enligt Fig. 17.35 *a* eftersom grönvärdet direkt registreras av pixeln och ingen interpolation behöver göras. För pixlar täckta med blå- eller rödfilter måste interpolation enligt Fig. 14.25 *b* göras för att få ett grönvärde. Då får vi en *psf* enligt Fig. 17.35 *b*. Genom att ta Fouriertransformen av dessa två *psf*:er får vi MTF-kurvor enligt Fig. 17.36. Det är uppenbart att interpolationen ger en mycket kraftig försämring av MTF redan vid ganska låga ortsfrekvenser. Nu kanske någon undrar hur man får fram total MTF för den gröna delbilden, dvs hur ska vi kombinera kurvorna i Fig. 17.36? Svaret är att det inte går. Eftersom skift-invarians inte råder gäller inte den matematiska teorin för MTF. Interpolationen kommer att introducera icke-lineariteter, som bland annat orsakar att det uppstår falska ortsfrekvenser som inte har någon motsvarighet i motivet.

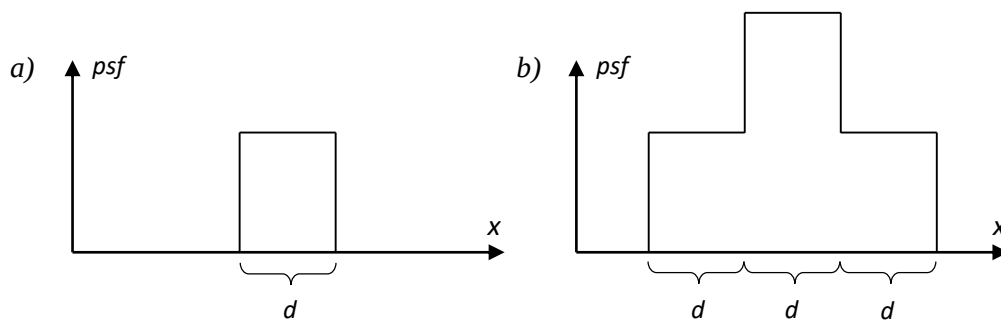


Fig. 17.35. Punktspridningsfunktionen, psf , för en grön pixels grönvärde a) och för blå och röda pixlars framinterpolerade grönvärde b) för sensorn i Fig. 17.34. Vi betraktar ett endimensionellt fall med ett linjemönster som bara varierar i horisontell riktning (x -riktning). d = pixelbredden.

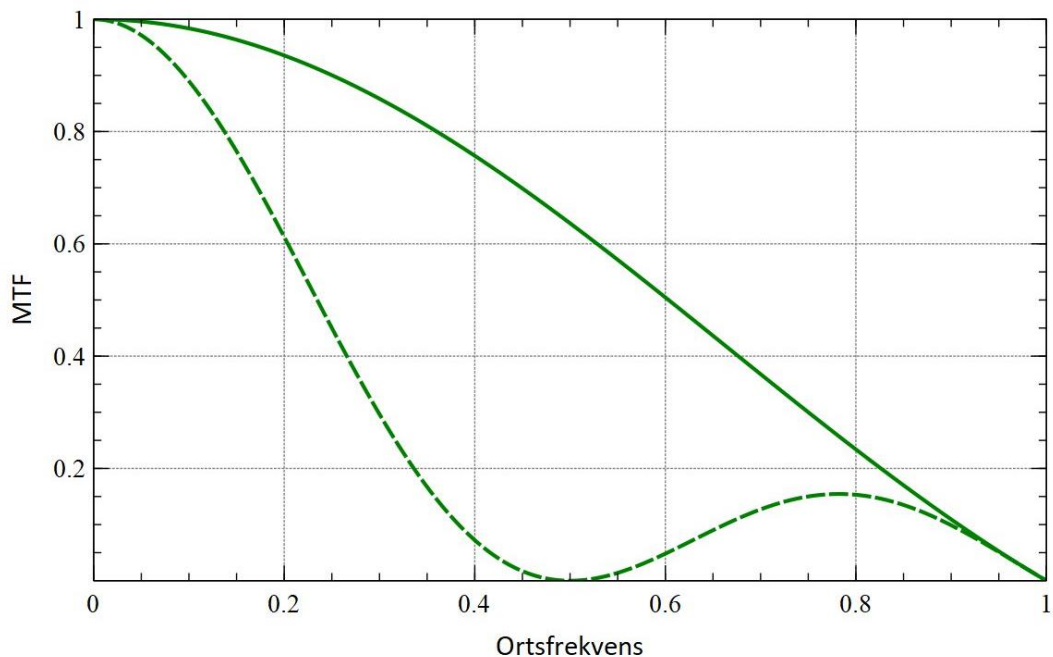


Fig. 17.36. Heldragna kurvan visar MTF för en grönpixels grönvärde (uppmätt), och den streckade kurvan MTF för det interpolerade grönvärdet för blå och röda pixlar. Ortsfrekvensskalan är normerad så att värdet 1 motsvarar frekvensen $\frac{1}{d}$, där d är pixelbredden.

Genom att titta på de MTF-kurvor som motsvarar de olika interpolationsfallen i Fig. 14.25 kan man se att de uppvisar ett likartat beteende. De faller kraftigt vid ökande Ortsfrekvens plus att de uppvisar ett eller flera sekundära maxima (som hamnar vid lite olika Ortsfrekvenser för de olika interpolationsfallen). Rent generellt kan vi därför säga att vi kan förvänta oss att färginterpolationen medför en klar skärpeförlust i bilderna, även om denna förlust inte kan beskrivas med en enda MTF-kurva. Forskning har utförts med målsättningen att trots de matematiska svårigheterna hitta någon sorts "ekvivalent MTF" för hela sensorn vid

färginterpolation*. Resultatet man kom fram till liknar ganska mycket en lite ”smoothad” version av medelvärdet för de olika interpolationsfallen, något i stil med det som visas i Fig. 17.37. Som synes kan man förvänta sig en ganska betydande försämring av bildskärpan på grund av färginterpoleringen. Det bör dock betonas att kurvorna som visas i Fig. 17.37 är mycket kraftiga approximationer som erhållits under förutsättning att pixlarna har helt jämn känslighet över ytan, att interpolationen är bilinjär (det är den inte i verkligheten) och att man gjort vissa matematiska förenklingar. Dessutom utnyttjas ofta korrelationer mellan färgkanalerna när man gör färginterpolation (se kap. 22.7.2) vilket ytterligare försvårar en beskrivning av avbildningsegenskaperna med hjälp av MTF. I praktiken är det därför knappast meningsfullt att använda annat än rena RAW-bilder (ingen färginterpolation eller annan bildbehandling gjord) vid framtagande av MTF (se även Fig. 17.25).

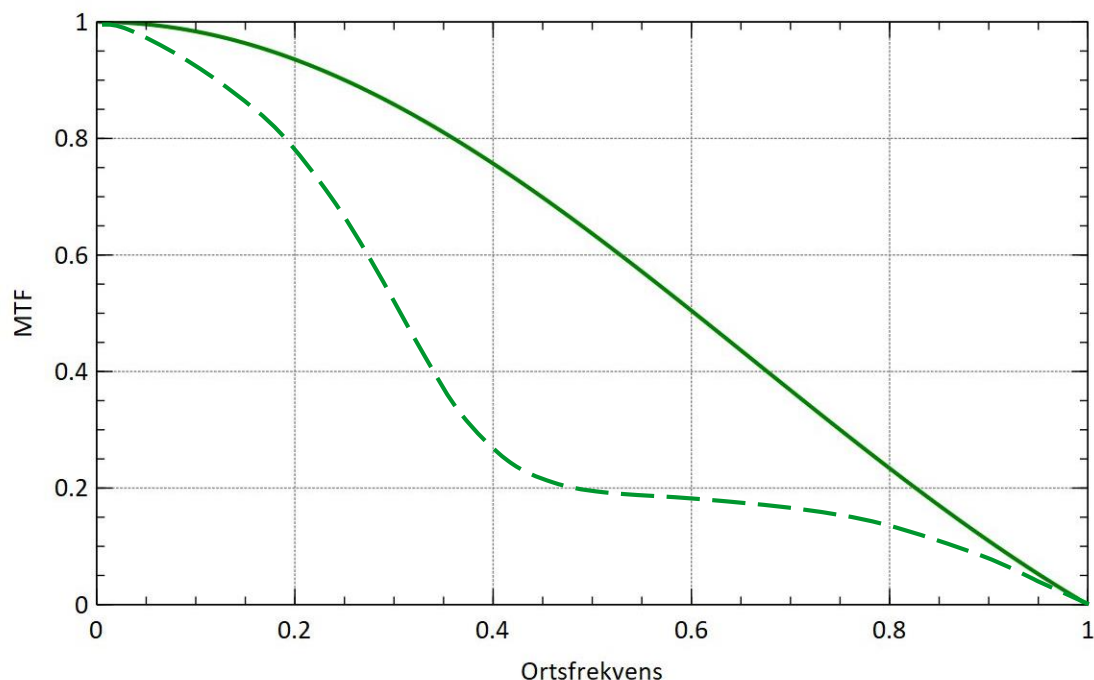


Fig. 17.37. Heldragna kurvan visar MTF för en ideal sensor utan färginterpolation. Streckade kurvan visar ungefär hur man kan förvänta sig att kontrasten i bilden minskar med ökande ortsfrekvens vid färginterpolation, se texten för detaljer. Ortsfrekvensskalan är normerad så att värdet 1 motsvarar frekvensen $\frac{1}{d}$, där d är pixelbredden.

17.13. Avfaltning

Om ett avbildande systems MTF är känd, så finns möjligheten att på matematisk väg kompensera för den skärpeförlust som det avbildande systemet orsakar. På detta sätt kan man ”skärpa upp” bilderna. Denna process kallas avfaltning (eng. deconvolution). Observera att avfaltning är något helt annat än de ”sharpening-

* Yotam, Ephri & Ami, MTF for Bayer pattern color detector, proc. SPIE vol. 6567 (2007).

rutiner” som utnyttjas i digitalkameror och diverse mjukvara för bildbehandling. Avfaltung bygger på en korrekt matematisk process för att återställa den kontrastförlust som skett i avbildningsprocessen, så att slutbilden så nära som möjligt återger det verkliga motivet. Sharpening å andra sidan innebär bara att man kosmetiskt framställer en bild som ser skarp ut för ögat, men som inte tar hänsyn till avbildningsprocessen eller det verkliga motivet.

Avfaltung bygger på att man med Fourieranalys kan få fram ortsfrekvensinnehållet i en bild. Om man känner MTF, så vet man också hur mycket de olika ortsfrekvenserna försvagats i avbildningsprocessen. Denna kunskap gör att man, åtminstone teoretiskt, kan förstärka upp de olika ortsfrekvenserna i bilden så att de motsvarar motivets ortsfrekvensinnehåll. Metoden illustreras i Fig. 17.38. Populärt kan man beskriva det som utförs så att man gör en kontrastförstärkning i bilden, men på så sätt att områden med små detaljer ökar mera i kontrast än bildpartier med grova detaljer (eftersom partier med små detaljer förlorar mest i kontrast vid bildregistreringen). Detta låter ju bra, eller hur? Men i verkligheten finns naturligtvis begränsningar för hur bra man kan återskapa motivets verkliga utseende.

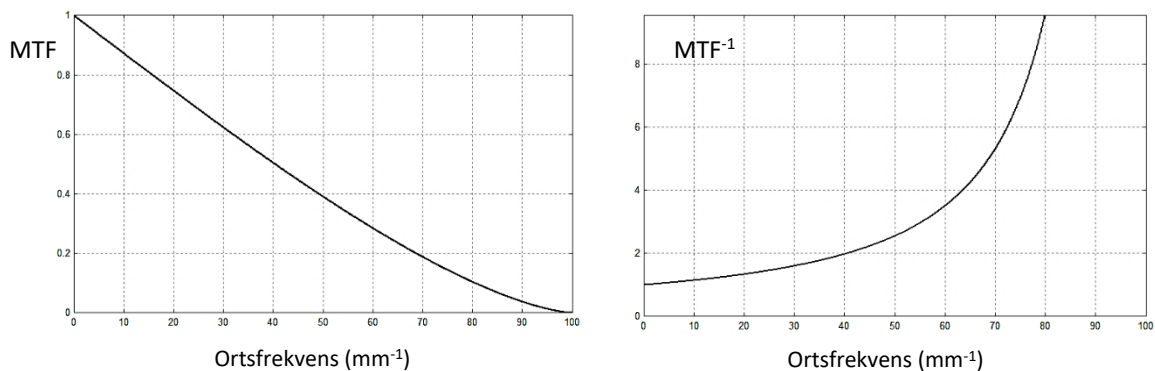


Fig. 17.38. Vänstra figuren visar ett exempel på en MTF-kurva för ett avbildande system. Högra kurvan visar inverterade MTF-värden som funktion av ortsfrekvensen. Om modulationsgraden för alla ortsfrekvenser i bilden multipliceras med dessa inverterade värden, kommer man att kompensera för den förlust i skärpa som orsakas av det avbildande systemet. I praktiken finns begränsningar som diskuteras i texten.

En begränsning är att man naturligtvis inte kan återskapa ortsfrekvenser som ligger över gränsfrekvensen för det avbildande systemet (100 mm^{-1} i Fig. 17.38). Information om dessa finns ju överhuvudtaget inte med i bilden. Och även frekvenser strax under gränsfrekvensen medför problem, eftersom modulationen för dessa frekvenser behöver förstärkas kraftigt. En sådan kraftig förstärkning kommer att medföra att även bruset som alltid finns i bilddata (kap. 14.5) också kommer att förstärkas upp kraftigt i det högre frekvensområdet. Vid avfaltung används därför ofta någon typ av brusreducerande åtgärder. Men detta är inte problemfritt, eftersom brusreduktion har en tendens att sudda till bilderna. Det gäller att hitta en lämplig balans.

Ett annat problem som kan dyka upp vid avfaltung är att antalet bitar i pixeldata (kap. 14.7) inte riktigt räcker till. Om modulationsgraden för en viss ortsfrekvens

är låg i bilden, så kommer pixelvärdena för ljusa och mörka linjer bara att skilja sig åt med några få nivåer. Efter en kraftig förstärkning så blir ljusa linjer ljusare och mörka linjer mörkare, men antalet digitala grånivåer är detsamma. Det kan medföra en "hackighet" i grånivåerna av den typ som visas i Fig. 14.18. Därför är det en fördel om bilder som ska avfaltas har pixelvärden med många bitars djup (12, 14 eller 16 istället för 8 som ofta är fallet).

Fig. 17.39 visar hur väl avfaltning kan fungera om MTF för avbildande systemet är välkänd, samt bilddata är av god kvalitet.



Fig. 17.39. Vänstra figuren visar en bild som erhållits då optiken varit felfokuserad. Bilden är allmänt suddig, och för vissa tätheter i "ekermönstret" försvinner kontrasten helt för att sedan återuppstå (men då har vita och svarta linjer bytt plats). Högra figuren visar samma bild efter avfaltning. Förutom att skärpan är mycket bättre, så ligger svarta och vita linjer på sin rätta plats. Vid linjetätheter där den vänstra figuren helt saknar kontrast kvarstår av naturliga skäl vissa felaktigheter. (Bildmaterialet har erhållits från Anna Burvall och Peter Unsbo.)

Ofta är det svårt att veta hur MTF för det avbildande systemet verkligen ser ut. En möjlighet att ändå avfalta bilder, är att utnyttja så kallad blind avfaltning (eng. blind deconvolution). Detta kan utföras genom t.ex. statistiska metoder för att uppskatta både avbildande systemets MTF, och det verkliga motivets utseende (till en viss gräns), ur originalbilden. Ibland används iterativa metoder vid dessa beräkningar. Problemet är matematiskt utmanande, och viss osäkerhet kan råda om hur verklighetstroget slutresultatet blir. Resultatet av blind avfaltning kan också påverkas av motivets utseende.

17.14. Sharpening

I stort sett alla bilder från digitalkameror genomgår någon form av ”sharpening” för att se ”skarpare” ut för betraktaren. På svenska används ofta beteckningen ”skärpa” för denna process, t.ex. i datorprogram för bildbehandling. En bättre beteckning vore ”uppskärpning”, men vi ska använda den engelska beteckningen ”sharpening” eftersom det är denna man oftast stöter på.

Sharpening kan utföras på lite olika sätt, men den vanligaste metoden är den som kallas oskarp mask (unsharp mask), och som vi ska titta lite närmare på i detta kapitel. Metoden går ut på att man först tillverkar en suddig kopia av bilden. Sedan subtraheras denna suddiga kopia, multiplicerad med en koefficient mindre än 1, från originalbilden. Koefficienten, samt graden av suddighet i kopian, kan styras med två parametrar (radie och vikt) som ska anges när sharpening utförs i bildbehandlingsprogram. Ibland ska också en tredje parameter, tröskelvärde, anges. Vi återkommer till dessa parametrar lite senare. Bilder som levereras i jpeg-format från en kamera är redan behandlade med sharpening av kamerans inbyggda processor. Fotografen kan vanligen ange med en enda parameter, t.ex. i en skala från -3 till +3 i en meny, hur kraftig sharpening som ska göras. Kameran väljer då automatiskt ”lämpliga” kombinationer av parametrar för att det inte ska bli för komplicerat för fotografen. Men om sharpening görs på en yttre dator har man större möjlighet att påverka resultatet, t.ex. genom att ta ut bilder ur kameran i RAW-format och behandla dessa med en RAW-converter.

Sharpening-processen kan beskrivas rent matematiskt på samma sätt som t.ex. optikens och sensors påverkan på bildens ordsfrekvensinnehåll. Motsvarigheten till kamerans punktspridningsfunktion, psf_{kamera} , se kap. 17.2, ser för sharpening-processen vanligen ut på följande sätt

$$psf_{sharp}(x, y) = \delta(x, y) - a \cdot \exp[b(x^2 + y^2)] \quad (17.5)$$

där δ står för deltafunktionen (kap. 17.2) och a och b är konstanter. I ekv. 17.5 representerar deltafunktionen originalbilden, och exponentialfunktionen den suddiga kopian av bilden (Gaussisk oskärpa). Genom att ta Fouriertransformen av psf_{kamera} fick vi kamerans MTF-funktion (kap. 17.2). På motsvarande sätt kan vi genom att ta Fouriertransformen av psf_{sharp} studera hur den kommer att påverka bildens ordsfrekvenser. Eftersom det nu handlar om rena datorberäkningar så blir detta inte något kvalitetsmått på kameran, och därför använder vi inte beteckningen MTF utan använder den mer neutrala beteckningen överföringsfunktion T (syftande på engelskans Transfer). Eftersom psf_{sharp} är rotationssymmetrisk i xy-planet så kommer även T att vara rotationssymmetrisk, dvs ordsfrekvenser påverkas likadant oberoende av mönsterorienteringen. Vi nöjer oss därför med att betrakta resultatet i en dimension. Vi får då följande utseende för T som funktion av ordsfrekvensen.

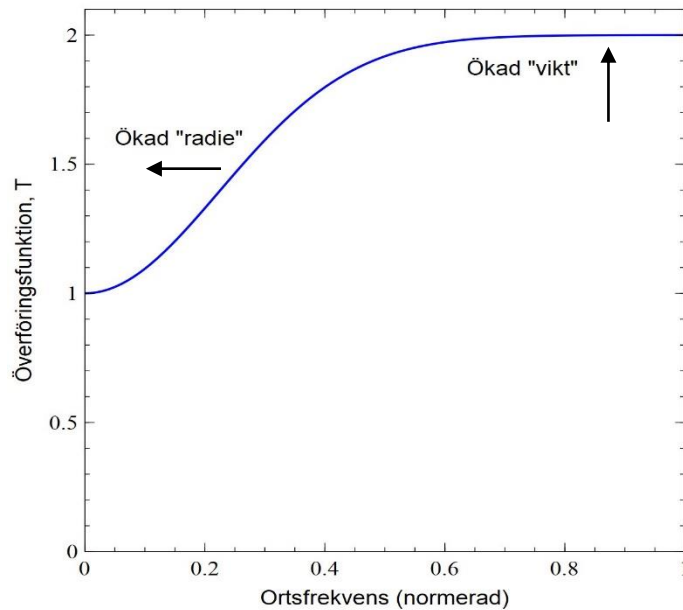


Fig. 17.40. Figuren visar hur sharpening-processen påverkar ortsfrekvenser i bilden. Låga ortsfrekvenser påverkas mycket lite, medan högre ortsfrekvenser förstärks. Graden av förstärkning, samt hur snabbt denna ska sätta in vid ökande ortsfrekvens, styrs genom parametervärderna "radie" och "vikt".

T -funktionen skiljer sig från MTF på en mycket väsentlig punkt, den antar värden större än 1. Detta beror på att psf_{sharp} antar negativa värden, vilket psf_{kamera} av naturliga skäl inte kan (ljusintensitet är alltid positiv). Genom att ändra parametervärdena a och b i ekv. 17.5 kan T -kurvans utseende ändras på så sätt som visas i Fig. 17.40. Parametern a påverkas genom valet av "vikt", och b påverkas av valet av "radie". Exakt hur denna påverkan görs kan skilja sig lite mellan olika program, men generellt gäller att ökad "vikt" förstärker höga ortsfrekvenser mera och större "radie" gör att denna ökning startar vid lägre ortsfrekvenser.

Genom att kombinera överföringsfunktionen T med kamerans MTF så erhålls den totala överföringsfunktionen från motiv till bild. Detta sker genom att multiplicera MTF_{kamera} med T . För att få en känsla för hur denna totala överföringsfunktion kan se ut så kan vi titta på fallet att MTF_{kamera} är ungefär lika med diffraktionsbegränsad MTF för objektivet (Fig. 17.8). I Fig. 17.41 visas denna totala överföringsfunktion för några olika värden på parametern "vikt" men där parametern "radie" är konstant. Man ser att sharpening ger en puckel i kurvan som blir mer uttalad ju högre vikt man använder. Om man minskar radien flyttas puckeln åt höger, och genom att anpassa både vikt och radie kan man få en kurva som blir mer "fyrkantig" och mera liknar "idealfallet" i Fig. 17.8. Exempel på detta visas i Fig. 17.42. Den röda kurvan innebär en kraftig kontrastökning för alla ortsfrekvenser utan att överförstärka något område så att kurvan blir högre än 1. Ganska optimalt kan det tyckas, men är det verkligen optimalt med en fyrkantig överföringsfunktion? Låt oss undersöka!

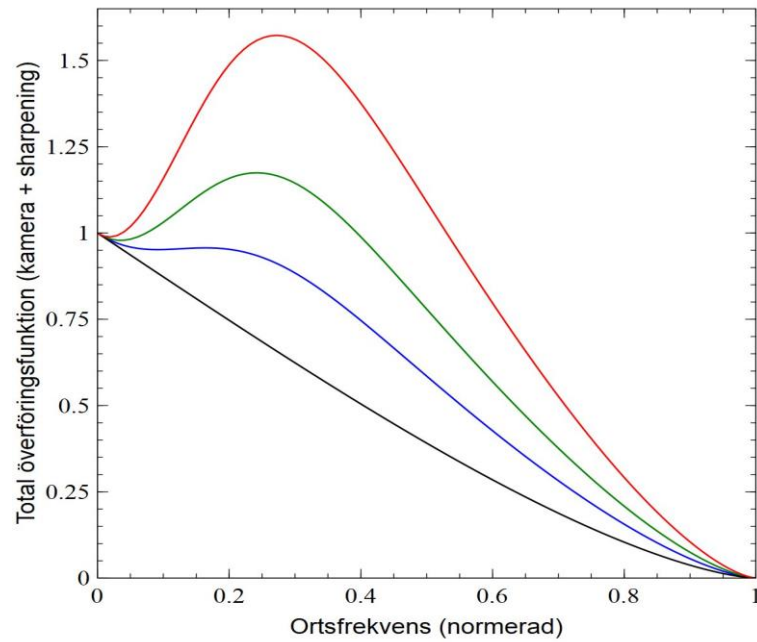


Fig. 17.41. Total överföringsfunktion när sharpening utförs olika kraftigt. Svart kurva är utan sharpening, och blå, grön och röd med ökande grad av sharpening. Om man jämför dessa teoretiskt beräknade kurvor med experimentella MTF-mätningar, Fig. 17.25 b, så ser man en kraftig överensstämmelse.

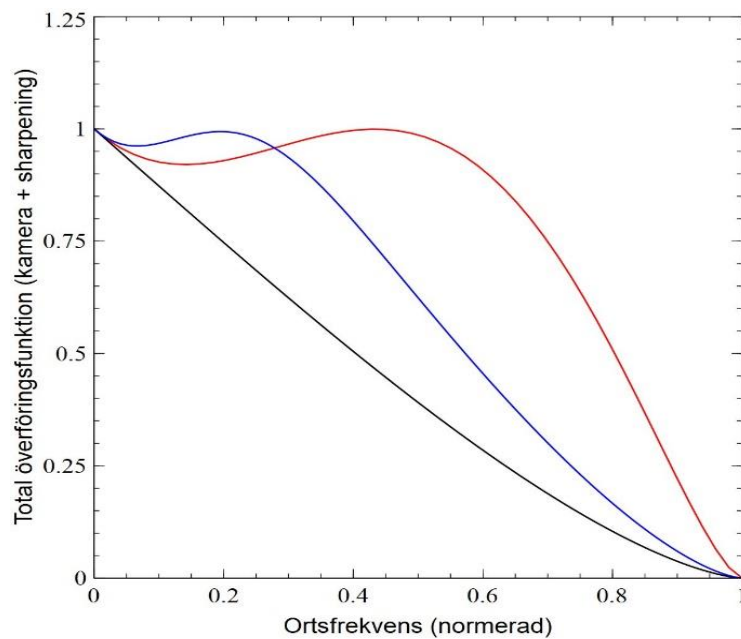


Fig. 17.42. Genom parameterintervall vid sharpening kan man påverka "puckelns" läge och höjd för att få en mer "fyrkantig" kurva. Genom att variera a - och b -värdena i ekv. 17.5 kan vi få t.ex. den blå eller röda kurvan. Den röda kurvan representerar ungefär det mest "fyrkantiga" som går att åstadkomma utifrån ekv. 17.5.

Frågeställningen är alltså om en ”fyrkantig” överföringsfunktion verkligen är optimal. För ett diffraktionsbegränsat objektiv får man en MTF enligt svarta kurvan i Fig. 17.43, där gränsfrekvensen $\frac{1}{\lambda F}$ (λ = ljusvåglängden och F = bländartalet) har normerats till 1. Över denna frekvens finns ingen information överhuvudtaget i bilden, men upp till denna frekvens kan man i princip förstärka upp modulationen (kontrasten) så att den motsvarar ett MTF-värde av 1, illustrerat av den röda kurvan.

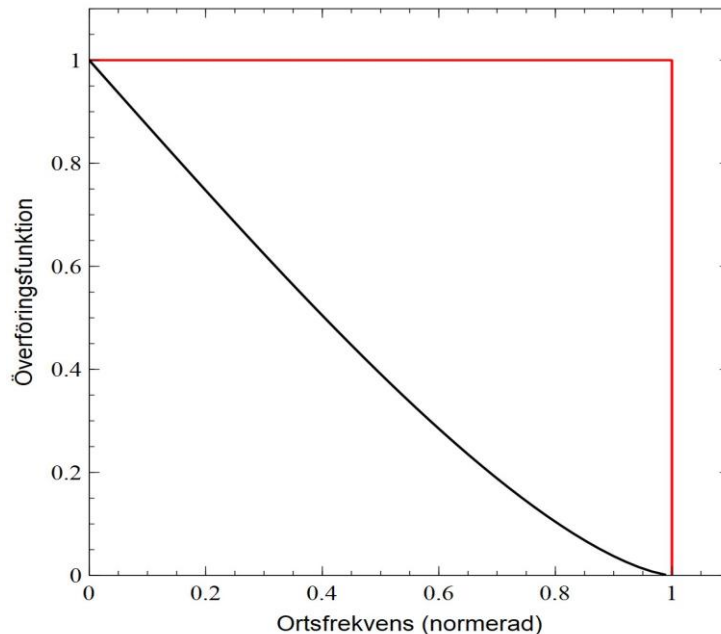


Fig. 17.43. Svart kurva visar MTF för ett diffraktionsbegränsat objektiv. Ortsfrekvensen är normerad så att 1 motsvarar objektivets gränsfrekvens. Den röda kurvan visar vad som skulle vara teoretiskt möjligt att åstadkomma genom att på lämpligt sätt förstärka upp högre Ortsfrekvenser.

Den röda kurvan i Fig. 17.43 svarar mot vad man skulle kunna åstadkomma genom avfaltung, se kap. 17.13. Med sharpening får man en mer ”svängig” kurva som framgår av Fig. 17.41 och 17.42. Ett problem i sammanhanget är att en så kraftig förstärkning av höga frekvenser som är fallet för de röda kurvorna i Fig. 17.42 och 17.43 skulle innebära att bilden blir väldigt brusig. Detta är ett problem som måste hanteras vid t.ex. avfaltung. Men låt oss för ögonblicket bortse från brus, och undersöka vad de två kurvorna i Fig. 17.43 skulle betyda vad gäller avbildningsegenskaperna. För att ge ett exempel på detta ska vi anta att vi har ett motiv bestående av en ljus homogen fyrkant på en mörk homogen bakgrund, Fig. 17.44 a. Vi låter detta motiv avbildas med system som har överföringsfunktioner enligt svart och röd kurva i Fig. 17.43. Resultatet framgår av Fig. 17.44 b, c.

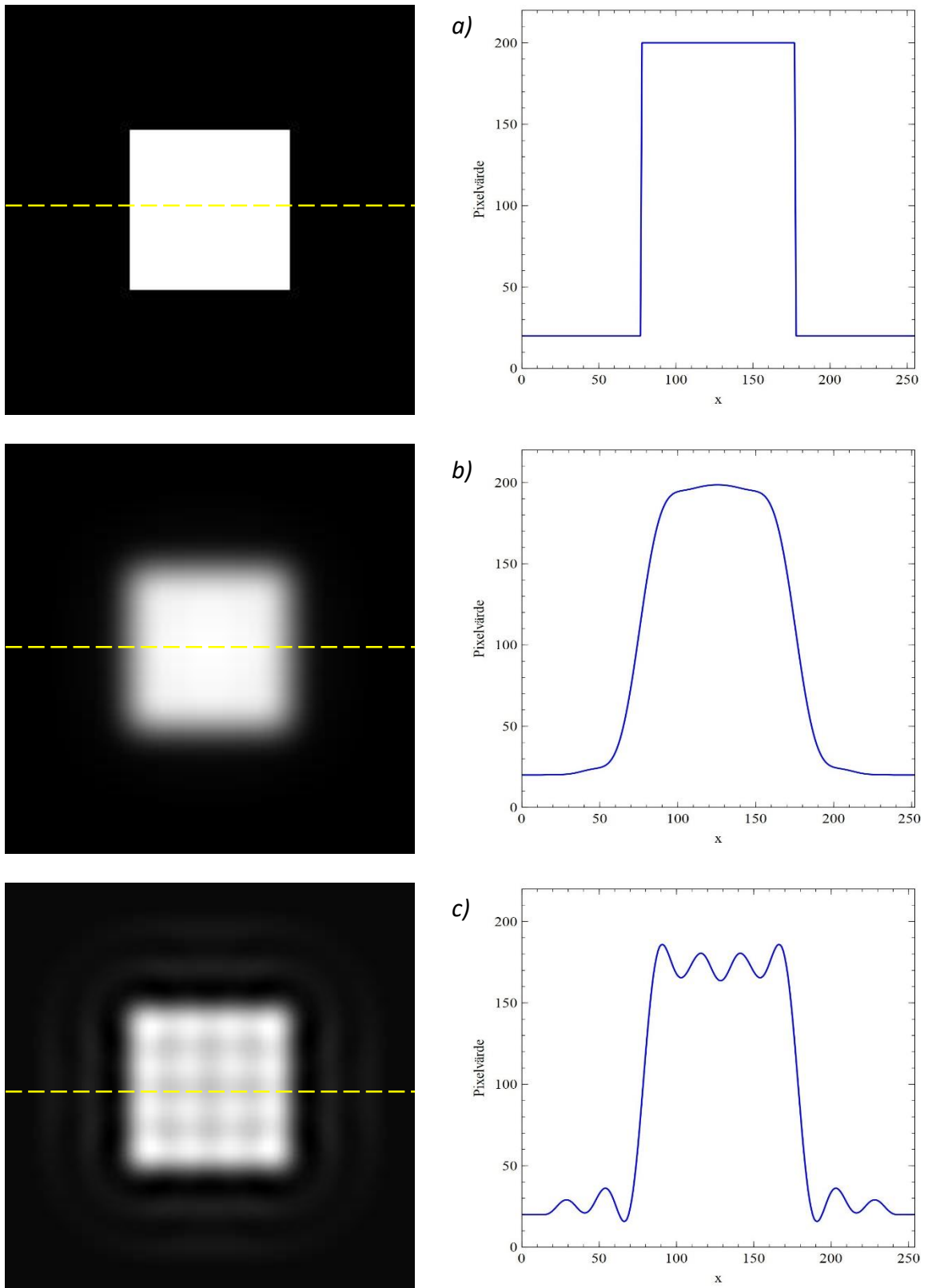
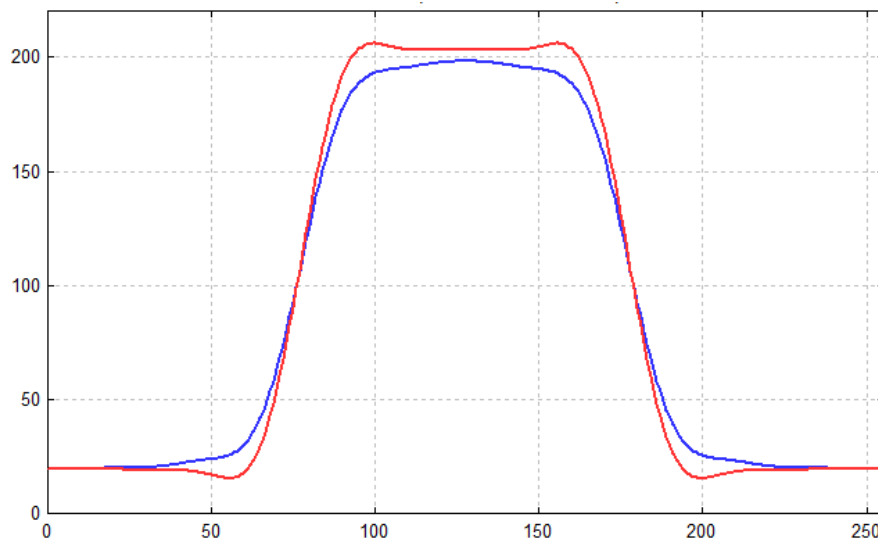


Fig. 17.44. a) Motivet och till höger en intensitetsprofil längs streckade linjen. b) Motivet avbildat enligt svart (diffraktionsbegränsad) kurva i Fig. 17.43. c) Motivet avbildat enligt röd (fyrkantig) kurva i Fig. 17.43.

Motivet avbildas helt klart ”skarpare” i Fig. 17.44 *c* än i *b*. Men samtidigt uppträder en mycket störande haloeffekt vid kanten av den ljusa fyrkanten. Detta syns också tydligt som över- och underslängar i intensitetsprofilen. Detta är ett exempel på ”Gibbs fenomen” som uppträder i alla typer av signaler när man gör en skarp bortfiltrering av höga frekvenser. För att undvika haloeffekter ska man alltså undvika den röda ”fyrkantiga” överföringsfunktionen i Fig. 17.43, och i stället försöka få till en kurva som ger en mer successiv dämpning när vi närmar oss gränsfrekvensen. Trots att överföringsfunktionen vid sharpening inte blir så fyrkantig som röda kurvan i Fig. 17.43, så är risken ändå stor för haloeffekter om inte parametrarna väljs på rätt sätt vilket framgår tydligt i Fig. 17.25 *a*. Uppgiften blir alltså att hitta en kompromiss som gör att bilden ser skarpare ut utan att man störs av haloeffekter. Tyvärr ser man ofta i praktiken att bilder har utsatts för alltför stark sharpening med åtföljande haloeffekter. Detta är särskilt vanligt med bilder från kameror av lägre kvalitet, där man försöker dölja kamerans bristande bildkvalitet med lite kraftigare sharpening.

Ett exempel på en lite lagom uppskärpning av bilden i Fig. 17.44 *b* ges av intensitetsprofilen i Fig. 17.45. Kanten avbildas skarpare än i originalbilden (kraftigare lutning), och att man har ytterst små under- och överlängar dvs. försumbar haloeffekt. Vid sharpening är det hela tiden en kompromiss mellan kantskärpa och haloeffekt. Man kan inte öka det ena utan att öka det andra. Men man ska komma ihåg att bildinnehållet också spelar roll. Innehåller inte motivet skarpa kanter med hög kontrast så tål bilden kraftigare sharpening.



*Fig. 17.45. Efter att bilden i Fig. 17.44 *b* behandlats med ”lite lagom” sharpening erhålls den röda intensitetsprofilen. Den ursprungliga profilen visas i blått. Kanten har blivit skarpare och haloeffekten är fortfarande försumbar.*

Som vi tidigare sett så innebär sharpening med oskarp mask att man ska välja lämpliga värden på ”radie” och ”vikt”, och i vissa program även ”tröskelvärde”. Så, hur ska man då välja lämpliga värden? Låt oss börja med radie och vikt. I Fig. 17.40 visas schematiskt hur överföringsfunktionen för sharpening påverkas av dessa parametrar. Överför vi det till totala överföringsfunktionen (kamera plus

sharpening), så innebär ökad vikt att puckeln i kurvan blir högre men ligger kvar på samma ställe (Fig. 17.41). Ökad radie innebär att puckeln flyttas mot lägre ortsfrekvenser (Fig. 17.42). Lämpligt läge och höjd på puckeln beror på ursprungsbildens utseende och vad man vill uppnå. Det kan vara lämpligt att börja med valet av radie. Ett lågt värde på radien ökar kontrasten mest på små detaljer i bilden, medan ett högre värde kommer att öka kontrasten även på lite större detaljer. En tumregel är att sätta radien till att motsvara de minsta detaljerna i bilden, eftersom det oftast är dessa man vill framhäva bättre. För en "normalbra" bild, dvs ingen felfokusering eller andra uppenbara fel, kan det därför vara lämpligt att börja med en radie av ca. 1 pixel. Detta är ofta defaultvärde i bildbehandlingsprogram. För bilder som är lite suddiga från början kan man öka på radien. Programmen brukar också föreslå ett vikt-värde som kan uttryckas lite olika i olika program, t.ex. som 50%. Man kan börja med det föreslagna vikt-värdet och se vad resultatet blir. Vikt-värdet justeras sedan så att man får optimal uppskärpning utan att få störande haloeffekter. Exempel på bildmässiga effekter av olika kraftig sharpening visas i Fig. 17.46, som är ett litet utsnitt ur en betydligt större ursprungsbild.

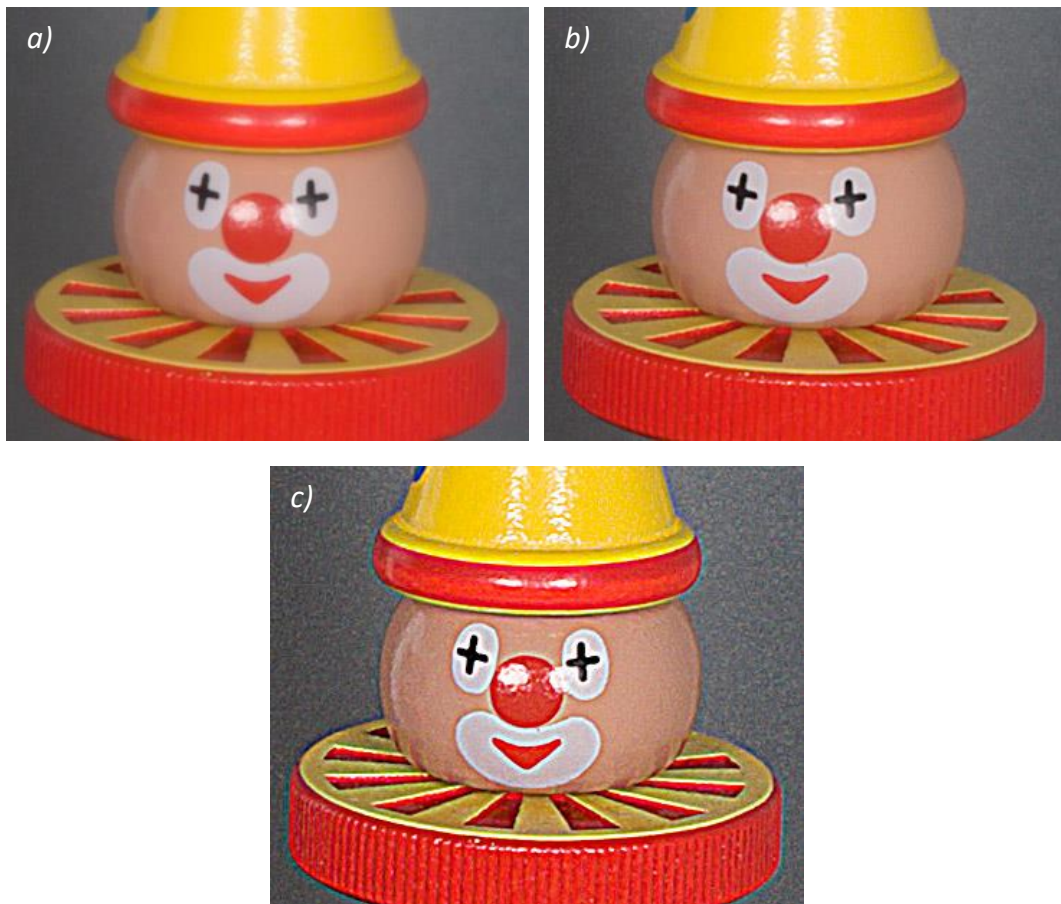


Fig. 17.46. Effekten av olika kraftig sharpening. a) Originalbild utan någon sharpening. b) En ganska lagom sharpening har utförts. c) För kraftig sharpening skapar en "hård" bild med haloeffekter, t.ex. vid ögon och mun. Notera också det ökade bruset i bakgrunden.

Trenden har under många år gått mot allt fler pixlar i sensorn, och det medför allt oftare att det är optiken som begränsar detaljrikedomen i bilderna. När man kommer ner på pixelnivå, så innebär detta att även bilder från de allra bästa kamerorna ser lite lätt suddiga ut. I sådana fall behöver man använda lite större värde på radien än 1 pixel för att få ett bra resultat, även när bilderna är tagna under optimala förhållanden.

Vissa bildbehandlingsprogram har också ett "tröskelvärde" som ska anges vid sharpening. Värdet är ofta satt till noll som default. Tröskelvärdets uppgift är att styra vilka områden i bilden som ska utsättas för sharpening. Värdet noll innebär att sharpening görs överallt. Ett värde större än noll innebär att homogena bildområden, dvs områden utan detaljer (t.ex. en enfärgad bakgrund), inte kommer att utsättas för sharpening. Att utföra sharpening i homogena områden kommer bara att öka bruset utan att man vinner något, se Fig. 17.46 c. Visserligen gör sharpening att bruset ökar också i detaljrika områden, men där är det inte lika störande. Ju högre tröskelvärde man väljer, desto mer inhomogent ska ett område vara innan det utsätts för sharpening. I Fig. 17.47 visas effekten av att öka tröskelvärdet i fallet som visades i Fig. 17.46 c.

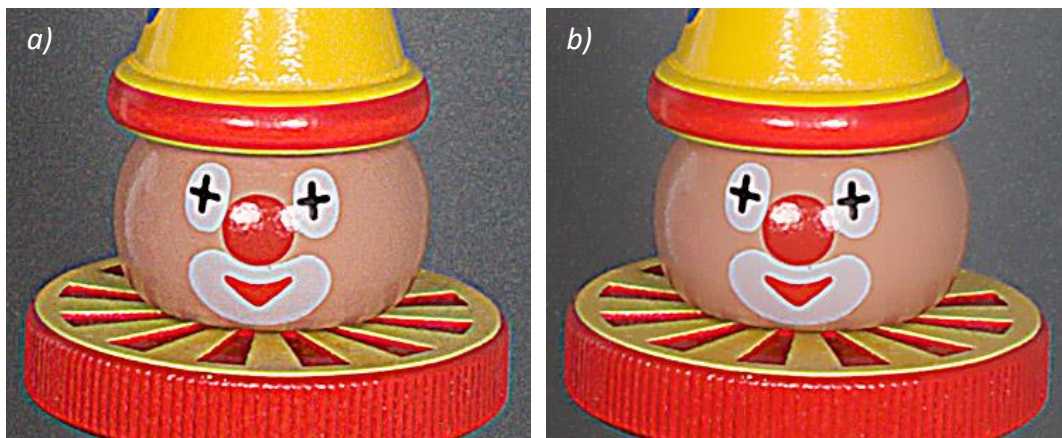


Fig. 17.47. a) Samma figur som 17.46 c. Tröskelvärdet var satt till noll. b) Effekten av att öka tröskelvärdet. Bruset i homogena bildpartier har minskat betydligt.

17.15. Softfokus

Genom att byta ut minustecknet i ekv. 17.5 mot ett plustecken, dvs addera den suddiga kopian och originalbilden, erhåller man något som kallas softfokus-effekt eller mjukteckning. Syftet med detta kan vara att göra vissa typer av bilder lite mjukare och mer stämningsfyllda, exempelvis när man jobbar med porträttfotografering. Men det handlar inte bara om att göra bilden oskarpare, vilket är vad som sker med verktyget "oskärpa" (blurring) i bildbehandlingsprogram. För att få softfokus-effekt är viktigt att det finns kvar skarpa detaljer i bilden, även om kontrasten på dessa är reducerad. Det hela kan liknas vid en skarp bild på vilken man överlagrat lite dis runt ljusa partier (så kallad överstrålningseffekt). Sådana effekter har sedan lång tid tillbaka åstadkommits med hjälp av speciella

mjuktecknande objektiv eller optiska filter. Men vi ska nu se vad man kan åstadkomma med bildbehandling.

Teorin bakom det hela är helt analog med den som presenterades i kap. 17.14, men med minustecknet i ekv. 17.5 utbytt mot ett plustecken. Detta kommer att ge en överföringsfunktion enligt Fig. 17.48 a, medan det gängse verktyget för oskärpa, Gaussisk oskärpa, ger en kurva enligt Fig. 17.48 b. Den bildmässiga skillnaden mellan dessa två fall kommer att illustreras lite senare.

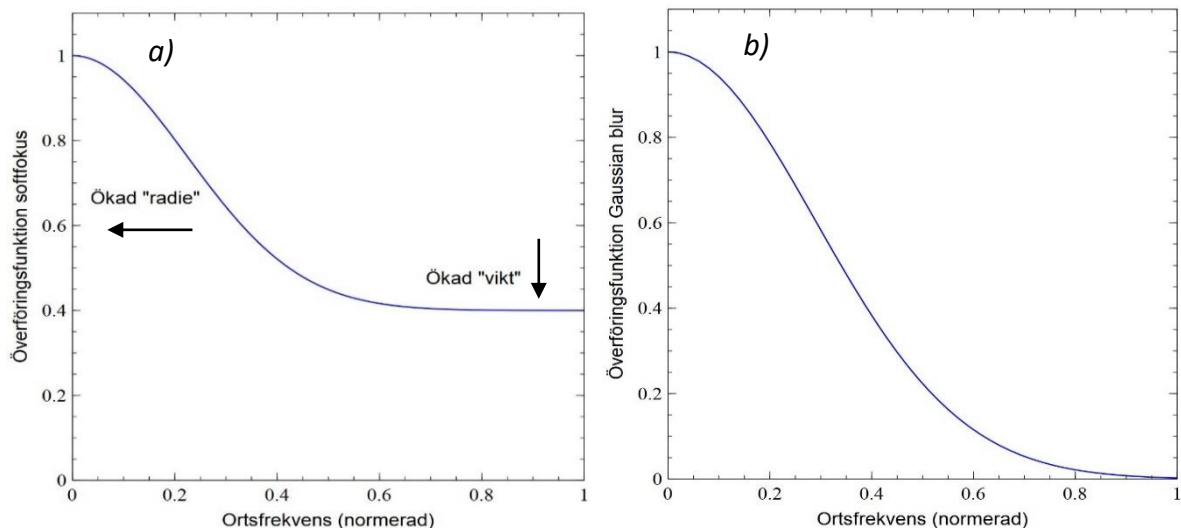


Fig. 17.48. a) Överföringsfunktionen för softfokus. Effekten är att dämpa höga ortsfrekvenser, men bara ner till en viss nivå. Jämför Fig. 17.40 som visar motsvarande kurva för sharpening. b) Överföringsfunktionen för "Gaussisk oskärpa", den funktion för oskärpa som vanligen används vid bildbehandling.

På motsvarande sätt som för sharpening, kan vi få totala överföringsfunktionen genom att multiplicera kamerans MTF med överföringsfunktionen för softfokus. Exempel på detta visas i Fig. 17.49 och 17.50, vilka kan jämföras med motsvarande fall för sharpening i Fig. 17.41 och 17.42. Precis som vid sharpening har man två parametrar att spela med, radie och vikt. Skillnaden är bara att större vikt nu ger en lägre överföringskurva istället för en kraftigare puckel som i fallet sharpening. En ökad radie kommer att innebära att kurvan faller snabbare i början. Ett praktiskt exempel på den påverkan softfokus och Gaussisk oskärpa har på en bild visas i Fig. 17.51, som är ett litet utsnitt ur en större bild.

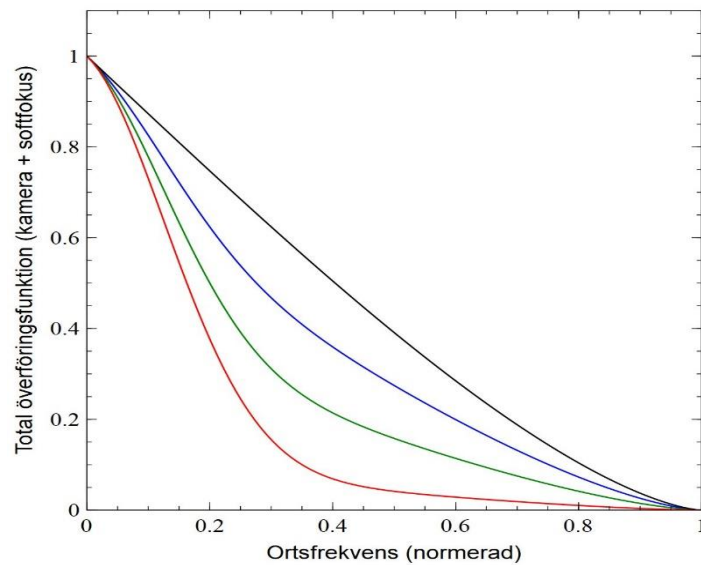


Fig. 17.49. Exempel på total överföringsfunktion vid softfokus-effekt. Svart kurva är utan softfokus, och blå, grön och röd kurva med en successiv ökning av parametern "vikt".

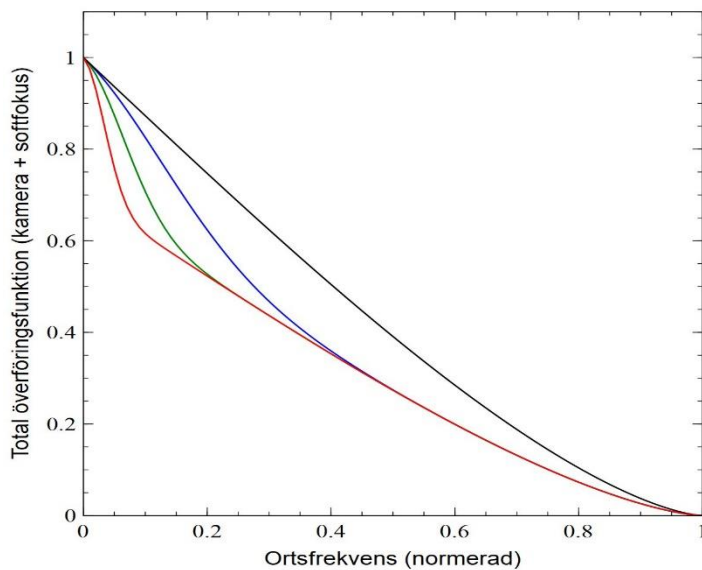


Fig. 17.50. Svart kurva är överföringsfunktionen utan soft-fokus. Blå, grön och röd kurva visar effekten av en ökning av parametern "radie". Man kan notera att gröna och röda kurvorna liknar överföringsfunktionen vid påverkan av ströljus, Fig. 17.13. Detta är förväntat eftersom addition av en extremt utsuddad bild liknar effekten av ströljus i kameran, vilket ger dålig softfokus-effekt eftersom man inte vill ha en jämn "ljusdimma" över bilden utan endast viss "överstrålning" från ljusa partier.

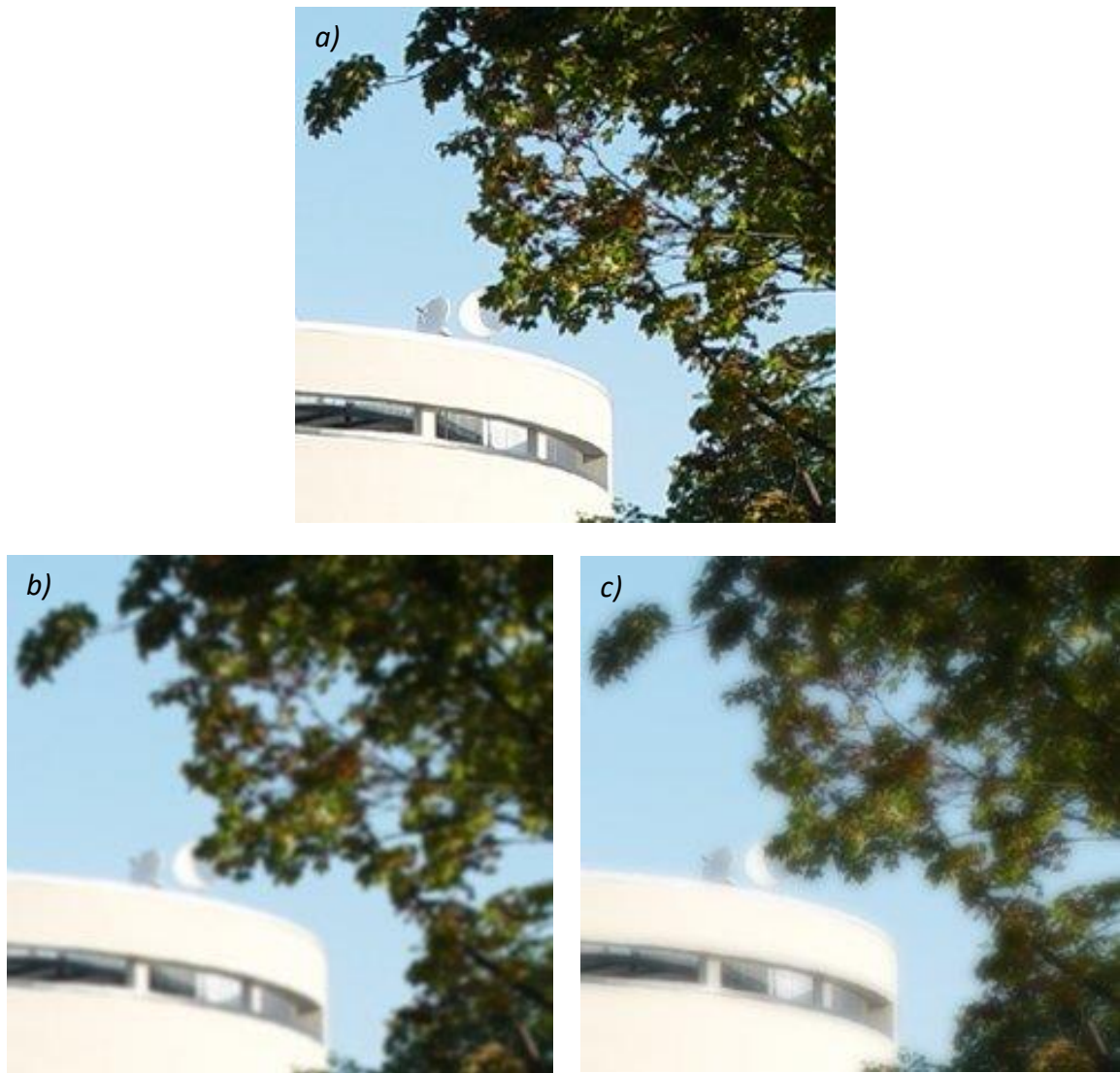


Fig. 17.51. Effekten av Gaussisk oskärpa jämfört med softfokus. a) Original. b) Gaussisk oskärpa. c) Softfokus. I fallet c) får man ett intryck av en skarp bild inbäddad i ett "romantiskt skimmer", vilket ofta gör sig bra vid t.ex. porträttfoto.

Jämfört med sharpening, är det svårare att ge allmänna råd angående val av radie och vikt när det gäller softfokus. Huvudsyftet med sharpening är ju i allmänhet att skapa en bild som ser så naturlig ut som möjligt. Softfokus å andra sidan vill ju skapa en konstnärlig stämningseffekt. Det betyder att man vid softfokus ofta kan ta i lite extra när man sätter parametervärden. Men precis som vid sharpening så är en god regel att det är bättre att ta för lite än för mycket. I bägge fallen påverkar också originalbildens utseende parametervälen.

18. Infrarödfotografi

Infrarödfotografi avser i allmänhet fotografering inom det så kallade nära IR-området, vilket sträcker sig från 700 nm upp till ca. 1.5 μm våglängd. Dessa bilder ska inte förväxlas med värmekamerabilder (termovision) vilka registrerar en tiopotens längre våglängder (ca. 10 μm), se kap. 25.

Vad finns det för intressanta saker att fotografera i nära IR-området? Låt oss först konstatera att om man ska fotografera egenutstrålningen från en kropp så bör dess temperatur vara över 300° C, och även då kan exponeringstider på åtskilliga minuter behövas. Det går alltså inte att fotografera temperaturfördelningen hos kroppar som har en temperatur nära rumstemperatur. Den vanligaste tillämpningen är istället att man, precis som vid vanlig fotografi, fotograferar strålning som reflekterats av motivet. Lämpliga strålningskällor vid IR-fotografering är solen och glödlampor, eftersom dessa utstrålar mycket IR. Ett intressant motiv för IR-foto är vegetation. De flesta typer av vegetation reflekterar betydligt mer i nära IR än i synligt ljus. Detta beror på att klorofyll absorberar det mesta av synliga ljuset. IR transmitteras däremot, och kan därmed reflekteras av underliggande skikt i växternas vävnader (gränsytor mellan celler och luftfyllda håligheter reflekterar IR bra). Särskilt hög IR-reflektion får man hos lövträd i samband med lövsprickningen. När löven åldras minskar i allmänhet IR-reflektionen. Om växter drabbas av någon sjukdom eller torka kan IR-reflektionen minska drastiskt. Ofta sker detta innan man ser någon påtaglig förändring i det synliga våglängdsområdet. IR-fotografering kan därför användas för att på ett tidigt stadium upptäcka "stressade" växter.

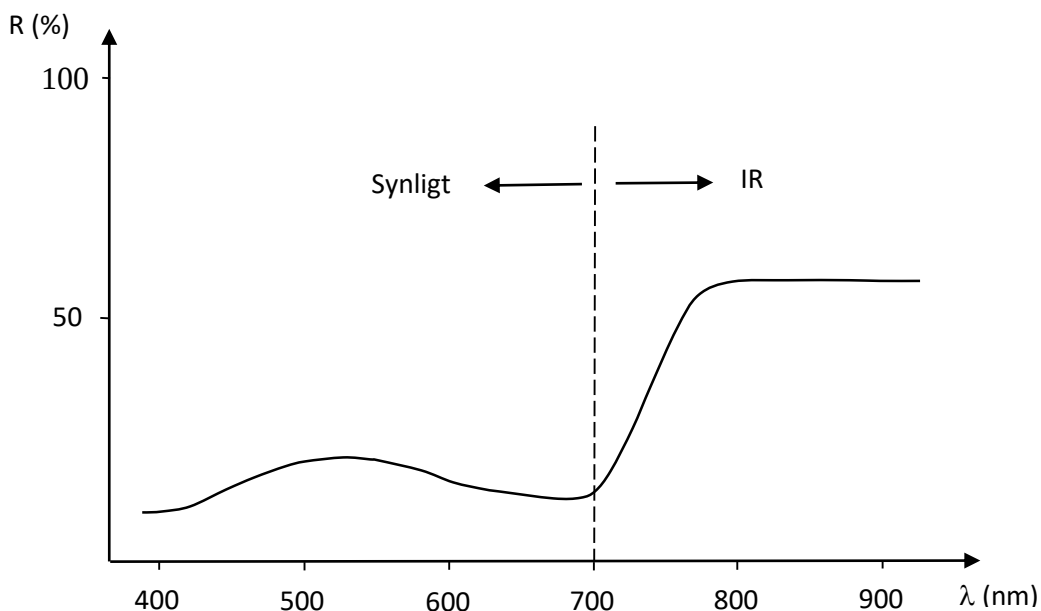


Fig. 18.1. Exempel på spektral reflektionsförmåga hos vegetation.

Ständigt gröna växter, t.ex. barrträd, har oftast en betydligt lägre IR-reflektion än lövträd, Fig. 18.2. Om man flygfotograferar en blandskog med en IR-känslig kamera, kan man vanligen mycket tydligt se vilka träd som är barr- respektive lövträd. På tal om flygfotografering, så kan det framhållas att IR-strålning tränger igenom avståndsdisk mycket effektivare än synligt ljus. Anledningen är att avståndsdisk uppkommer genom ljusspridning mot mycket små partiklar (s.k. Rayleighspridning). Man får då en absorptionskoefficient som är proportionell mot $\frac{1}{\lambda^4}$. Detta innebär att man gör stora vinster genom att öka våglängden. Denna vinst gäller för avståndsdisk, medan kraftigare disk och dimma (dvs. små vattendroppar) ej kan genomträngas av IR-strålningen.



Fig. 18.2. Fotografier tagna i våglängdsintervallet 700-900 nm. Sädesfälten i bildernas förgrund reflekterar IR kraftigt och ser därför ljusa ut. Blått himmelsljus innehåller knappt någon IR och ser därför svart ut. Notera i undre bilden att lövträd reflekterar mera IR än barrträd.

IR-fotografering kan även användas för medicinska tillämpningar. Genom att IR tränger igenom vävnader bättre än synligt ljus, kan man t.ex. avbilda blodkärl som ligger en liten bit under huden. Förändringar hos dessa kärl uppträder vid vissa sjukdomstillstånd. En annan tillämpning av IR-foto är dokumentation av gamla skrifter som blivit hårt angripna av tidens tand (missfärgning, mögel etc.). Ofta framträder i sådana fall skriften betydligt bättre i IR-området. Detsamma gäller dokument som avsiktligt påverkats genom övermålning med t.ex. tusch, eller som har raderats med kemikalier. I sådana fall kan ofta den ursprungliga texten vara fullt synlig i IR-området.

För att ta IR-bilder användes tidigare fotografisk film speciellt anpassad för detta ändamål (bilderna i Fig. 18.2 togs på detta sätt). Elektroniska sensorer i digitalkameror (CCD eller CMOS) har stor potential för användning inom IR-foto, eftersom känsligheten är god inom nära IR-området (Fig. 14.14). Denna känslighet är så stor att den är störande för vanlig fotografering i det synliga området (RGB-filtren i Bayer-mosaiken släpper igenom IR). Därför sitter i digitalkameror ett filter som absorberar IR-strålning. Hos vissa tidiga digitalkameror kunde detta filter fällas bort ur strålgången, varvid kameran blev en utmärkt IR-kamera. För att ta bort synligt ljus använder man ett så kallat svartfilter, som bara släpper igenom våglängder över ca. 700 nm. Bilder som erhålls på detta sätt blir naturligtvis svartvita. Vita eller ljusgrå partier motsvarar hög IR-strålning, medan mörka partier motsvarar låg IR-strålning. Även vissa digitalkameror där filtret som absorberar IR inte går att fälla bort kan vara användbara för IR-foto. Filtret har kanske en viss resttransmission, och om man använder långa exponeringstider och stora bländaröppningar kan man ibland få användbara bilder.

Vid all IR-fotografering ska man vara beredd på att bildskärpan blir lite sämre än för synligt ljus. Kameraoptik är nämligen optimerad för synliga området, och därför kan man förvänta sig att avbildningsfelen blir större i IR. Dessutom blir diffraktionen större vid längre våglängder. Även CCD- och CMOS-sensorerna ger sämre bildskärpa i IR-området.

Förutom de svartvita IR-bilder som beskrivits ovan, så förekommer ofta så kallade färg-IR-bilder. Tidigare användes särskild färgfilm för IR-foto, medan man idag kan köpa speciella digitalkameror för detta syfte. Precis som i vanlig färgfotografi, så registreras tre våglängdsband som sedan återges med tre olika färger i den slutliga bilden. Till skillnad från vanliga färgfotografier, så återger IR-bilder inte motivets naturliga färger utan man får så kallade "falska" färger i bilden. I Fig. 18.3 visas hur en IR-kamera för färgbilder kan vara uppbyggd.

För att tolka en IR-färgbild behöver man förstå vad de olika "falska" färgerna representerar i verkligheten. Detta är mycket enkelt om man bara kommer ihåg att alla färger återges "ett snäpp" mot kortare våglängder, Fig. 18.4. Sålunda kommer IR i motivet att återges rött, rött i motivet återges grönt, grönt i motivet återges blått och det som är blått i motivet återges inte alls. Man offerar det blå våglängdsområdet därför att det är minst intressant i de flesta tillämpningar. I tabellen på nästa sida ges exempel på färgåtergivningen som erhålls vid IR-fotografering.

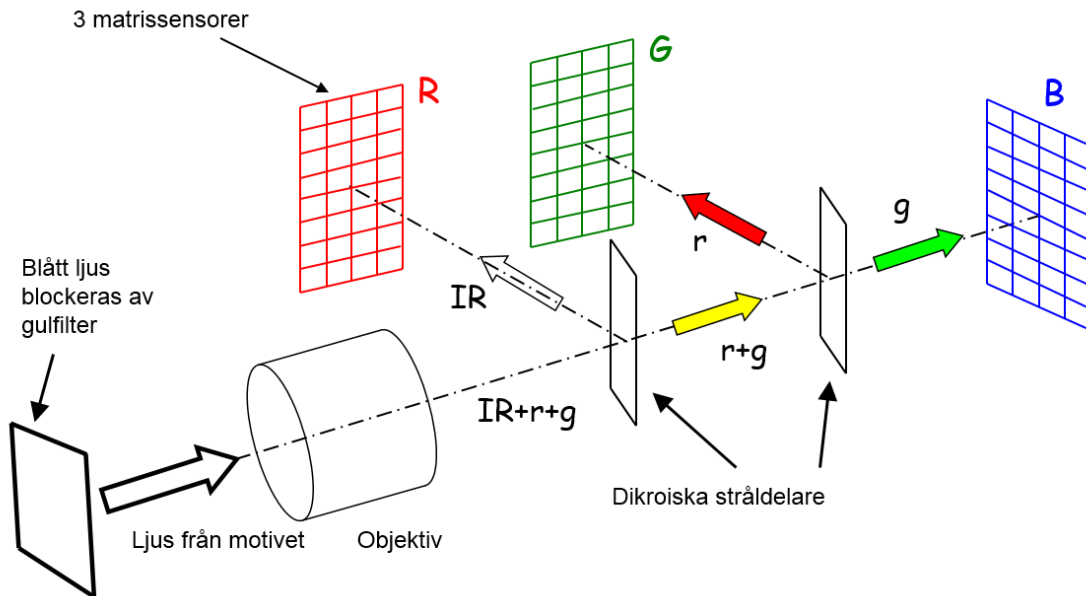


Fig. 18.3. Uppbyggnaden hos en färgkamera för infrarödfotografi. Dikroiska stråldelare reflekterar vissa våglängder medan andra transmittas. Första stråldelaren reflekterar sålunda IR, men transmittar allt synligt ljus. Andra stråldelaren reflekterar rött ljus, medan kortare våglängder transmittas. Signalerna från de tre sensorerna färgkodas i slutbilden med rött, grönt och blått enligt figuren. I den slutliga bilden återges därför IR-strålning från motivet med rött, röd motivfärg återges grönt och grön motivfärg återges blå. Blått ljus från motivet kommer att absorberas av gulfiltret på kameran, och därför kommer blått att återges som svart.

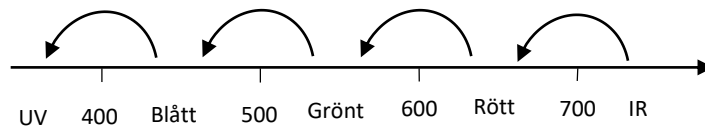


Fig. 18.4. I en IR-färgbild återges motivets alla "färger" ett "snäpp" åt kortare våglängder. I tabellen nedan ges exempel på hur detta påverkar bildens utseende.

<u>Motivets "färg"</u>	<u>Bildens färg</u>
IR	Röd
Röd	Grön
Grön	Blå
Blå	Svart
Gul (R + G)	Cyan (G + B)
Magenta (R + B)	Grön (motivets blå färg återges ej)
Cyan (G + B)	Blå (motivets blå färg återges ej)

Eftersom vegetation ofta reflekterar stora mängder IR-strålning, kommer IR-färgbilder av landskap att ha en kraftig röd färgton. Förutom vegetation så finns det många andra motiv som reflekterar stora mängder IR, t.ex. vissa tyger. Låt oss titta på ett exempel.

Antag att man fotograferar svenska flaggan med en IR-färgkamera. Antag vidare att flaggans tyg reflekterar IR över hela ytan (dvs. både i gula korset och blå bakgrunden). Hur kommer flaggan att återges på IR-bilden? Jo, bakgrunden strålar ut blå + IR. Blått absorberas i gulfiltret, medan IR ger röd färg i bilden. Bakgrunden återges alltså röd. Gula korset strålar ut G + R + IR. Dessa kommer att återges som B + G + R = vitt. Korset återges alltså i vit färg. Flaggan kommer alltså att återges som danska flaggan på IR-bilden (förutsatt att hela tyget reflekterar IR). I Fig. 18.5 visas exempel på hur några föremål återges i synligt ljus respektive med IR-foto.

Gamla IR-färgbilder upptagna med film uppvisar samma typ av färgskala som IR-färgbilder fotograferade med digitalkamera. Så vad gäller bildtolkningen spelar det ingen roll om det rör sig om film eller digitalbilder.

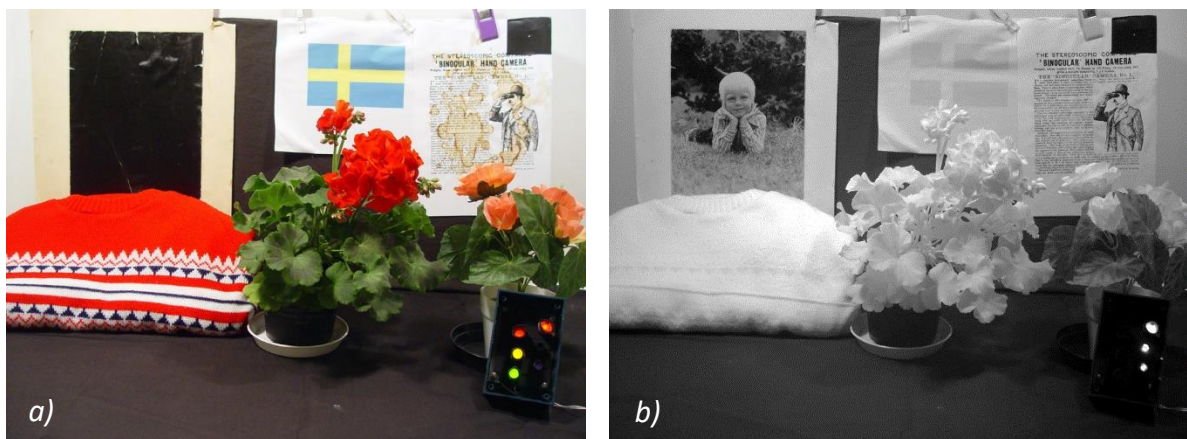


Fig. 18.5. a) Vanligt färgfoto. b) IR-foto (svartvitt).

c) IR-färgbild. Notera att det tuschövermålade barnporträttet framträder tydligt i IR, samt att de bruna kaffebläckarna försvinner. Den vänstra krukväxten är riktig (hög IR-reflektion), medan den högra är konstgjord (låg IR-reflektion). Den rödfärgade tröjan har hög IR-reflektion, och svenska flaggan återges snarast som dansk i IR-färgbilden (se texten ovan). De färgade ljuspunkterna är lysdioder plus en glödlampa med rödfilter.

(Kamera: Sony DSC-F717. Filter IR720 användes för bild b. Bild c framställdes genom att kombinera bild b med röda och gröna färgkanalerna i bild a.)



19. Ultraviolett- och fluorescensfotografi

I föregående kapitel behandlades fotografering med strålning vars våglängd var längre än synligt ljus. Vi ska nu gå åt andra hållet och titta på våglängder kortare än synligt ljus, dvs. ultravioletta våglängder ($\lambda < 400$ nm). Vi ska också titta på fluorescensfotografi, dvs då man belyser motivet med intensiv strålning av kort våglängd (vanligen ultravioletta, blåa eller gröna våglängder) varvid det lyser upp med en längre våglängd. Fig. 19.1 visar exempel på fluorescensbilder från mikroskop som registrerats vid fysikinstitutionen på KTH i samarbete med medicinska institutioner.

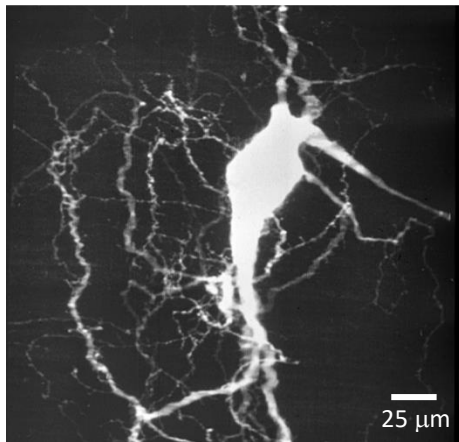


Fig. 19.1. a) Mikroskopibild av nervcell från nejonöga. Cellen var inmärkt med det fluorescerande färgämnet Lucifer Yellow. Belysningen utgjordes av laserljus med våglängden 488 nm, varvid grönt ljus (ca. 530 nm) emitterades.

Bildregistreringen gjordes med ett scannande, s.k. konfokalt, mikroskop. (Prep.: Dr. Peter Wallén, Inst. för Neurofysiologi, Karolinska Inst.)

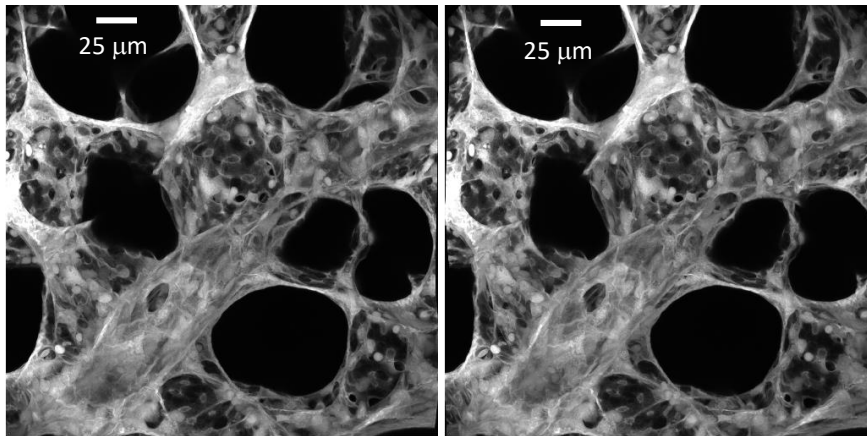


Fig. 19.1. b) Stereobild (jfr. kap. 8) av lungvävnad inläst med konfokal mikroskopi. Samma färgämne och våglängder användes som i Fig. 19.1 a). (Prep.: Dr. E. Oldmixon, Brown University, USA.)

Sensorer i vanliga digitalkameror har i allmänhet inte så hög känslighet i ultraviolet (UV) området (Fig. 14.14). Det finns dock speciella (dyra) digitalkamerasensorer som utvecklats speciellt för UV-fotografi. Dessa har ofta god känslighet ner till 200 nm. Det kan nämnas att vanlig fotografisk film är användbar ner till våglängder runt 300 nm.

En annan begränsande faktor vid UV-fotografering utgörs av optiken, eftersom vanliga glassorter absorberar strålning under ca. 350 nm. Visserligen finns specialoptik för kortare våglängder, men den är mycket dyr och urvalet är litet. En ytterligare komplikation vid fotografering med korta våglängder är att luft absorberar denna strålning kraftigt (absorption i ozon). I stort sett inget av solens strålning under 300 nm tränger igenom jordatmosfären, och under 200 nm kan man inte ens jobba i en liten laboratorieuppställning utan att pumpa vakuum. På grund av alla dessa begränsningar, och främst de som gäller optiken, brukar man vid UV-fotografering ofta jobba i våglängdsintervallet 350 – 400 nm. Som användningsområden för UV-fotografi kan nämnas medicin och kriminalteknik. Inom medicinen används den ibland för att synliggöra hudpigment. Detta absorberar nämligen UV-strålning mycket kraftigt (och ger oss solbränna, och, i värsta fall, malignt melanom). Pigmentfria delar av huden reflekterar däremot UV-strålningen. Inom kriminaltekniken kan UV-fotografering vara ett värdefullt hjälpmedel för att upptäcka förfälskningar, studera textilfibrer mm. I viss mån används UV-fotografi inom mikroskopi, bl.a. för att genom den kortare våglängden uppnå en högre upplösning. För att undvika att synliga våglängder exponerar sensorn vid UV-fotografering, så kan man använda speciella filter som transmittar UV men blockerar synligt ljus. Alternativt kan man använda en ljuskälla som bara ger UV-strålning. Vanligaste ljuskällan är en kvicksilverlampa, vilken ger ett antal våglängder i framförallt UV och blått. Oftast använd är spektrallinjen vid 365 nm som dels är mycket stark, dels ligger i ett lämpligt våglängdsintervall som gör att vanlig optik kan användas. Vid användning av kvicksilverlampor (och ev. andra typer av UV-lampor) ska man iaktta försiktighet. Dels arbetar dessa lampor ofta under högt tryck vilket innebär en explosionsrisk, dels kan den utsända strålningen ge skador på hud och ögon. I en del fall bildas också ansevärliga mängder giftigt ozon genom att strålningen bryter ned luftens syremolekyler. Ozonbildning märks dock lätt genom den karakteristiska lukten.

Renodlad ultraviolettphotografi används relativt sällan, det är vanligare med fluorescensfotografi. Speciellt har det skett ett stort uppsving inom detta område på mikroskopisidan under senare år, Fig. 19.1. Vid fluorescensfotografi belyses vanligtvis motivet med UV eller kortvågigt ljus, och det fluorescensljus som därvid utsänds vid längre våglängder (ofta grönt, gult eller rött) fotograferas. Bland fluorescerande ämnen finns många mineraler och organiska ämnen. Genom att olika ämnen sänder ut fluorescensljus med olika färger, kan man se t.ex. fördelningen av olika mineraler i en bergart. Många växter och djur uppvisar också fluorescens, speciellt vid belysning med UV. Dessutom kan molekyler av fluorescerande ämnen bindas till antikroppar, och därmed kan man fluorescensmärka t.ex. signalsubstanser i nervceller, DNA och andra kemiska ämnen i biologiska prover. Detta utnyttjas mycket inom biologisk och medicinsk mikroskopi. Ofta gör man i sådana sammanhang multipelinmärkningar med flera olika fluoroforer i olika färger, och som binder till olika substanser i provet. Tidigare använde man oftast fluoroforer som skulle belysas med UV eller blått ljus, men idag finns även fluoroforer som ska belysas med grönt, gult och även rött ljus. Eftersom fluorescensljuset har längre våglängd än det belysande ljuset innebär det att man vid rödexcitation får fluorescens i IR. Fluorescensfotografi används även i medicinska sammanhang, och i makroskopisk skala. Ett exempel på detta är så kallad fluorescensangiografi, varvid man intravenöst injicerar natriumfluorescein i

patienten. Genom en serie fotografier studerar man sedan hur snabbt detta ämne tränger fram i olika delar av patientens hud, vilket ger värdefull information om den perifera blodcirkulationen (dvs. hur bra kapillärnätet fungerar). Denna information är värdefull för bedömning av effekten av olika cirkulationsbefrämjande läkemedel, samt som underlag vid amputationer mm. Fluorescensangiografi används också ofta vid undersökningar av blodcirkulationen i ögats näthinna.

I Fig. 19.2 visas ett exempel på hur en uppställning för fluorescensfotografi kan se ut.

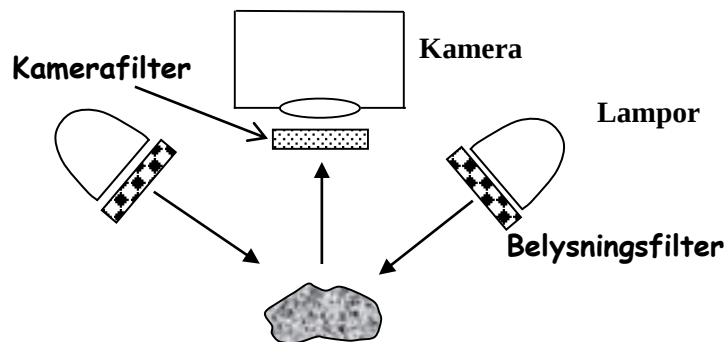


Fig. 19.2. Uppställning för fluorescensfotografi. Belysningsfilterna släpper igenom kortvågig strålning (ofta UV eller blått ljus) och kamerafiltret släpper igenom det mer långvågiga fluorescensljuset.

Filtervalet är av stor betydelse vid fluorescensfotografi. Motivet ska belysas med våglängder som effektivt absorberas av det fluorescerande ämnet, varvid dess molekyler intar ett exciterat tillstånd. Genom lämpligt val av ljuskälla och belysningsfilter kan detta uppnås. Framför kameran måste ett så kallat barriärfilter monteras. Dess uppgift är att blockera reflekterad strålning från motivet, men samtidigt transmitta fluorescensljuset. Eftersom fluorescensljuset oftast är flera storleksordningar svagare än belysningsljuset, är det mycket viktigt att kamerafiltret kraftigt dämpar reflekterad strålning. I Fig. 19.3 visas typexempel på absorptions- och emissionsspektra för ett fluorescerande färgämne. Dessutom är lämpliga filterkurvor för belysnings- och kamerafilter utritade tillsammans med den starka spektrallinjen vid 365 nm för en kvicksilverlampa. Vid fluorescensfotografering kan vanligtvis normal fotografisk utrustning användas eftersom fluorescensljuset nästan alltid är synligt.

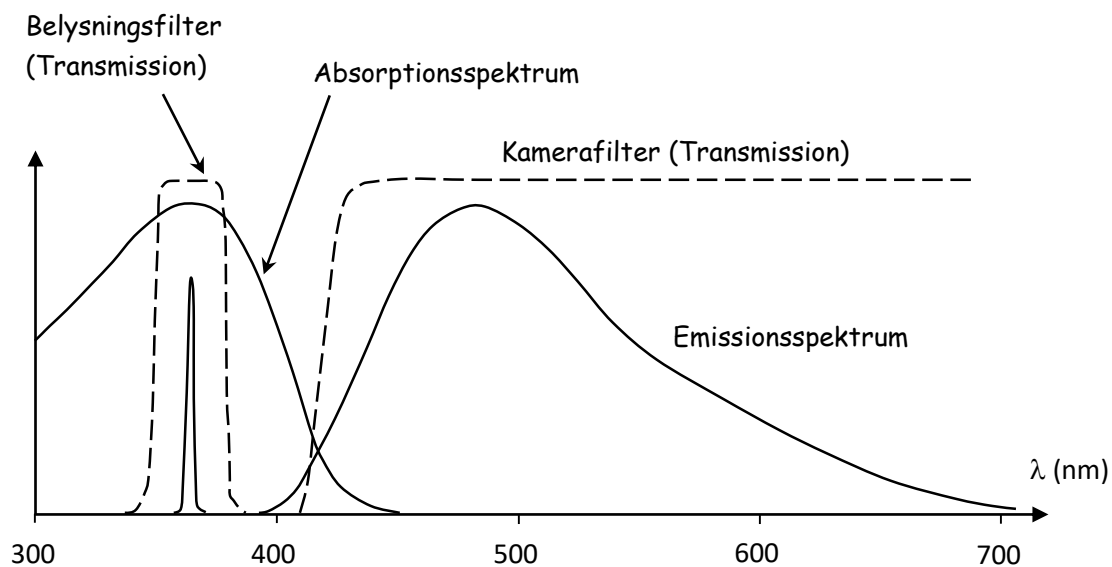


Fig. 19.3. Exempel på absorptions- och emissionsspektrum för ett fluorescerande färgämne. I figuren finns även utritat en lämplig spektrallinje för excitation (365 nm hos kvicksilver), samt lämpliga spektrala transmissionskurvor för belysnings- och kamerafilter.

20. Röntgenfotografi

20.1. Inledning

Röntgenstrålning är elektromagnetisk strålning i våglängdsintervallet ca. 0.001 – 50 nm. Vid röntgenfotografi utnyttjas röntgenstrålarnas stora penetrationsförmåga, vilken ökar ju kortare våglängden är. Detta gör att man med röntgenstrålar kan genomlysa människokroppar, flygplansvingar, brobalkar mm. På så sätt har man ett icke-förstörande redskap för att undersöka hur saker ser ut inuti. Genom att olika material har olika stor röntgenabsorption uppstår kontrast i den registrerade bilden. Sålunda kan man se skillnad på ben och mjukdelar i människokroppen, samt solitt järn respektive luftfyllda sprickor i en brobalk. Röntgenfotografi skiljer sig på flera sätt från vanlig fotografi. Dessa skillnader inkluderar strålningskällan, avbildningen och strålningsdetektionen.

20.2. Röntgenröret

Strålningskällan vid röntgenfotografering utgörs vanligtvis av ett så kallat röntgenrör, Fig. 20.1.

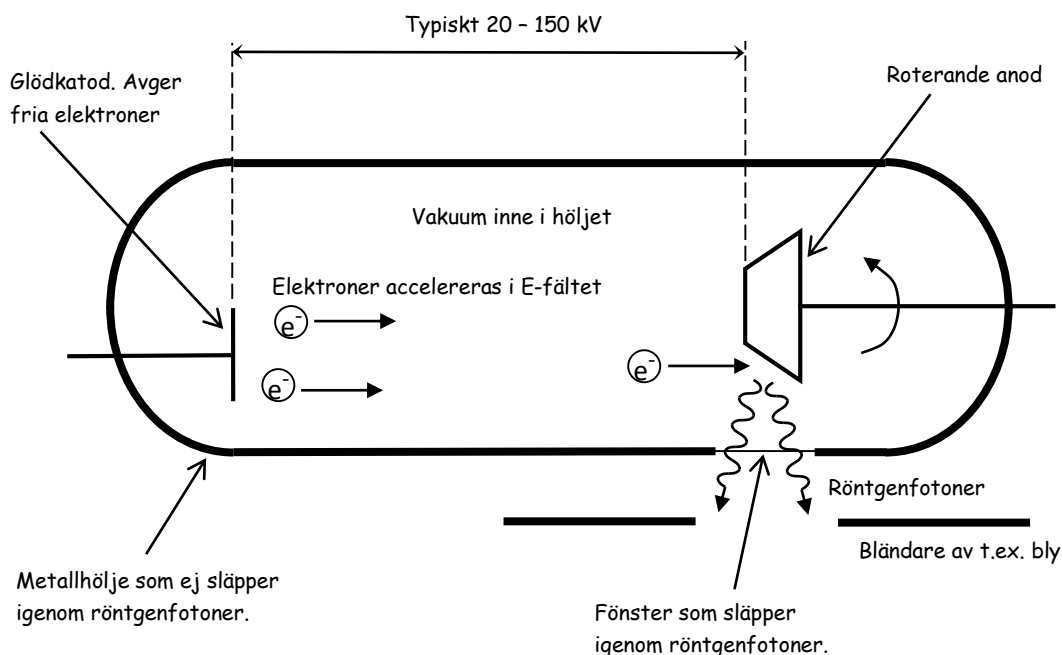


Fig. 20.1. Principen för ett röntgenrör.

I röntgenröret avges fria elektroner från en glödkatod. Dessa elektroner accelereras av en hög spänning mot anoden som vanligen består av någon tung och värmetålig metall, t.ex. volfram. När elektronerna hastigt bromsas in i anodmaterialet omvandlas den kinetiska energin delvis till röntgenstrålning, så kallad bromsstrålning, som ger upphov till ett kontinuerligt spektrum. Dessutom kan de

infallande elektronerna slå loss elektroner i inre skal i anodmaterialalets atomer. De luckor som på detta sätt uppstår i de inre skalerna utfylls snabbt genom att elektroner faller ner från högre liggande skal, varvid överskottsenergin utsänds i form av röntgenfotoner. På detta sätt uppstår ett linjespektrum som överlagras på det kontinuerliga bromsstrålningsspektrat.

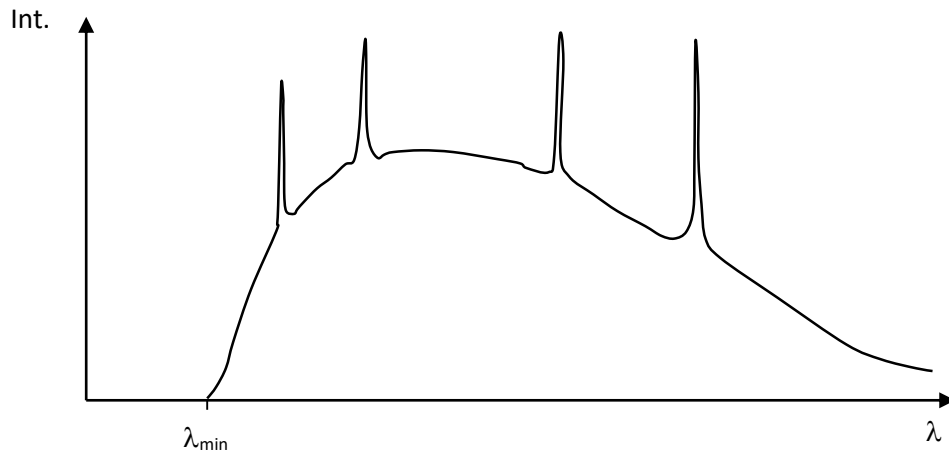


Fig. 20.2. Spektrum för ett röntgenrör.

$\lambda_{\min}$ svarar mot att elektronens hela kinetiska energi övergår till en enda röntgenfoton när den bromsas upp i anoden. Vi får då att

$$e \cdot U = \frac{hc}{\lambda_{\min}} \Rightarrow \lambda_{\min} = \frac{hc}{eU}, \quad (20.1)$$

där h är Plancks konstant, c är ljushastigheten, e är elektronladdningen och U är spänningen över röret. Exempelvis ger $U = 100$ kV ett $\lambda_{\min}$ -värde av ca. 0.01 nm. Vid medicinsk röntgen används vanligen våglängder av 0.01 – 0.1 nm. Långa våglängder (mjukröntgen) används vid fotografering av tunna och/eller svagt absorberande motiv (t.ex. insekter och andra små djur). Korta våglängder (hårdröntgen) används vid studier av tjocka och/eller täta material (t.ex. brobalkar). I första fallet använder man låg spänning över röntgenröret (några kilovolt), och i senare fallet används hög spänning (upp till flera megavolt). För att få bra skärpa vid röntgenfotografering är det viktigt att elektronerna träffar anoden på en så liten yta som möjligt, så att man får en strålningskälla som är så nära punktformig som möjligt (anoden roterar snabbt för att inte denna punkt ska hettas upp för mycket). Kravet på punktformighet beror på att man vanligen inte har någon avbildande optik vid röntgenfotografering, utan man fotograferar helt enkelt motivets skuggbild. Ju mindre strålningskällan är, desto skarpare blir skuggbilden.

20.3. Uppställning för röntgenfotografering

Det är mycket svårt att bygga optik som fungerar i röntgenområdet. De flesta material har till exempel ett brytningsindex som ligger mycket nära ett, vilket gör det svårt att konstruera linser. Spegeloptik vore ett tänkbart alternativ, men problemet är att reflektionsförmågan är mycket dålig för praktiskt taget alla material. Därför används vanligen "skuggfotografering" enligt Fig. 20.3. En nackdel med detta är att sensorn måste vara minst lika stor som den motivdel som ska avbildas, vilket ger stora problem om man ska röntgenundersöka t.ex. en elefant.

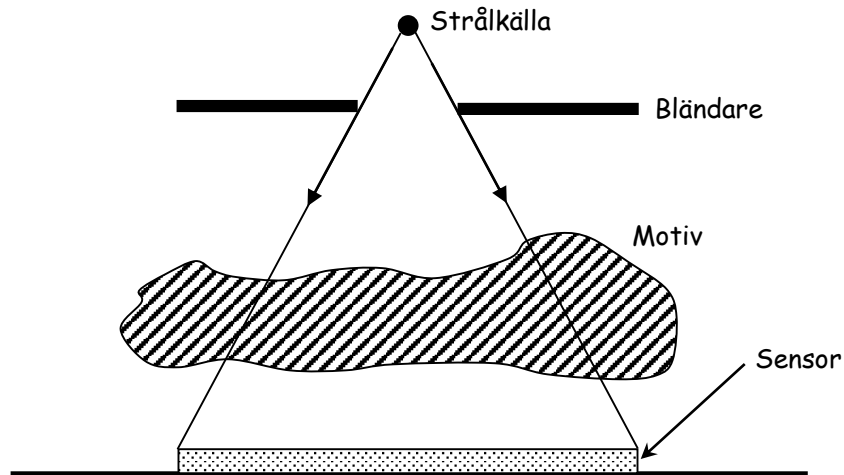


Fig. 20.3. Uppställning för röntgenfotografering.

Det finns dock en variant av röntgenundersökning som har den fördelen att man kan använda en vanlig kamera även vid fotografering av stora motiv. Denna metod kallas skärmbildsfotografering, Fig. 20.4.

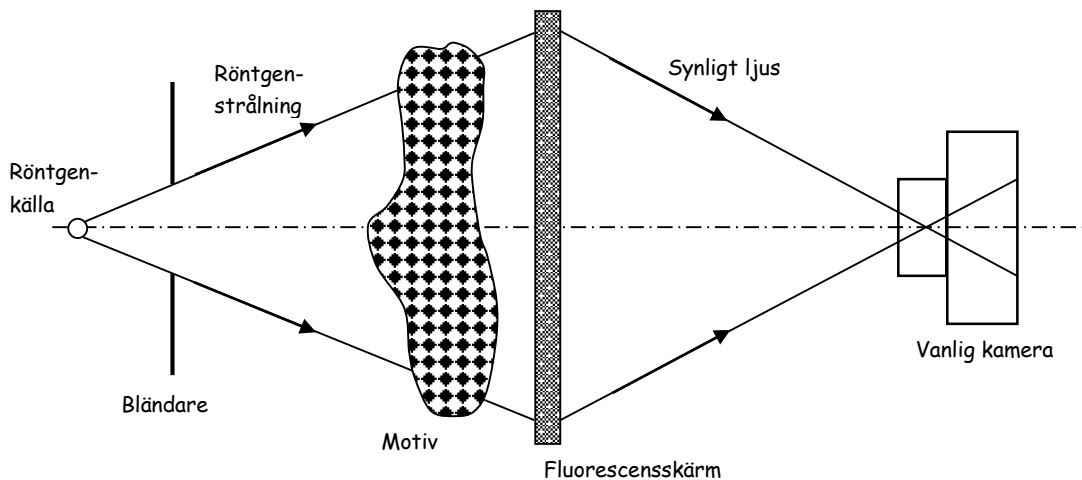


Fig. 20.4. Uppställning för skärmbildsfotografering.

Röntgenstrålningen genomtränger motivet och träffar en fluorescensskärm, som därvid utsänder synligt ljus. Denna skärm avfotograferas med en vanlig kamera. Man kan på detta sätt få en valfri avbildningsskala. Fluorescensljuset är ofta svagt, vilket gör att ljusstarka objekt och högt ISO-tal bör användas.

20.4. Sensorer för röntgenstrålning

Fotografisk film har använts för att detektera röntgenstrålning alltsedan Wilhelm Röntgen utförde sina första experiment 1895. Film har dock många nackdelar, speciellt vad gäller medicinsk röntgen. Den totala detektionseffektiviteten är begränsad, vilket gör att man ofta tvingas använda högre stråldoser än vad som skulle vara önskvärt. Andra nackdelar med användande av film är att den är dyr, samt kräver kemisk framkallning. Den färdigframkallade filmen kräver dessutom stort lagringsutrymme, samt kräver en omständlig scanningprocedur om den ska överföras till datorkompatibel form. Allt detta har gjort att man precis som för vanlig fotografi har övergått till elektroniska matrissensorer. Dentala sektorn var tidigast ute, eftersom man där kan nöja sig med betydligt mindre storlek på sensorn än för medicinsk röntgen.

De CCD- eller CMOS-sensorer som används för dentalröntgen fungerar efter samma princip som de som används i digital fotografi, se kap. 14. På grund av de höga fotonenergierna i röntgenstrålning är förutsättningarna goda att skapa laddningsbärare i sensorn. Emellertid så medför de höga fotonenergierna att en hel del fotoner passerar genom sensorn utan att absorberas (den höga penetrationsförmågan är ju något som utnyttjas i röntgenundersökningar), och vidare kan den höga fotonenergin skada sensorn. För att åtgärda bägge dessa saker använder man ett fluorescerande skikt (ofta kallat scintillator) som absorberar röntgenfotoner och omvandlar dessa till synligt ljus. Typiska data för en dental sensor kan vara:

- Aktiv sensoryta 20 mm x 30 mm.
- Pixelstorlek 25 μm x 25 μm (ca. 1 Mpixel totalt).
- Upplösningsförmåga ca. 20 linjepar/mm.

När man behöver större röntgendetektorer än vad som är möjligt med CCD-eller CMOS-matriser kan man använda så kallade flatpanelsdetektorer. Dessa har typiskt en storlek av ca. 40 cm x 40 cm, och lämpar sig därför till medicinsk röntgen. Precis som CCD- och CMOS-sensorer, så har flatpanelsdetektorn den fördelen gentemot film att man drastiskt kan reducera stråldosen som patienten utsätts för. Flatpanelsdetektorn har också en mycket högre dynamik än filmen. Detta innebär att den kan registrera ett större exponeringsomfång, och att den därför är mindre känslig för felexponering. Detta är egenskaper som är mycket värdefulla, speciellt i medicinsk röntgen.

Flatpanelsdetektorn liknar till uppbyggnad och funktionssätt en CMOS-sensor av den typ som används i digitalkameror. Det är alltså en detektormatris, där varje pixel består av en fotodiod plus elektroniska komponenter för signalutläsning. Varje pixel är dock betydligt större än i en digitalkamera, ett typiskt värde kan vara drygt 100 μm kantlängd. Precis som för en CMOS-sensor, kan pixlarna i en

flatpanelsdetektor individuellt adresseras och läsas ut till en dator. Denna möjlighet att läsa ut bara delar av totala datamängden, gör att man kan få en hög bildfrekvens vilket ger möjlighet att spela in rörliga röntgenbilder.

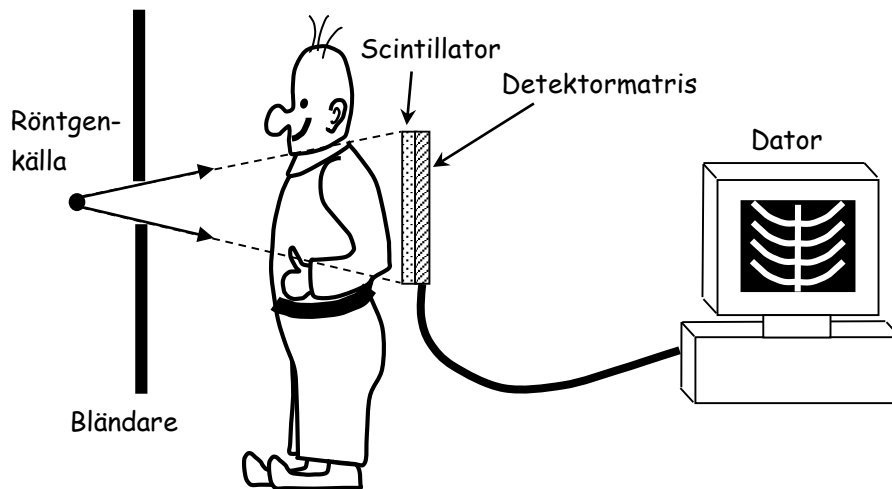


Fig. 20.5. Flatpanelsdetektorer har idag ersatt röntgenfilm i medicinska tillämpningar. Röntgenstrålningen omvandlas av ett scintillatorskikt till synligt ljus som detekteras av en stor detektormatris baserad på amorft kisel.

En flatpanelsdetektor skiljer sig från en CMOS-sensor på så sätt att man inte använder kristallint kisel, eftersom detta är omöjligt på grund av storleken. Istället använder man amorft kisel, där atomerna inte ligger i någon regelbunden kristallin struktur utan är mera oordnade. Strukturen liknar mycket den som finns i TFT (Thin Film Transistor) displayer på laptop-datorer. Precis som för CCD-matriser för tandröntgen, så innehåller en flatpanelsdetektor ett scintillator-lager som omvandlar röntgenstrålning till synlig strålning. Detta ger en kraftig ökning i känsligheten, vilket medför att stråldosen till patienten kan minskas.



Röntgenfoto: Medicinsk Röntgen, Läkarhuset, Odenplan, Stockholm.

Fig. 20.6. Röntgenfoto som visar författarens nackparti samt övre delen av ryggraden (sidovy). Man ser tydligt skruvarna som håller huvudet på plats!

Skruvar och andra metalledar är gjorda av titan, vilket har en högre absorption av röntgen-strålning än omgivande ben och vävnad. De ser därför ljusa ut på denna bild som motsvarar ett fotografiskt negativ, dvs ju högre exponering desto mera svärtning.

21. Höghastighetsfotografi

21.1. Inledning

Många gånger finns det behov av att kunna avbilda mycket snabba förlopp för att studera vad det är som egentligen händer. Ofta finns också ett behov av att kvantifiera informationen genom att mäta upp positioner, hastigheter mm. Exempel på detta är krockstudier med bilar, explosioner, vibrationer mm. Beroende på vilken typ av information man är ute efter kan man utnyttja två principiellt olika metoder. Är man ute efter att i detalj studera dynamiken i ett förlopp krävs att man studerar ett stort antal bilder registrerade i mycket snabb följd, Fig. 21.1. Detta kan uppnås med höghastighetsvideo. Om man däremot är nöjd med bara en ögonblicksbild av ett förlopp, räcker det med en bild tagen i rätt ögonblick och med mycket kort exponeringstid. Alternativt gör man ibland multipla korta exponeringar på en och samma bildruta, Fig. 21.2. Denna senare variant kan åstadkommas med blyxt- eller strobofoto om blyxttiden är tillräckligt kort.



Image courtesy of The Cooke Corporation, www.cookecorp.com.

Fig. 21.1. Bildsekvens som visar hur en kula lämnar mynningen på ett vapen. Sekvensen har tagits med en videokamera som tillåter upp till 10 miljoner bilder per sekund.



Image courtesy of The Cooke Corporation, www.cookecorp.com.

Fig. 21.2. Fotografi som visar hur en gevärskula med hastigheten 800 m/s genomtränger ett äpple. Bilden togs med en vanlig kamera vars slutare var öppen under hela exponeringsförloppet. Som ljuskälla användes ett blyxtaggregat som gav ca. $0.5 \mu\text{s}$ långa blyxtar som upprepades med ca. $100 \mu\text{s}$ tidsintervall.

21.2. Höghastighetsvideo

Gränsen mellan videokameror (för rörliga bilder) och vanliga digitalkameror är idag ganska utsuddad, eftersom de flesta digitalkameror också kan spela in videosekvenser. Detta är möjligt eftersom det i grunden är samma komponenter som ingår i bägge kameratyperna. Vid videoinspelningar använder man dock färre pixlar per bildruta, eftersom antalet pixlar som kan läsas ut per sekund från en kamera är begränsat. Fler bilder per sekund betyder alltså färre pixlar per bild. Genom att använda CMOS-sensorer kan man lätt ändra antalet pixlar i en utläst bild eftersom pixlarna är individuellt adresserbara. I detta avseende har CMOS en stor fördel gentemot CCD, eftersom man i en CCD inte kan läsa ut delbilder av totala antalet pixlar på samma sätt. CMOS-kretsar är också generellt sett snabbare än CCD. Förutsättningarna att uppnå höga bildfrekvenser är alltså gynnsamma vid användande av CMOS-teknik. För höghastighetsvideo är det huvudsakligen två faktorer som begränsar bildfrekvensen:

- 1) *Hastigheten med vilken pixeldata kan läsas ut ur elektroniska sensorn och lagras i ett minne.*
- 2) *Att detektera tillräckligt många fotoner eftersom exponeringstiden per bild blir mycket kort.*

Vad gäller den första punkten kan här ges ett exempel på vad en typisk kommersiell kamera kan prestera. Den kan ha en (i höghastighetssammanhang) högupplösande CMOS-sensor med 1.5 Mpixel, och klara av en maximal bildfrekvens av 5000 s^{-1} . Bilddata lagras upp i ett lokalt minne i kameran, vilket gör att bara 6 sekunders inspelningstid klaras av vid maximal bildfrekvens. Ofta räcker det med mindre än 1.5 Mpixel per bild, och då kan bildfrekvensen ökas eftersom det är lagringshastigheten av data som begränsar. Om man bara tittar på ett delutsnitt bestående av 320×200 pixlar kan man sålunda komma upp i 50 000 bilder per sekund. Om man drastiskt reducerar totala antalet registrerade bilder, kan man i vissa kameror komma upp i betydligt högre bildfrekvens. Som exempel kan nämnas att det finns en kommersiell kamera som klarar av att registrera 3.3×10^8 bilder per sekund, där varje bild består av 1.3 Mpixel. Haken med denna höga hastighet är att totalt bara fyra bilder kan registreras. En sådan kamera platsar därför inte riktigt i kategorin höghastighetsvideo, utan vi återkommer till den när vi tittar på teknik för att registrera enstaka ögonblicksbilder.

Under punkt 2 ovan nämns problemet med att hinna registrera tillräckligt många fotoner vid höga bildfrekvenser. Om man till exempel använder en bildfrekvens av $1.0 \times 10^5 \text{ s}^{-1}$ så blir den maximala exponeringstiden per bild $10 \text{ } \mu\text{s}$. Detta är två till tre tiopotenser kortare än vad som vanligen används i traditionell fotografi. Som en konsekvens av detta krävs att motivet belyses mycket kraftigare. Låt oss göra ett överslag av hur kraftig belysning av motivet som krävs. Eftersom data och prestanda för både kameror och ljuskällor kan variera en hel del så betonas att beräkningen bara är ett överslag för att bestämma storleksordningen.

Räkneexempel 1 (överkurs):

Om man tittar i databladet för en viss videokamera kan man se följande data:

Varje detektorelement (pixel) har en storlek av $6.7 \mu\text{m} \times 6.7 \mu\text{m}$. Detektorn har en kvantverkningsgrad av 40%, och elementen bottenar när 25 000 fotoner har detekterats.

Vi antar att ljuskällan är lite "rödförskjuten" relativt dagsljus så att medelvåglängden kan sättas till 600 nm. Detta innebär att förhållandet mellan ljusflöde och strålningseffekt kan sättas till 400 lumen/W (fås ur tekniska uppslagsverk).

Vi antar att motivet har en medelreflektans av 18% (används ofta som typiskt värde inom fotografi).

Vi vill att i medeltal 10 000 fotoner ska registreras per pixel (svarande mot medelgrå nivå).

Antag att objektivet arbetar vid bländartal $F = 1.4$.

Lösning:

För att få 10 000 detekterade fotoner på en pixel krävs att $\frac{10000}{0.4} = 25000$

fotoner träffar pixeln under exponeringstiden. Detta motsvarar en instrålad effekt på en pixel av

$$\frac{25000 \times hc}{\lambda \times t_{\text{exp}}} = \frac{25000 \times 6.63 \times 10^{-34} \times 3.0 \times 10^8}{600 \times 10^{-9} \times 1.0 \times 10^{-5}} = 8.3 \times 10^{-10} \text{ W.}$$
 Irradiansen i

detektorplanet blir då $\frac{8.3 \times 10^{-10}}{(6.7 \times 10^{-6})^2} = 18.5 \text{ W/m}^2$. Eftersom vi antagit att en

watt svarar mot 400 lumen så blir belysningen i detektorplanet $E = 18.5 \times 400 = 7.4 \times 10^3 \text{ lumen/m}^2$ (lux). Förhållandet mellan motivluminans, L ,

och belysning i objektivets bildplan, E , ges av ekv. 9.9: $E = \frac{L\pi}{4F^2}$. Med ett

bländartal av 1.4, och $E = 7.4 \times 10^3 \text{ lux}$, får vi att motivluminansen måste vara

$$L = \frac{4 \times (1.4)^2 \times 7.4 \times 10^3}{\pi} = 1.84 \times 10^4 \text{ cd/m}^2.$$
 Om vi antar att motivet diffust

reflekterar infallande belysning, Fig. 9.5, samt har en medelreflektans av $R = 18\%$, får vi att belysningen av motivet, E_{motiv} , måste vara

$$E_{\text{motiv}} = \frac{L \times \pi}{R} = 3.2 \times 10^5 \text{ lux}$$
 (här har gjorts en integrering över en total

rymdvinkel 2π). På grund av alla approximationer som gjorts bör vi inte svara med större noggrannhet än att belysningen troligen bör ligga i området $10^5 - 10^6 \text{ lux}$. Detta är en extremt hög belysning. Som jämförelse kan nämnas att vanlig rumsbelysning kanske uppgår till hundra, eller några hundra, lux. Belysningen utomhus en solig dag på sommaren uppgår maximalt till ca. 10^5 lux , och detta räcker alltså knappt till för höghastighetsvideo. Det kan nämnas att det finns speciella spotlights för höghastighetsfoto som inom ett litet område (ca. 20 cm diameter) kan ge en belysning av 10^5 lux , men en sådan spotlight drar flera kilowatt elektrisk effekt och sprider mycket värme. Att

uppnå tillräcklig belysning är alltså ofta ett stort problem för upptagningar med höghastighetsvideo.

21.3. Enkelbild och strobofotografi

Såsom tidigare nämnts behöver man i många höghastighetstillämpningar inte använda en höghastighetsvideo. Detta är fallet när det räcker med en enda ögonblicksbild, eller en multipelexponering på samma "bildruta". För detta behövs i princip ingen speciell kamera, förutsatt att man har en pulsad ljuskälla med mycket kort brinntid. Oftast används speciella fotoblixtar för detta syfte, men i vetenskapliga sammanhang används ibland också pulsade lasrar. Det är alltså i dessa fall ljuskällan, och inte kameran, som bestämmer exponeringstiden. Genom att låta ljuskällan blinka upprepade gånger (stroboskop) under exponering av en och samma bild får man multipelexponeringar. Detta kan vara användbart till exempel för att studera en projektil som flyger förbi kamerans synfält. Kraven för att lyckas med denna typ av fotografi är att man har en ljuskälla med tillräckligt kort brinntid, samtidigt som intensiteten måste vara hög. Dessutom måste den tändas i rätt ögonblick, vilket ofta kräver rörelsedetektorer i kombination med speciell elektronik. För att få korta, intensiva blixtar kan man använda speciella blyxtaggregat som utvecklats för denna typ av foto. Exempel på prestanda för ett sådant aggregat kan vara blyxttider från 5 till 200 μs , och pulsenergier från 40 till 200 joule (dvs. effekter i megawatt-området). Om man kan nöja sig med pulsenergier på någon enstaka joule är det möjligt att komma ner till blyxttider på ca. 0.2 μs .

21.4. Extremt korta exponeringstider, bildförstärkare

För att få exponeringstider nere i nanosekundområdet krävs vanligtvis pulsade lasrar, eller att man har någon typ av ultrasnabb slutarfunktion i kameran. Ett intressant exempel på det senare är användande av en "gateable micro-channel plate", MCP, (bildförstärkare på svenska). Denna fungerar både som en snabb slutare och fotonförstärkare. Den sistnämnda funktionen är mycket fördelaktig, eftersom man ofta har problem med lågt fotonantal. Med en MCP kan man komma ner i exponeringstider på 3 ns eller lägre i kommersiellt tillgänglig utrustning. Låt oss titta på hur en MCP fungerar. I Fig. 21.3 a) ges en översiktsbild, medan Fig. 21.3 b) visar ett tvärsnitt som förklarar funktionen.

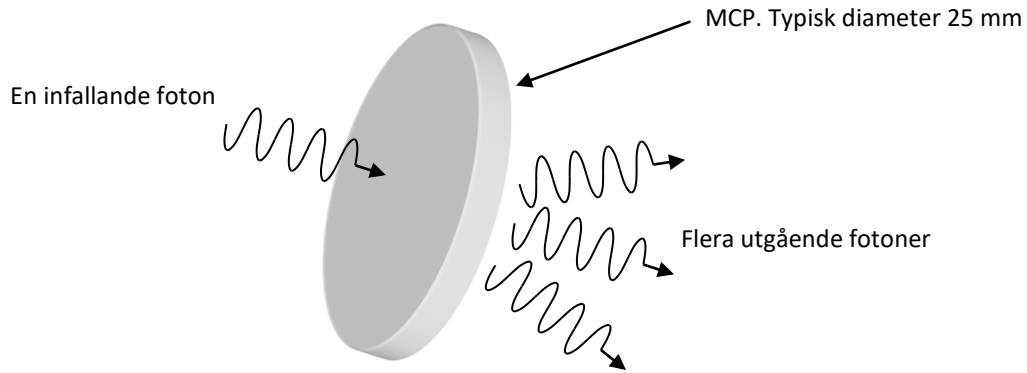


Fig. 21.3. a) Översiktsbild av en bildförstärkare av typ MCP (MicroChannel Plate).

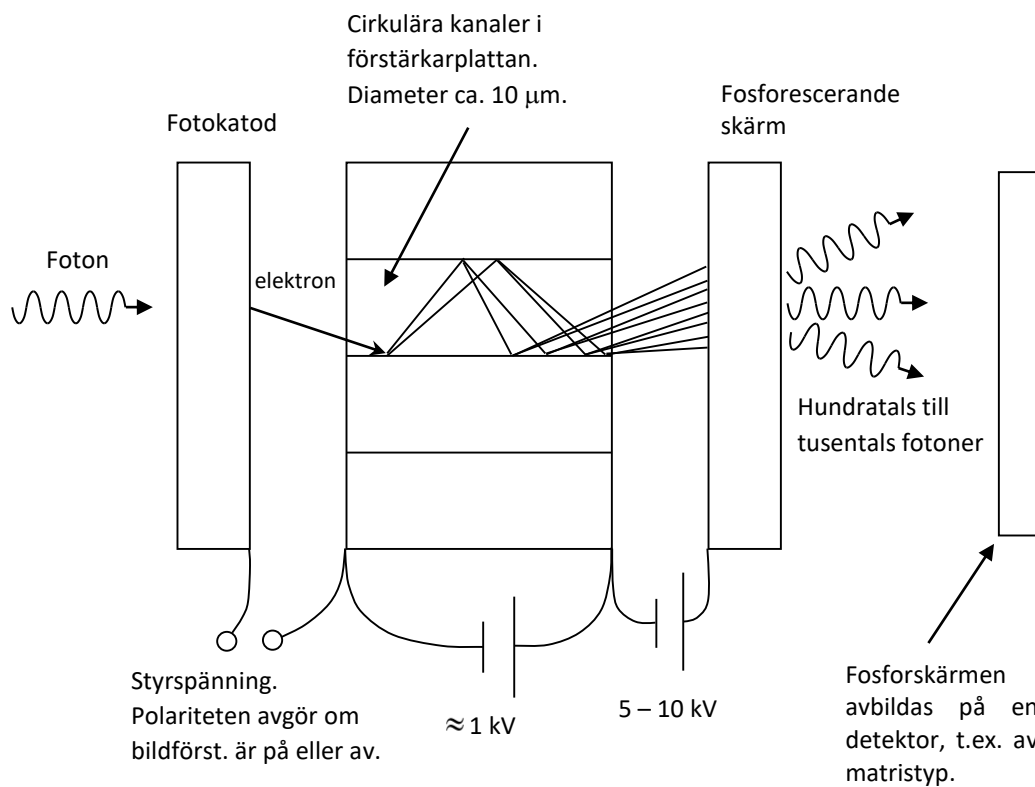


Fig. 21.3. b) Funktionsprincipen för en MCP.

Bildförstärkaren fungerar så att en infallande foton kan slå ut en elektron i fotokatoden. Om styrspänningen mellan fotokatoden och förstärkarplattan är sådan att fotokatoden är negativ i förhållande till plattan, kommer elektronen att accelereras mot plattan och slår emot väggen i en av de cirkulära kanalerna. Kanalväggen är belagd med ett material som har låg utträdesenergi för elektroner, vilket gör att den infallande elektronen kan slå ut flera sekundära elektroner (precis som i en fotomultiplikator). Inne i kanalen finns ett elektriskt fält som accelererar

de frigjorda elektronerna åt höger. Dessa elektroner kommer att med hög hastighet slå i väggen på nytt, och kan då i sin tur slå ut flera nya elektroner var. Denna process upprepas och antalet elektroner växer lavinartat. Ett stort antal elektroner kommer ut från kanalens högra sida och accelereras av en hög spänning mot en fosforescerande skärm där de ger upphov till fotoner. På detta sätt kan en enda foton förstärkas till att bli flera tusen. Detta är mycket användbart i situationer där man har ont om ljus, men det ska observeras att en hög ljusförstärkning även ger ett högt brus. En annan fördel med bildförstärkaren är att den reagerar mycket snabbt på ändringar i styrspänningen mellan fotokatod och förstärkarplatta. På detta sätt får man samma funktion som med en extremt snabb slutare, vilken kan ge exponeringstider ner till 3 ns eller ännu lägre. Detta är mycket användbart för att minska rörelseoskärpan vid höghastighetsfotografi. Den tidigare nämnda kameran som klarar av att registrera fyra bilder med frekvensen 3.3×10^8 bilder per sekund använder just bildförstärkare av den typ som beskrivits. Egentligen består denna kamera av fyra ”stillbildskameror”, var och en med sin egen bildförstärkare och efterföljande matrisdetektor. Dessa kameror exponeras i snabb följd, men att sedan läsa ut bilddata från matrisdetektorerna tar mycket längre tid. Detta är orsaken till att bara fyra bilder kan registreras i snabb följd.

21.5. Försöksplanering, exempel

Vid alla tillämpningar av höghastighetsfotografi måste man bestämma dels vilken exponeringstid man kan tillåta, dels hur lång tid som ska förflyta mellan successiva exponeringar (såvida det inte räcker med en enda ögonblicksbild). Detta kräver ingen avancerad matematik, och ofta räcker det med överslagsräkningar. Ett tänkt exempel finns beskrivet nedan.

Räkneexempel 2:

Antag att vi vill göra en multipelexponerad bild som visar från sidan hur en gevärskula träffar och genomtränger en tunn plåt. Vi använder en kamera med matrisdetektor med en area av 8.6 mm x 7.2 mm, bestående av 1200 x 1000 pixlar. Kulans hastighet är 800 m/s, och den är 20 mm lång. Vi önskar en bild med utseende enligt Fig. 21.4, och rörelseoskärpan måste vara försumbar. Hur kort exponeringstid krävs, och hur lång tid ska förflyta mellan successiva exponeringar?

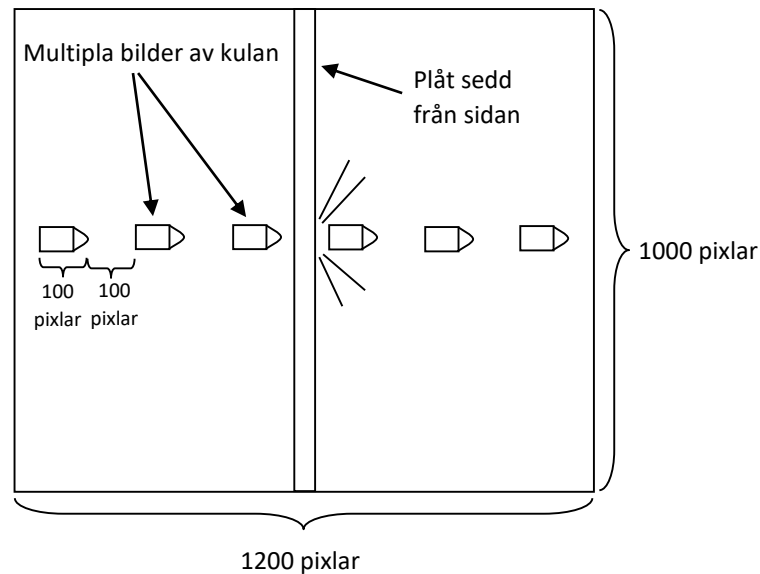


Fig. 21.4. Önskad bild som genom multiplexponering visar hur en gevärskula genomtränger en plåt.

Lösning:

I den optiska bilden som faller på detektormatrisen ska kulans längd vara $\frac{100}{1200} \times 8.6 = 0.72$ mm. Detta ger en avbildningsskala av $\frac{0.72}{20} = 0.036$, vilket innebär att den optiska bilden av kulan rör sig med hastigheten $800 \times 0.036 = 28.7$ m/s över detektormatrisen. En försumbar rörelseoskärpa får vi om bilden flyttar sig maximalt en pixelbredd under exponeringstiden. En pixelbredd är $\frac{8.6}{1200} = 0.0072$ mm = $7.2 \mu\text{m}$. Tiden som det tar för bilden

att flytta sig den sträckan blir $t_{\text{exp}} = \frac{7.2 \times 10^{-6}}{28.7} = 2.5 \times 10^{-7} \text{ s} = 250 \text{ ns}$. Tiden

mellan exponeringarna ska vara 200 gånger så lång (motsvarande en förflyttning av 200 pixlar), vilket blir $50 \mu\text{s}$. Detta klaras lätt av en kamera med en bildförstärkare av den typ som beskrivits.

Vid så här korta exponeringstider kan man naturligtvis fråga sig om en kontinuerlig ljuskälla kan ge tillräckligt hög belysning av motivet, eller om man tvingas använda en pulsad ljuskälla. I det tidigare räknade exemplet med höghastighetsvideo fann vi att med en exponeringstid av $10 \mu\text{s}$ så krävdes en motivbelysning i storleksordningen $10^5 - 10^6$ lux, vilket motsvarar upp till 10 gånger fullt dagsljus. Här har vi en exponeringstid som är 40 gånger kortare, vilket skulle innebära 40 gånger högre belysning. Situationen räddas emellertid om vi har en kamera med bildförstärkare. Med ljusstark optik eller fibermatning mellan fosforskärm och detektormatris kan man då klara sig med lägre motivbelysningar än i fallet med höghastighetsvideo. Nackdelen med bildförstärkare är, såsom tidigare nämnts, att bruset ökar samt, beroende på fosforskärmens egenskaper, att bildskärpan kan bli sämre.

22. Sampling

22.1. Inledning

För att lagra en bild i digital form måste den ”samplas”. Detta brukar förklaras som att ljusvärden bara registreras för vissa punkter i den optiska bilden, och vad som händer däremellan känner man inte till. I verkligheten kan man naturligtvis inte göra några perfekta punktmätningar, utan vad som mäts upp är medelvärdet över en liten area, t.ex. en enskild pixel i en CCD- eller CMOS-sensor. Detta påverkar emellertid inte de samplingregler, felaktigheter etc. som presenteras i detta kapitel. Enda effekten av mätytans storlek blir en kontrastminskning för högre ortsfrekvenser, vilket illustreras i exempel 2 i detta kapitel. Vid användning av en matrissensor ligger samplingen så att säga inbyggd i själva sensorn – samplingpositionerna ges av detektorelementens lägen. Ju färre pixelvärden som registreras, desto grövre och kantigare kommer bilden att se ut, men den upptar mindre lagringsutrymme. I detta kapitel ska vi titta på vilka regler som gäller för sampling, och vad som händer om man bryter mot dessa regler.

22.2. Ett exempel

Låt oss titta på ett enkelt, endimensionellt exempel på punktsampling. Vi antar att vi har en sinusformigt varierande belysning längs x-axeln, samt att vi har frihet att välja hur tätt vi vill sampla funktionen. Man har väl en intuitiv känsla av att ju tätare man samplar, desto mer information får man om den kontinuerliga originalfunktionen. Stämmer detta? En intressant fråga är också vilka konstigheter och felaktigheter som dyker upp om man samplar glest. Slutligen är det intressant att veta om man vid tillräckligt tät sampling kan skaffa sig så mycket information att man exakt kan återskapa belysningsfördelningen i den ursprungliga optiska bilden. Vi börjar med ett praktiskt experiment, där vi helt enkelt samplar sinusfunktionen med tre olika samplingtätheter enligt Fig. 22.1. Täthet 1 innebär ”många” samplingpunkter per period. Täthet 2 är precis två samplingpunkter per period (men dom behöver naturligtvis inte ligga på toppen och i botten som i figuren). Täthet 3 innebär mindre än två samplingpunkter per period.

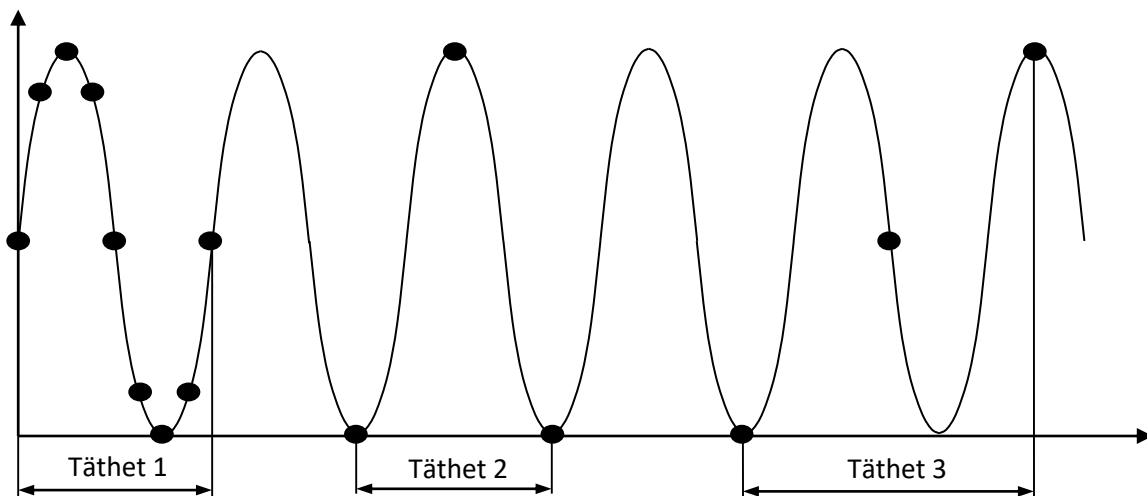


Fig. 22.1. Sampling av en periodisk funktion. Tre fall med olika samplingtätheter illustreras.

Antag att vi försöker göra en enkel rekonstruktion av vår originalkurva utgående från samplade värden. Detta gör vi genom att helt enkelt sammanbinda samplingpunkterna med rätta linjer (linjär interpolation). Resultaten för de tre samplingtätheterna illustreras i Fig. 22.2.

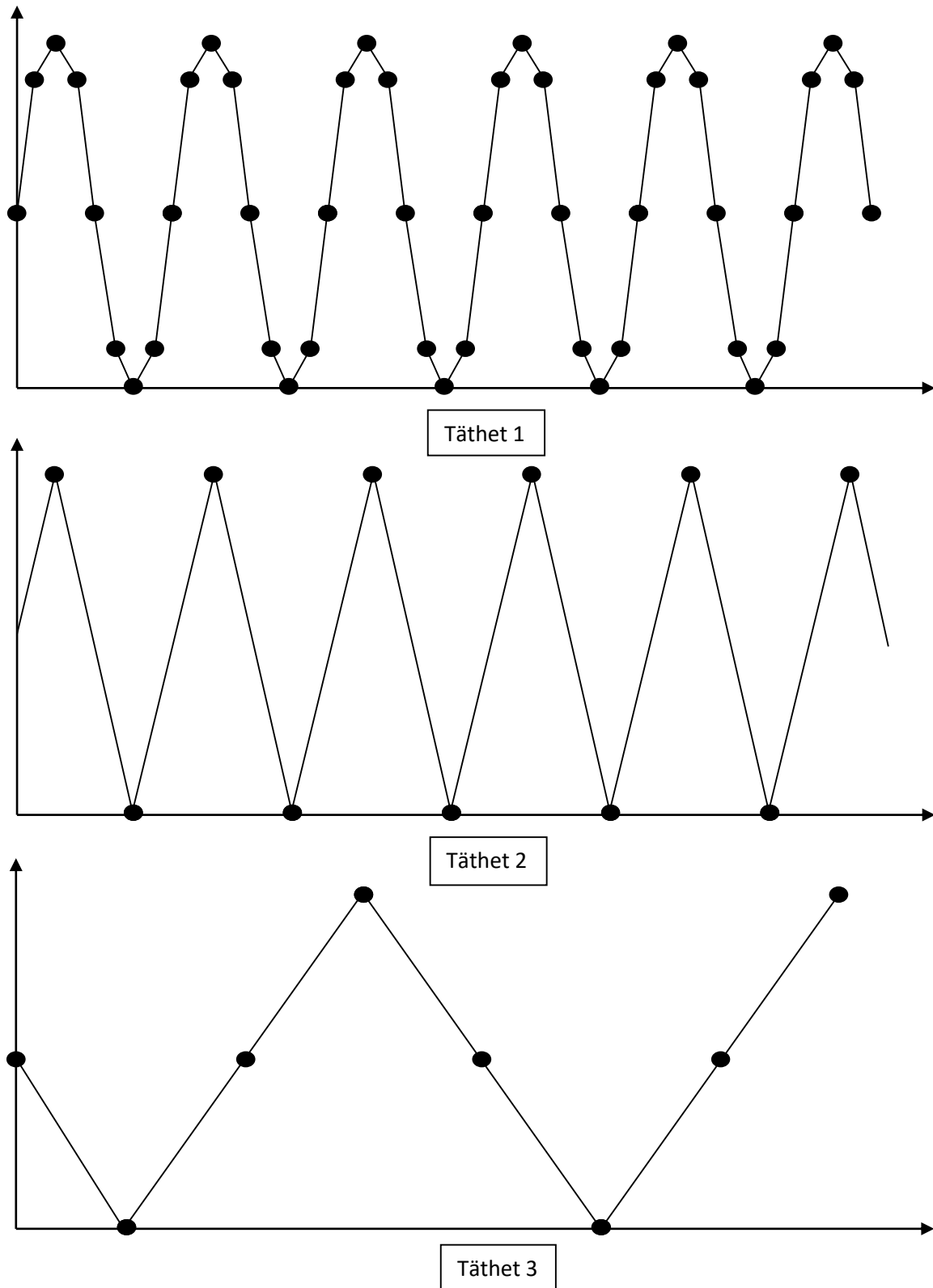


Fig. 22.2. Försök till rekonstruktion av originalfunktionen med linjär interpolation.

Inte i något fall får vi en funktion som ser ut som originalet (sinus). Mest likt blir (föga förvånande) fallet 1. I fallet 2 får vi en triangelvåg vars ortsfrekvens överensstämmer med originalfunktionens. Även i fall 3 får vi en triangelvåg, men med felaktig ortsfrekvens. Resultatet visar alltså att ju glesare vi samplar, desto sämre överensstämmer resultatet med originalfunktionen. Man skulle därför vara frestad att dra slutsatsen att för att få ett bra resultat ska man sampla så tätt som möjligt, och endast vid oändligt tät sampling kan man exakt återskapa originalet. Detta är emellertid fel.

22.3. Samplingkriteriet

Man kan visa att i ovanstående exempel så är ett tillräckligt villkor att avståndet mellan samplingpunkterna är mindre än halva perioden. Gränsfallet går alltså vid täthet 2 ovan. Om detta är uppfyllt kan man ur samplade data återskapa den kontinuerliga originalfunktionen helt och hållet. Detta kan tyckas motsägas av resultaten ovan, men ett korrekt återskapande kräver lite mer raffinerade metoder än den linjärinterpolation som gjordes (den intresserade hänvisas till K. Carlsson, Imaging Physics, KTH 2016).

I en verklig bild finns i allmänhet ett stort antal olika ortsfrekvenser representerade, dock inte hur höga som helst. I avsnittet om MTF såg vi att optiken har en högsta frekvens som den kan avbilda, och även andra komponenter kan verka begränsande i detta avseende. Det innebär att det i bilden finns en maximal ortsfrekvens som vi kan kalla $v_{\max}$. I detta fall gäller att om man samplar med en samplingfrekvens v_s (samplingfrekvensen = $1/\text{samplingavståndet}$) som är högre än $2v_{\max}$ så kan man helt och hållet återskapa originalbilden. Detta kallas samplingkriteriet och kan uttryckas i följande formel

$$v_s > 2v_{\max} \quad (22.1)$$

Exempel 1: För ett objektiv som används vid bländartal $F = 8$, så gäller att

optikens $v_{\max} = \frac{1}{\lambda \cdot F} \approx \frac{1}{550 \times 10^{-9} \cdot 8} = 2.27 \times 10^5 \text{ m}^{-1}$. Om vi i bildplanet

använder en samplingfrekvens $v_s > 2 \cdot 2.27 \times 10^5 = 4.55 \times 10^5 \text{ m}^{-1}$ så räcker det. Vi kan inte genom tätare sampling få mer information. För att inte behöva jobba med så stora tal så anges ortsfrekvenser och samplingfrekvenser ofta i sorten mm^{-1} . I detta fall får vi då villkoret $v_s > 455 \text{ mm}^{-1}$, vilket innebär ett avstånd av $2.2 \text{ }\mu\text{m}$ mellan samplingpunkterna.

I praktiken kan rekonstruktionen underlättas om man använder lite tätare sampling än vad som ges av samplingteoremet. Därför används i t.ex. ljud-CD 44 kHz samplingfrekvens trots att frekvenser över 20 kHz är ointressanta i ljudsammanslagning.

22.4. Aliasing (sv. vikning)

Om man bryter mot samplingkriteriet och samplar för glest, vad händer då? Jo, detta har vi redan sett ett exempel på för samplingtätet 3 i Fig. 22.2. Den ur samplingvärdena rekonstruerade funktionen kommer då att få en för låg ortsfrekvens (detta gäller också för mer sofistikerade rekonstruktionsalgoritmer än linjärinterpolation). Detta fenomen kallas på engelska för "aliasing" och på svenska "vikning". Det brukar ibland också benämnas moiré-mönster. Exempel på detta kan man ofta se vid återgivande av täta linje- eller rutnät, typ fiskbensmönstrade kostymer (när man ser detta fenomen på en TV- eller datorskärm kan det också bero på tätheten av färgpunkter på skärmen). Aliasingfenomen gör ofta inte så mycket i vanliga TV-sammanhang eller i amatörfotografi. Om man däremot vill använda bilderna kvantitativt för t.ex. fjärranalys eller som kartunderlag kan aliasing vara mycket besvärande. Det kan vara svårt eller omöjligt att upptäcka felet, och det kan grovt förvanska geometriska strukturer i bilderna. Speciellt problematiskt är aliasing när man använder rad- eller matrissensorer. Här har man inte något val eftersom samplingtäteten finns inbyggd i själva sensorn, och dessutom ligger den ofta på ett för lågt värde vilket framgår av exemplet nedan.

Exempel 2: Antag att vi har en sensor där de individuella elementen har en bredd av $5.0\ \mu\text{m}$ och ligger på ett centrum-till-centrum avstånd av $10\ \mu\text{m}$.

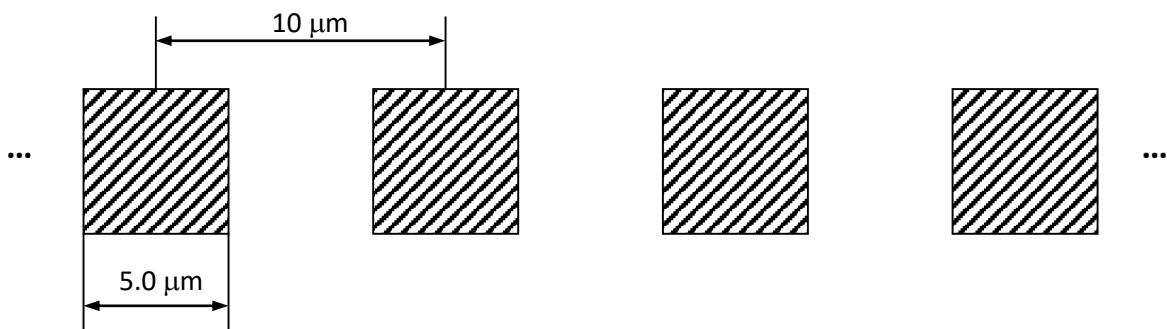


Fig. 22.3. Detektorelementens placering i en sensor.

Den bildregistrering som görs med denna sensor blir precis densamma som om vi hade ett enda detektorelement som flyttas $10\ \mu\text{m}$ i sidled mellan avläsningar. Samplingavståndet är alltså $10\ \mu\text{m}$, vilket svarar mot en samplingfrekvens av $100\ \text{mm}^{-1}$. Detta innebär enligt samplingteoremet att ortsfrekvenser som är $< 50\ \text{mm}^{-1}$ kan korrekt registreras. Fönsterbredden för ett detektorelement är $5.0\ \mu\text{m}$, vilket innebär att dess MTF kan approximeras

med (se kap. 17.6)
$$\text{MTF}(\nu) = \left| \frac{\sin(5.0 \times 10^{-6} \pi \nu)}{5.0 \times 10^{-6} \pi \nu} \right|.$$
 Denna MTF finns

illustrerad i Fig. 22.4. Som synes kan sensorn med god modulation registrera ortsfrekvenser långt högre än $50\ \text{mm}^{-1}$, vilka kommer att ge upphov till aliasing. Det är alltså mycket stor risk att råka ut för aliasing med denna typ av sensor.

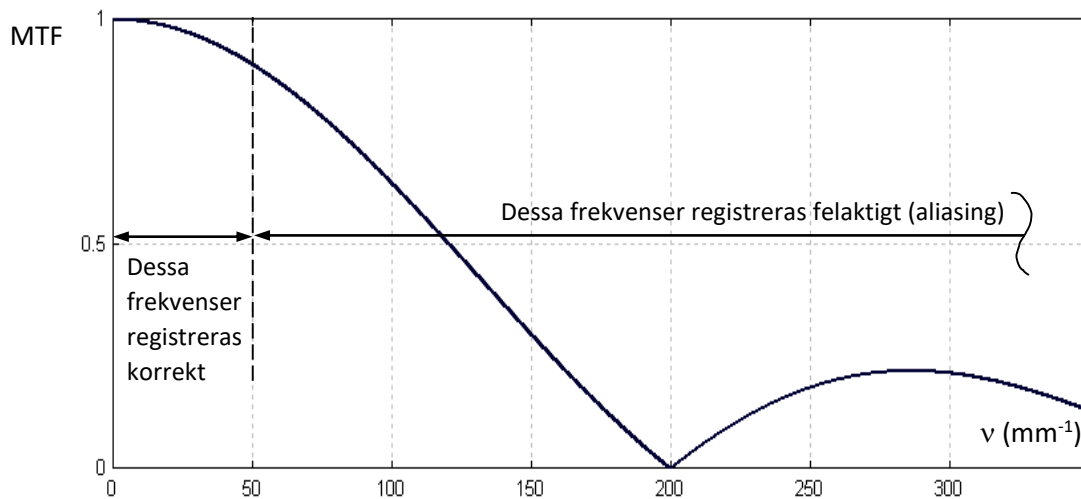


Fig. 22.4. MTF för sensorn i Fig. 22.3. Den kan med god kontrast återge ortsfrekvenser som är betydligt högre än de som kan samplas korrekt.

Ett mycket vanligt fall där man stöter på aliasing är när man ska fotografera en rasterad bild. Originalbildens rastertäthet är ofta ungefär lika stor som samplings-tätheten i den digitala bilden som registreras, vilket ger aliasing-problem av den typ som illustreras i Fig. 22.5. Rastringens höga ortsfrekvens har här genom aliasing omvandlats till en mycket lägre ortsfrekvens, vilket är mycket störande.

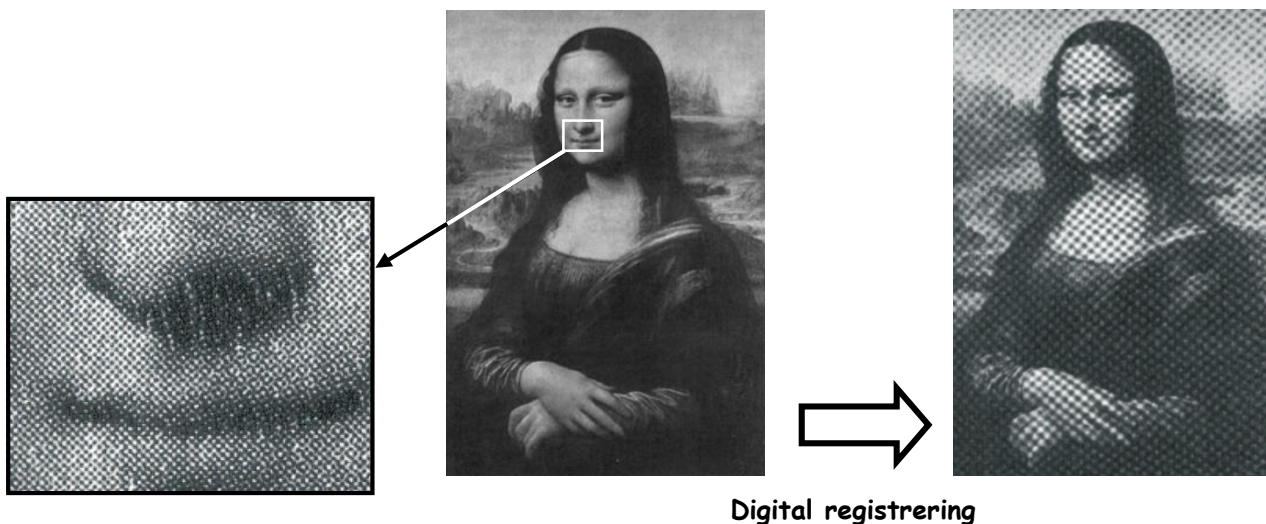


Fig. 22.5. Exempel på aliasing-fenomen, i vardagligt språk ofta kallat moiré-effekt. Denna effekt kan uppstå när man fotograferar en digital bild av ett rasterat original. Samplings-tätheten har i detta fall legat nära rastertätheten i originalbilden.

22.5. Antialiasing-filter

För att minska aliasing kan man använda ett så kallat anti-aliasing filter (även kallat optiskt lågpasfilter) som monteras precis framför sensormatrisen. Detta filter består av material som är optiskt dubbelbrytande, vilket resulterar i att man får en dubbelbild av motivet projicerad på sensorn. Förskjutningen mellan dubbelbilderna är ungefär lika stor som centrum-till-centrum avståndet mellan pixlarna, dvs. vanligen någon mikrometer. För att motverka aliasing i både horisontell och vertikal riktning använder man två filter lagda ovanpå varandra, och med förskjutningsriktningarna vinkelräta mot varandra. Resultatet blir då fyra överlagrade bilder, förskjutna både horisontellt och vertikalt. Effekten av ett anti-aliasing filter kan beskrivas med hjälp av MTF. Om vi tittar på ett endimensionellt fall så gäller att

$$MTF_{anti-alias} = |\cos(\pi v D)| \quad (22.2)$$

där v är ortsfrekvensen och D är förskjutningen mellan dubbelbilderna.

$MTF_{anti-alias}$ kommer att dra ner kamerans totala MTF till noll (eller i praktiken nära noll) ungefär vid den högsta ortsfrekvens som korrekt kan registreras, dvs 0.5 perioder/pixel* (denna frekvens brukar kallas Nyquistfrekvensen). Fig. 22.6 visar effekten som ett sådant anti-aliasing filter kan förväntas ha på kamerans MTF.

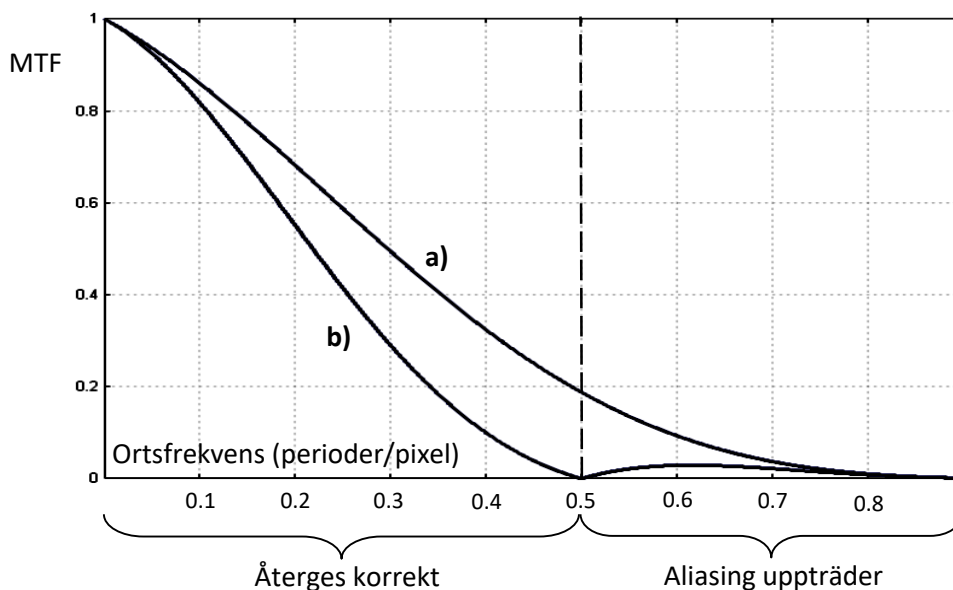


Fig. 22.6. Exempel på effekten av ett anti-aliasing filter. (Ortsfrekvensen anges här i perioder/pixel, se kap. 17.9) a) Typisk MTF för en digitalkamera inställd på bländartal 8. b) MTF för samma kamera som i a), men utrustad med anti-aliasing filter. För ortsfrekvenser över 0.5 perioder/pixel (Nyquistfrekvensen) uppträder aliasing (moiré-mönster). Effekten av filtret är att dessa frekvenser dämpas kraftigt så att aliasing praktiskt taget uteblir. Priset man får betala för detta är emellertid att MTF blir klart sämre också för ortsfrekvenser under 0.5 perioder/pixel.

* 0.5 perioder per pixel innebär att vi har ett linjemönster vars periodlängd = 2 gånger centrum-till-centrum avståndet mellan pixlarna. Detta innebär att linjemönstrets ortsfrekvens är hälften av samplingfrekvensen, vilket är gränsen när aliasing börjar uppstå.

Som synes i Fig. 22.6 kommer anti-aliasing filtret att dra ner MTF kraftigt även för frekvenser under Nyquistgränsen. För att minska denna icke önskvärda effekt, så lägger man ibland filtrets nollställe lite över Nyquistfrekvensen. Så verkar vara fallet i Fig. 22.7, som visar en uppmätt kurva på MTF_{sensor} tillsammans med teoretiska kurvor för en sensor med och utan anti-aliasing filter. Man kan se att den uppmätta kurvan relativt väl ansluter till den teoretiska kurvan för en sensor med anti-aliasing filter som har nollställe vid 0.65 perioder/pixel. Att den verkliga kurvan ligger lite lägre beror dels på den interpolation som görs för sensorer med Bayer-mosaik, dels på att sensorn inte är helt ideal (laddningsdiffusion etc.).

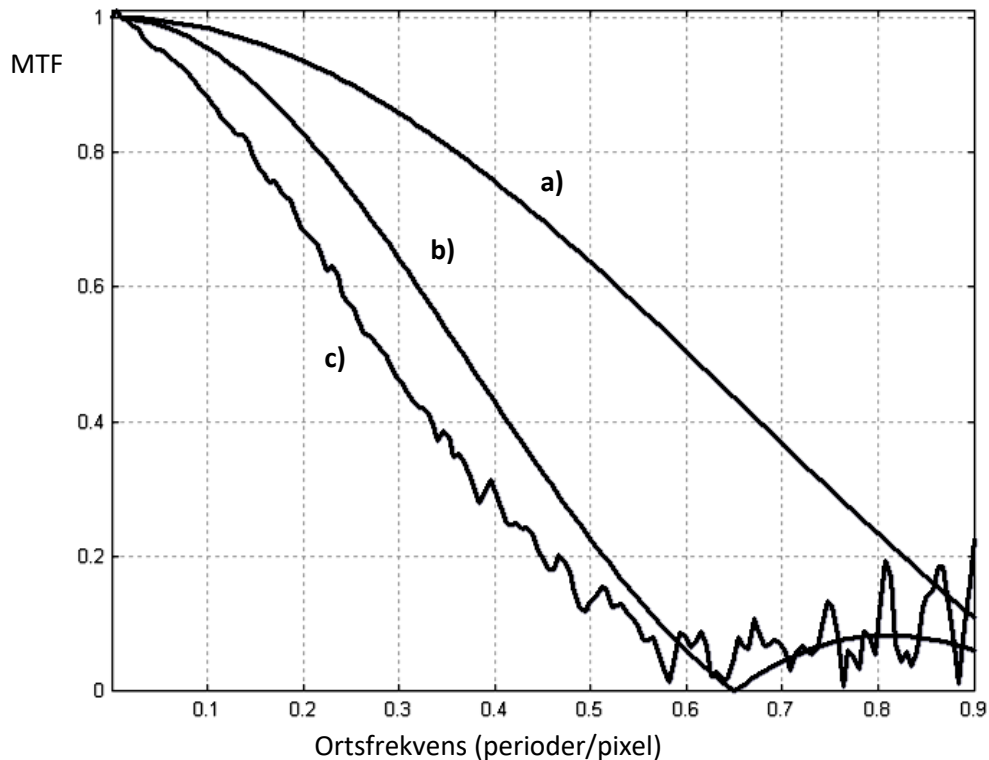


Fig. 22.7. Jämförelse mellan uppmätt MTF_{sensor} och teoretiska kurvor. a) Kurva för ideal sensor enligt kap. 17.6. b) Kurva för ideal sensor utrustad med idealt anti-aliasing filter med nollställe vid 0.65 perioder/pixel. c) Verklig MTF_{sensor} uppmätt med metoden som beskrivs i kap. 17.9. Kurvan är osäker för ortsfrekvenser över 0.6 pga. brus.

Kameran var en SONY DSC-R1, vilken är utrustad med anti-aliasing filter.

Det är framförallt på kameror med lite större sensorer, och därmed oftast större pixelstorlek (typiskt runt 5 μm), som anti-aliasing filter används. För kameror med små sensorer och liten pixelstorlek (typiskt runt 1 μm) kommer optiken att fungera som ett effektivt anti-aliasingfilter på grund av dess begränsade upplösning.

22.6. Sampling i två dimensioner

Så här långt har vi behandlat sampling som ett endimensionellt problem med utgångspunkt från Fig. 22.1. Men i verkligheten är ju bilder tvådimensionella, och

CCD- och CMOS-kretsar utför därför sampling i ett tvådimensionellt mönster enligt Fig. 22.8.

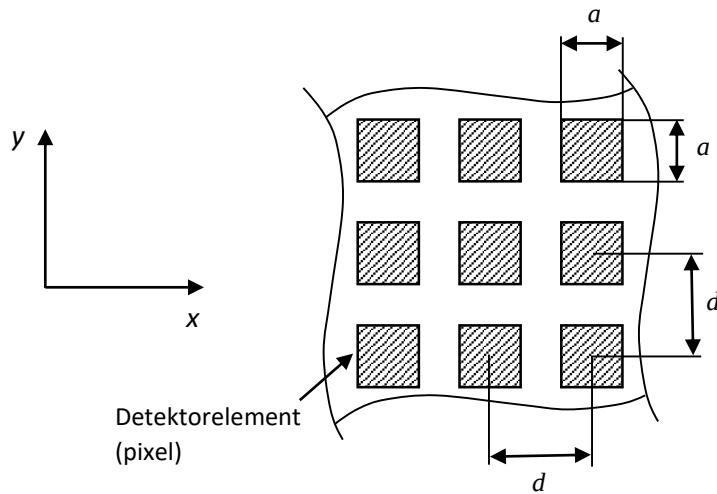


Fig. 22.8. Typisk layout för pixlarna i en kamerasensor. De skuggade områdena visar pixlarnas ljuskänsliga ytor som har kantlängden a . Värdet på a kommer att påverka sensorns MTF enligt kap. 17.6, men påverkar inte samplingen. Det som påverkar samplingen är centrum-till-centrum avståndet mellan pixlarna, d .

Vi ska beteckna samplingfrekvenserna i x - och y -riktningarna med $\nu_{s,x}$ och $\nu_{s,y}$. Dessa samplingfrekvenser kan vara olika (vilket ofta är fallet i t.ex. dokumentscannern), men i praktiken är de lika stora i digitalkameror. För sensorn i Fig. 22.8 har vi att $\nu_{s,x} = \nu_{s,y} = \frac{1}{d}$. Det endimensionella samplingfallet som vi hittills betraktat motsvarar att vi avbildar ett streckmönster med horisontell eller vertikal orientering på sensormatrisen, dvs fall a och b i Fig. 22.9. Då behöver vi bara betrakta sampling i en riktning (x -riktningen för a -fallet, och y -riktningen för b -fallet), eftersom man i vinkelräta riktningen inte har någon mönstervariation alls.

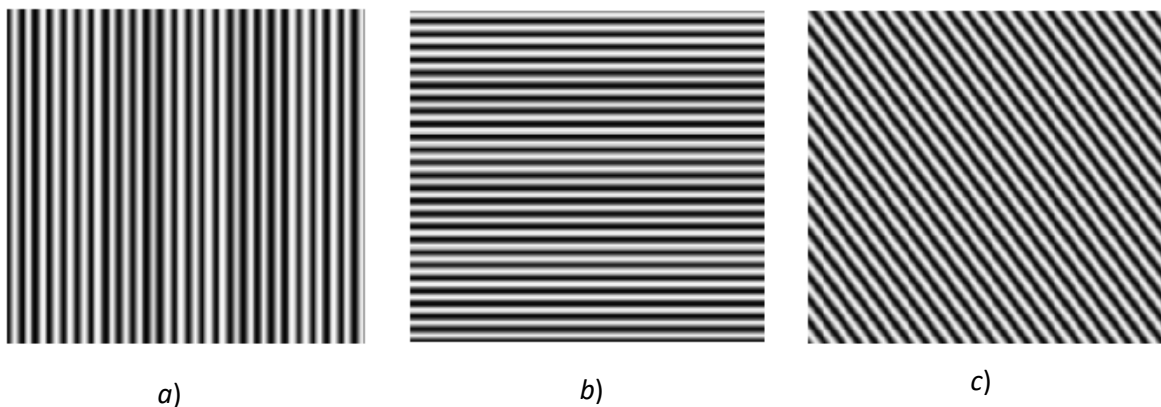


Fig. 22.9. Tre linjemönster med samma linjetäthet, dvs samma ortsfrekvens, men olika orientering.

Låt oss beteckna periodlängden för mönstren i Fig. 22.9 med Δs , dvs ortsfrekvensen $\nu = \frac{1}{\Delta s}$. Villkoret för att undvika aliasing i *a*- och *b*-fallen är då detsamma som vi såg i kap. 22.3, nämligen att $\nu_s > 2\nu$, vilket är samma sak som att $d < \frac{\Delta s}{2}$. I *c*-fallet har vi emellertid ett mönster som varierar i både *x*- och *y*-riktningarna. Detta fall är schematiskt illustrerat i Fig. 22.10.

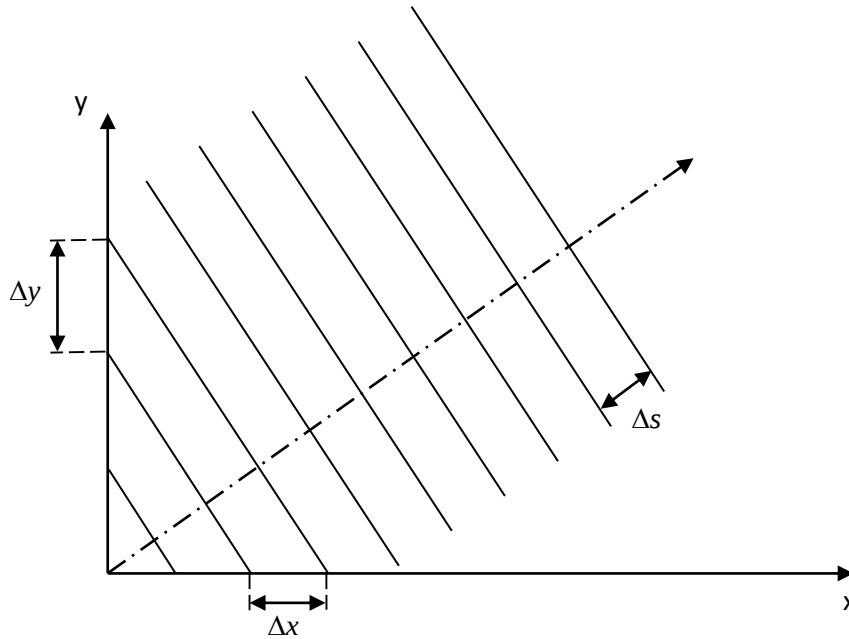


Fig. 22.10. Ett linjemönster med godtycklig orientering kan beskrivas med två ortsfrekvenskomponenter $\nu_x = \frac{1}{\Delta x}$ och $\nu_y = \frac{1}{\Delta y}$.

Mönstrets verkliga ortsfrekvens $\nu = \frac{1}{\Delta s}$ kan delas upp i två komponenter, $\nu_x = \frac{1}{\Delta x}$ och $\nu_y = \frac{1}{\Delta y}$. När vi rör oss längs *x*-axeln ser vi en skenbar ortsfrekvens ν_x , och när vi rör oss längs *y*-axeln ser vi en skenbar ortsfrekvens ν_y . ν_x och ν_y kommer att bero på linjemönstrets orientering, men är alltid $\leq \nu$.

Samplingkriteriet i två dimensioner kan skrivas som

$$\nu_{s,x} > 2\nu_x \text{ och } \nu_{s,y} > 2\nu_y \quad (22.3)$$

Bägge villkoren måste vara uppfyllda samtidigt för att man ska undvika aliasing. Låt oss titta på vad detta innebär genom att räkna ett litet exempel:

Exempel 3: Antag att vi har en matrissensor enligt Fig. 22.8 med $d = 5.0 \mu\text{m}$. Vilken är den högsta ortsfrekvens, ν , som denna sensor kan korrekt registrera för mönsterorienteringar enligt Fig. 22.9 a-c?

Sensorn har $\nu_{s,x} = \nu_{s,y} = \frac{1}{d} = \frac{1}{5.0 \times 10^{-6}} = 2.0 \times 10^5 \text{ m}^{-1} = 200 \text{ mm}^{-1}$.

För mönsterorientering enligt Fig. 22.9 a) så är $\nu_x = \nu$ och $\nu_y = 0$. För att uppfylla ekv. 22.3 krävs att $\nu_{s,x} > 2\nu_x = 2\nu$. Det innebär att ortsfrekvenser under $\frac{\nu_{s,x}}{2} = 100 \text{ mm}^{-1}$ kommer att registreras korrekt. På grund av symmetri blir resultatet detsamma för mönsterorientering b).

För mönsterorientering c) med 45 graders lutning gäller att $\nu_x = \nu_y = \frac{\nu}{\sqrt{2}}$. Ekv. 22.3 ger nu att $\nu_{s,x} > 2\nu_x = \sqrt{2}\nu$ och $\nu_{s,y} > 2\nu_y = \sqrt{2}\nu$. Bägge villkoren ger samma resultat, nämligen att ortsfrekvenser under $\frac{200}{\sqrt{2}} \approx 140 \text{ mm}^{-1}$ kommer att registreras korrekt. Vi kan alltså korrekt registrera ca. 40% högre ortsfrekvenser om de är orienterade diagonalt med 45 graders lutning jämfört med lodrät och vågrät orientering.

I exemplet ovan ser vi att den högsta ortsfrekvens en sensor kan registrera korrekt beror på orienteringen av mönstret. Man klarar av tätare mönster om dessa är orienterade diagonalt istället för vertikal och horisontellt. När man talar om maximal frekvens som kan samplas korrekt (Nyquistfrekvensen) så brukar man utgå från ett ”worst case scenario”, dvs vertikal eller horisontell orientering. I exempel 3 skulle man alltså säga att Nyquistfrekvensen för sensorn är 100 mm^{-1} . Men faktum är alltså att den med rätt mönsterorientering klarar högre frekvenser än så (upp till $\sqrt{2}$ högre).

22.7. Sampling med Bayer mosaik

Så gott som alla digitalkameror är utrustade med färgade filter ovanpå pixlarna för att kunna registrera färgbilder. Filterna är vanligen utlagda i ett så kallat Bayer mosaik-mönster enligt Fig. 22.11.

R	G	R	G
G	B	G	B
R	G	R	G
G	B	G	B

Fig. 22.11. Ett Bayer mosaik-mönster bestående av röda, gröna och blå filter är utlagt över pixlarna i de flesta sensorer. Detta möjliggör fotografering i färg.

Vi ska först titta rent teoretiskt på hur höga ortsfrekvenser som kan korrekt registreras i de olika färgbanden med Bayer mosaik. Detta görs i kap. 22.7.1. Därefter ska vi i kap. 22.7.2 titta lite på experimentella resultat som erhållits från en digitalkamera, och jämföra med de teoretiska förväntningarna.

22.7.1. Sampling med Bayer mosaik: Teori

Som synes i Fig. 22.11 så medför Bayer-mönstret att man har lite olika samplingtäthet i olika färgband. Tätast sampling har man i det gröna spektralbandet, medan de röda och blå banden har glesare sampling. Detta medför att aliasing kommer att inträffa vid olika ortsfrekvenser i olika färgband, och dessutom kommer orienteringen att spela in som vi såg i kap. 22.6. Eftersom rött och blått samplas med samma täthet, så får vi bara två olika fall, grönt och rött/blått. Låt oss börja med att titta på rött/blåa fallet.

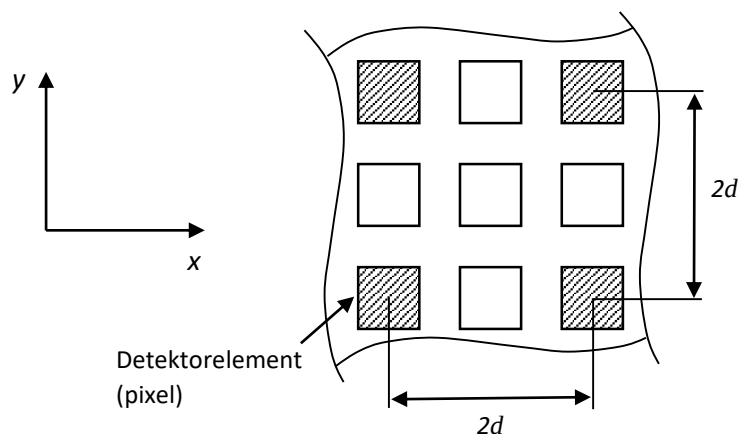


Fig. 22.12. I en Bayermosaik är varannan pixel i både horisontell och vertikal riktning röd (se Fig. 22.11). Detta innebär att avståndet mellan röda samplingpunkter är $2d$ både i x - och y -riktningarna (d är centrum-till-centrum avståndet mellan pixlarna, se Fig. 22.8). Samma samplingtäthet gäller för blåa pixlar.

Samplingen i de röda och blå färgbanden är alltså ekvivalent med samplingen som beskrevs i kap. 22.6 (sensor utan Bayer mosaik) förutom att samplingfrekvenserna bara är hälften så höga, dvs $\nu_{s,x} = \nu_{s,y} = \frac{1}{2d}$. Det innebär att vi bara klarar ortsfrekvenser som är hälften så höga jämfört med en (svartvit) sensor utan Bayer mosaik.

Låt oss nu betrakta samplingen i det gröna färgbandet. De gröna pixlarna bildar ett mönster enligt Fig. 22.13. Det som kan verka lite förvirrande när man tittar på Fig. 22.13 a) är att de gröna pixlarna inte bildar sammanhängande rader och kolumner. Om vi emellertid vrider sensorn 45 grader får vi Fig. 22.13 b), och då ser vi att de gröna pixlarna bildar sammanhängande rader och kolumner. Så låt oss utgå från 45 graders-läget i vårt resonemang. Vi har nu samma situation som vi redan tittat på i Fig. 22.8 och 22.12, men vi ser i Fig. 22.13 att centrum-till-centrum avståndet

mellan pixlarna är $\sqrt{2}d$. Detta ger oss samplingfrekvenserna $\nu_{s,x} = \nu_{s,y} = \frac{1}{\sqrt{2}d}$. Högsta ortsfrekvens som kan registreras korrekt blir då hälften så hög, dvs $\frac{1}{2\sqrt{2}d}$ för vertikal och horisontell mönsterorientering. Som vi emellertid såg tidigare så klarar man $\sqrt{2}$ gånger så hög frekvens om mönstret är orienterat i 45 grader, dvs då får vi maxfrekvensen $\frac{1}{2d}$. Men eftersom vi redan hade vridit sensorn 45 grader när vi började resonemanget, så innebär det att om vi håller kameran på normalt sätt (ovriden eller 90 grader vriden) så kommer vi att klara vertikala och horisontella linjemönster upp till en ortsfrekvens av $\frac{1}{2d}$, dvs samma som för en svartvit sensor enligt Fig. 22.8. Skillnaden mellan de bägge fallen är att man med sensorn i Fig. 22.8 klarar ännu högre frekvenser om mönstret lutar, medan motsatsen gäller för sensorn i Fig. 22.13 (man klarar bara lägre frekvenser om mönstret lutar). Det är lätt att bli förvirrad av alla dessa olika samplingfall, och därför finns en sammanfattning av resultaten i tabellform under Fig. 22.13. Ett praktiskt fall räknas också igenom i Exempel 4.

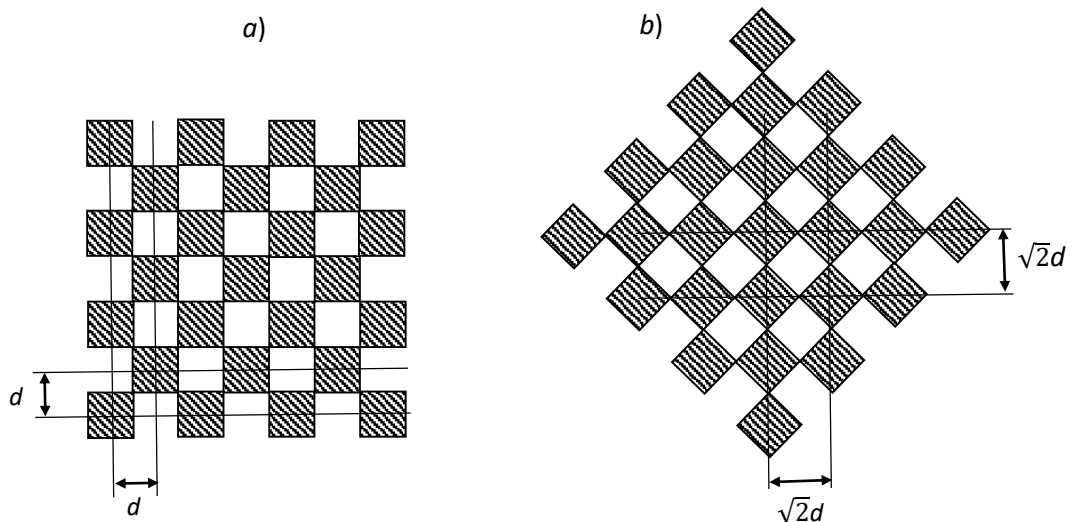
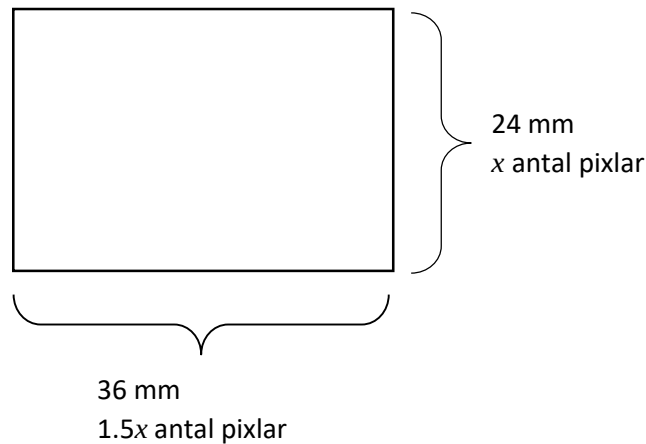


Fig. 22.13. a) Layout för gröna pixlar (skuggade) i en sensor med Bayer mosaik, jfr. Fig. 22.11. Som tidigare representerar d centrum-till-centrum avståndet för pixlarna. b) Samma pixelmönster som i a), men vridet 45 grader så att det bildar horisontella rader och vertikala kolumner. Centrum-till-centrum avståndet mellan gröna pixlar är nu $\sqrt{2}d$.

Sensor (d = centrum-till-centrum avst. för pixlar)	Max. ortsfrekv. för vertikala/horisontella mönster	Max. ortsfrekv. för mönster med 45 graders lutning
Svartvit	$\frac{1}{2d}$	$\frac{1}{\sqrt{2}d}$
Bayer, röda/blå pixlar	$\frac{1}{4d}$	$\frac{1}{2\sqrt{2}d}$
Bayer, gröna pixlar	$\frac{1}{2d}$	$\frac{1}{2\sqrt{2}d}$

Exempel 4: Digitalkameran Sony A7R är utrustad med en fullformatsensor (24 mm x 36 mm) som har Bayer-mosaik och 36 megapixlar totalt (alla färger inräknade). Den är inte utrustad med anti-aliasing filter. Låt oss först räkna ut centrum-till-centrum avståndet för pixlarna. Låt oss anta att vi har x pixlar längs sensorns 24 mm sida. Då har vi $1.5x$ pixlar längs 36 mm sidan (sensorn har bredd/höjdförhållandet 1.5).



Totalt antal pixlar = $1.5x \times x = 36 \times 10^6$, vilket ger $x = 4899$. Alltså går det 4899 pixlar på 24 mm, vilket ger en pixelstorlek (centrum-till-centrum avstånd) av $d = 4.9 \mu\text{m}$. Om vi stoppar in detta d -värde i tabellen på föregående sida får vi följande maximala ortsfrekvenser som sensorn klarar av att registrera utan aliasing.

Sensor ($d = 4.9 \mu\text{m}$)	Max. ortsfrekv. för vertikala/horisontella mönster (sort mm^{-1})	Max. ortsfrekv. för mönster med 45 graders lutning (sort mm^{-1})
Röd/blå färgkanal	$\frac{1}{4d} = 51$	$\frac{1}{2\sqrt{2}d} = 72$
Grön färgkanal	$\frac{1}{2d} = 102$	$\frac{1}{2\sqrt{2}d} = 72$

För andra mönsterorienteringar kommer max ortsfrekvens som klaras av att ligga mellan de värden som angivits för vertikal/horisontell och 45 graders lutning. Att aliasingeffekter uppträder vid olika ortsfrekvenser i olika färgband, medför att vi kan förvänta oss att det dyker upp färgade störningar i bilderna. Detta är också mycket riktigt det man ser i digitala bilder, om man betraktar små strukturer i motivet.

22.7.2. Sampling med Bayer mosaik: Experimentella resultat

Låt oss göra ett litet experiment för att se hur väl värdena i tabellen i Fig. 22.13 tycks stämma med verkligheten. En bild av streckmönster med olika orientering fotograferades med en Sony A7R (sensor med 7392×4920 pixlar). Bilden togs ut från kameran både i RAW-format och som jpeg. Från RAW-filen plockades enbart de röda pixelvärdena ut, vilket gav en bild med 3696×2460 pixlar. En detalj av denna bild visas i Fig. 22.14a. Sedan interpolerades denna rödbild upp till dubbel bredd och höjd (7392×4920) med hjälp av bikubisk interpolation (vilket är en

hyggligt bra metod, se kap. 14.9). En detalj av den interpolerade bilden visas i Fig. 22.14b.

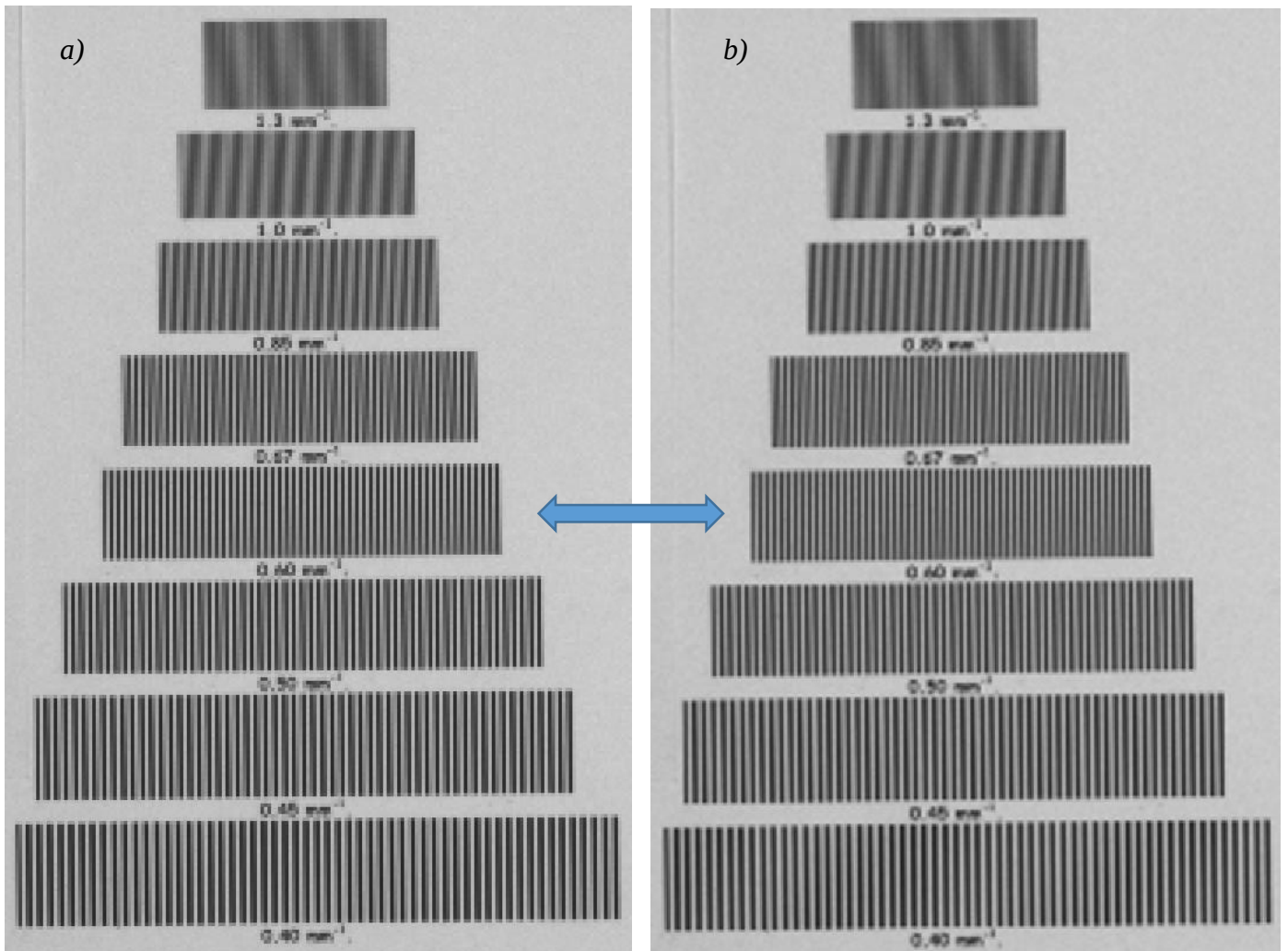


Fig. 22.14. Bilder uppbyggda av enbart röda pixlar extraherade ur RAW-filen. a) Detalj direkt ur R-bilden. b) Samma bild som a), men genom bikubisk interpolation har pixelantalet dubblerats både på bredden och höjden.

Anledningen till att RAW-data användes vid framställning av bilden i Fig. 22.14, är att man då har full kontroll på bildbehandlingen som används. På så vis kan man undvika påverkan av diverse okända ”bildförbättringsrutiner” som används i kameran vid framställning av jpeg. Vid fotograferingen avbildades 0.60 mm^{-1} mönstret, markerat med blå pil, på sensorn med ortsfrekvensen $\frac{1}{4d}$ (se tabell i Fig. 22.13), dvs den maximala frekvensen som kan återges utan moiré (samplingfrekvensen i rött motsvarade alltså 1.2 mm^{-1}). Genom att utvärdera bilderna med ImageJ, kunde det verifieras att ortsfrekvenser i motivet upp till 0.60 mm^{-1} återgavs korrekt, medan högre frekvenser gav moiré-effekt. Ingen skillnad förelåg mellan den icke-interpolerade och den interpolerade bilden. För frekvenser över 0.60 mm^{-1} så gällde att 0.67 återgavs som 0.53 och 0.85 återgavs som 0.35 .

Detta stämmer perfekt med samplingteorin, som säger att över gränshänsen så återges linjemönster med en frekvens som är skillnaden mellan samplingfrekvens och verklig frekvens ($1.2 - 0.67 = 0.53$ och $1.2 - 0.85 = 0.35$).

Som väntat medför inte interpolationen att man kan återge högre ortsfrekvenser, men däremot återges lägre ortsfrekvenser på ett bättre sätt. Jämför t.ex. det understa mönstret (0.40 mm^{-1}) i Fig. 22.14 a & b. Den interpolerade bilden återger helt klart mönstret på ett mera korrekt sätt, trots att ingen ny information har tillförts bilden under interpolationen. Det beror på att interpolationen (bikubisk) är en ganska god approximation av hur man ska göra för att återskapa en originalfunktion ur samplade data (se vidare kap. 22.9).

Låt oss nu titta på mönster med 45 graders orientering. Fig. 22.15 visar en detalj ur samma originalbild som Fig. 22.14b. Här återges mönster upp till och med 0.85 mm^{-1} på ett korrekt sätt, medan 1.0 och 1.3 ger moiré (kvantitativt kontrollerat med ImageJ). Detta är helt i enlighet med tabellen i Fig 22.13 som säger att man klarar $\sqrt{2}$ gånger så hög frekvens i 45 graders riktning jämfört med horisontell/vertikal orientering ($\sqrt{2} \times 0.60 = 0.85$).

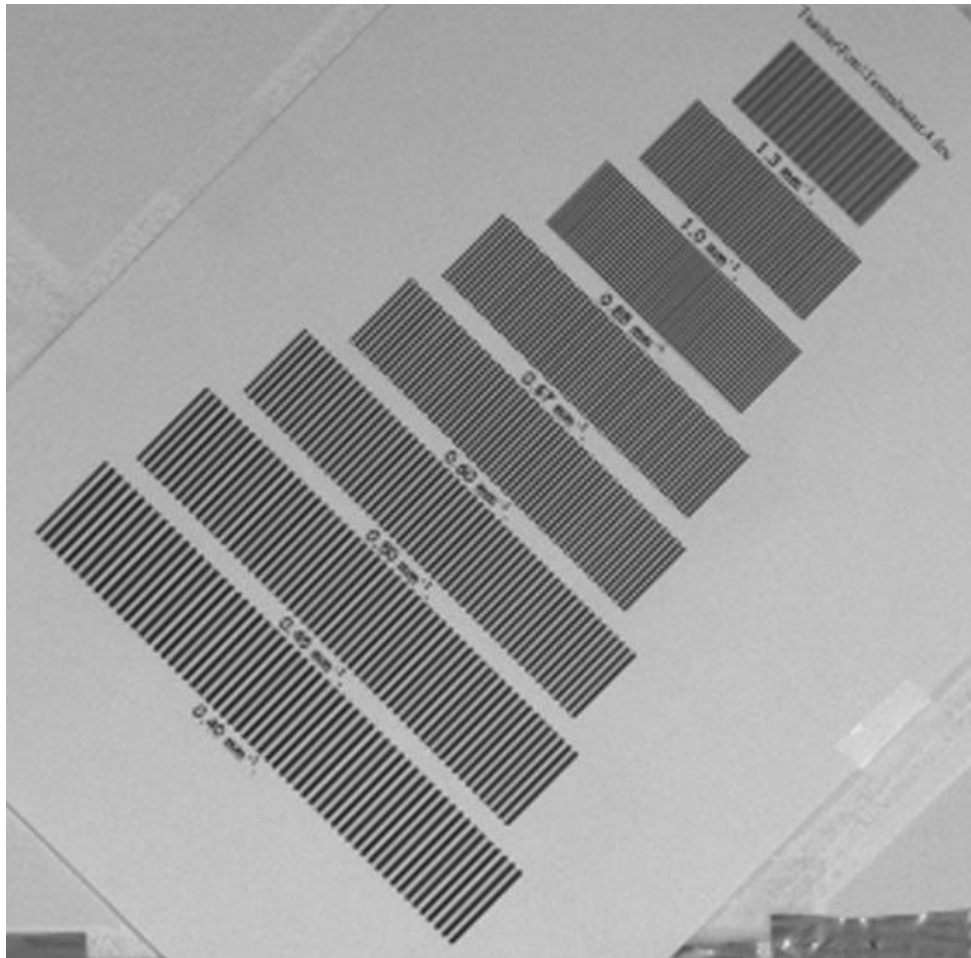


Fig. 22.15. Detalj ur samma interpolerade bild som Fig. 22.14b.

Så här långt stämmer alltså allting precis enligt teorin, tack vare att vi jobbat med data direkt från RAW-filen. Vi hade inte fått samma resultat om vi hade tittat på rödkanalen i jpeg-bilden från kameran (mer om detta senare).

Antag nu att vi gör liknande interpolationer för blå och gröna pixlar¹. Vi har då fått tre ”fullpixlade” delbilder (7392 x 4920) som representerar pixelvärden i rött, grönt och blått. Sätter vi samman dessa till en färgbild så blir resultatet som i Fig. 22.16.

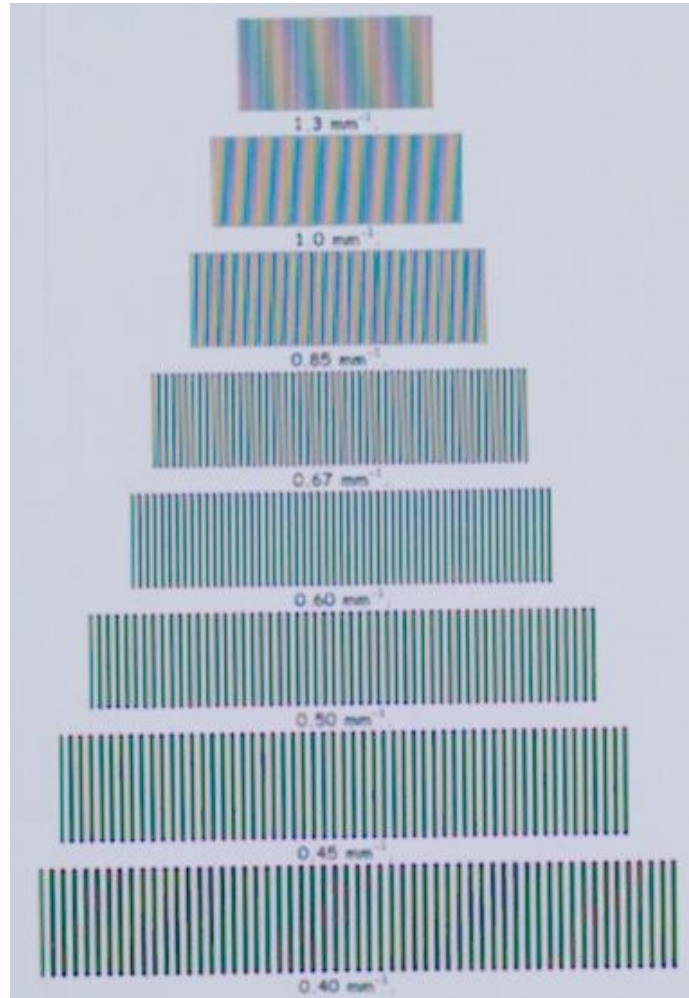


Fig. 22.16. Detalj ur färgbild visande samma utsnitt som i Fig. 22.14. Bilden framställdes ur framinterpolerade RGB-delbilder från en RAW-fil. (Demosaicing algorithm VNG4)

Föga förvånande så blir resultatet bra upp till samma ortsfrekvens som vi tidigare såg i R-delbilden (0.60 mm^{-1}), och däröver får vi tydlig moiré med kraftiga färginslag. Att vi får färginslag är naturligt, eftersom gröna pixlarna kan korrekt registrera högre ortsfrekvenser än vad röda och blå pixlar klarar av (enligt tabellen i Fig. 22.13). Om vi däremot betraktar mönster i 45 graders riktning, Fig. 22.17, så ser vi knappt några störande färgmönster alls, och vi klarar 0.85 mm^{-1} utan moiré (precis som i R-delbilden). Detta resultat är också naturligt, eftersom man för 45

¹ I gröna fallet behöver vi bara dubblera pixelantalet genom interpolation, inte fyrdubbla som för rött och blått, men vi går inte in i detaljer om det.

graders orientering klarar lika höga ortsfrekvenser i alla tre färgbanden (enligt tabellen i Fig. 22.13).

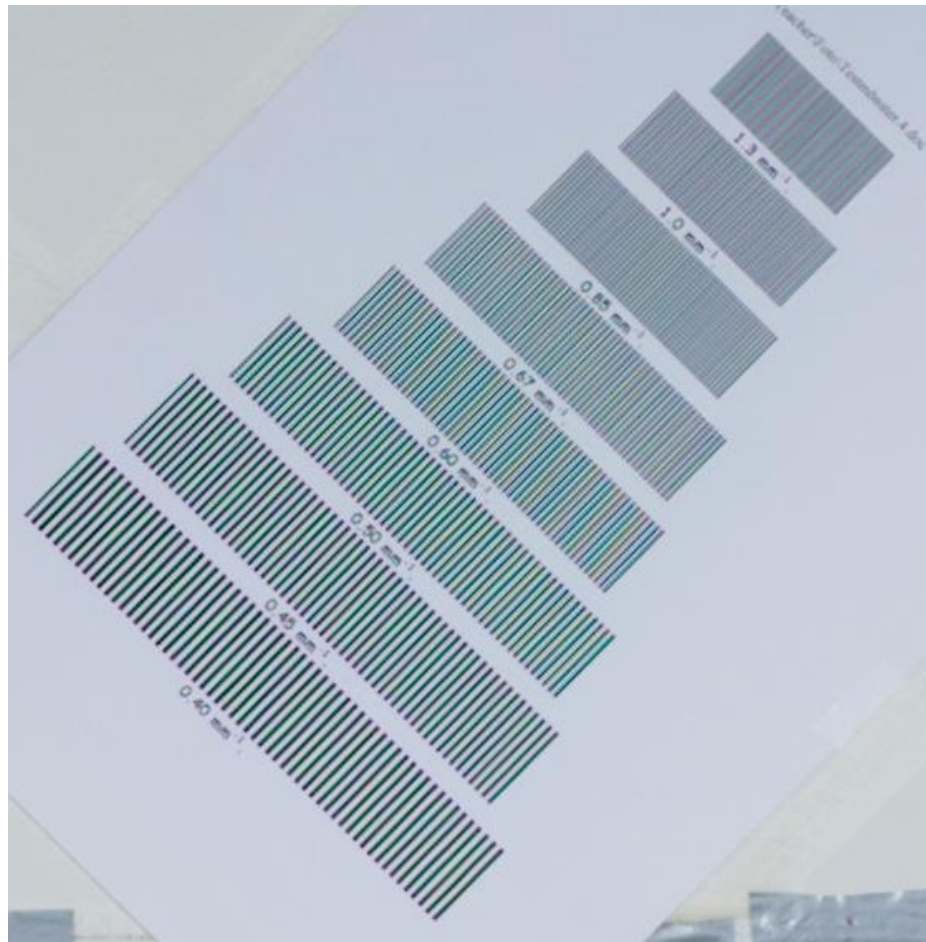


Fig. 22.17. Detalj ur färgbild visande samma utsnitt som i Fig. 22.15. Bilden framställdes ur framinterpolerade RGB-delbilder på samma sätt som Fig. 22.16.

Så här långt har vi inte stött på några överraskningar, allt verkar stämma enligt teorin. Låt oss nu titta på jpeg-filen som kameran levererade av samma motiv, Fig. 22.18. Denna jpeg-bild har alltså av kameran räknats fram från samma RAW-data som vi har använt i Fig. 22.14-22.17. Jämfört med Fig. 22.14b och Fig. 22.16, så klarar jpeg-bilden i Fig. 22.18 av att återge klart högre ortsfrekvenser på ett korrekt sätt. Både 0.67 och 0.85 mm^{-1} återges korrekt, medan 1.0 mm^{-1} och högre återges utan tydliga linjer och med färgstörningar. Detta resultat kan tyckas mycket förbluffande, och strida mot grundläggande fysikaliska principer. Varför kan t.ex. Fig. 22.18b (R-kanalen i jpeg-bilden) korrekt återge betydligt högre ortsfrekvenser än Fig. 22.14b (R-kanalen erhållen ur RAW-data)? Innehållet i RAW-filens röda kanal innehåller ju inte information om högre ortsfrekvenser än 0.60 mm^{-1} ! Svaret är att när jpeg-bildens RGB-kanaler framställdes, så användes inte bara data inom respektive färgband. När t.ex. R-kanalen framställdes så utnyttjades information inte bara från röda pixlar, utan även från gröna och blå pixlar. Man antar då att det finns en viss korrelation (spatiell och/eller spektral) mellan röda, gröna och blå pixlar.

Kort sagt man antar att motivet har ett visst utseende, t.ex. att inom ett väldigt litet område så har pixlarna ungefär samma färgton. I de flesta fall så stämmer dessa antaganden om korrelation väl med hur verkliga motiv ser ut, vilket ger ett bildresultat som är märkbart bättre än om man bara utför interpolation mellan pixlar i samma färg. Men man får komma ihåg att det rör sig om gissningar om motivets utseende, och dessa kan naturligtvis slå fel i vissa extrema fall. Särskilt i vetenskapliga sammanhang bör man ha detta med jpeg-bildernas antagande om pixelkorrelation i åtanke, plus att jpeg-bilder ger en kraftigt icke-linjär återgivning av motivets ljusnivåer².

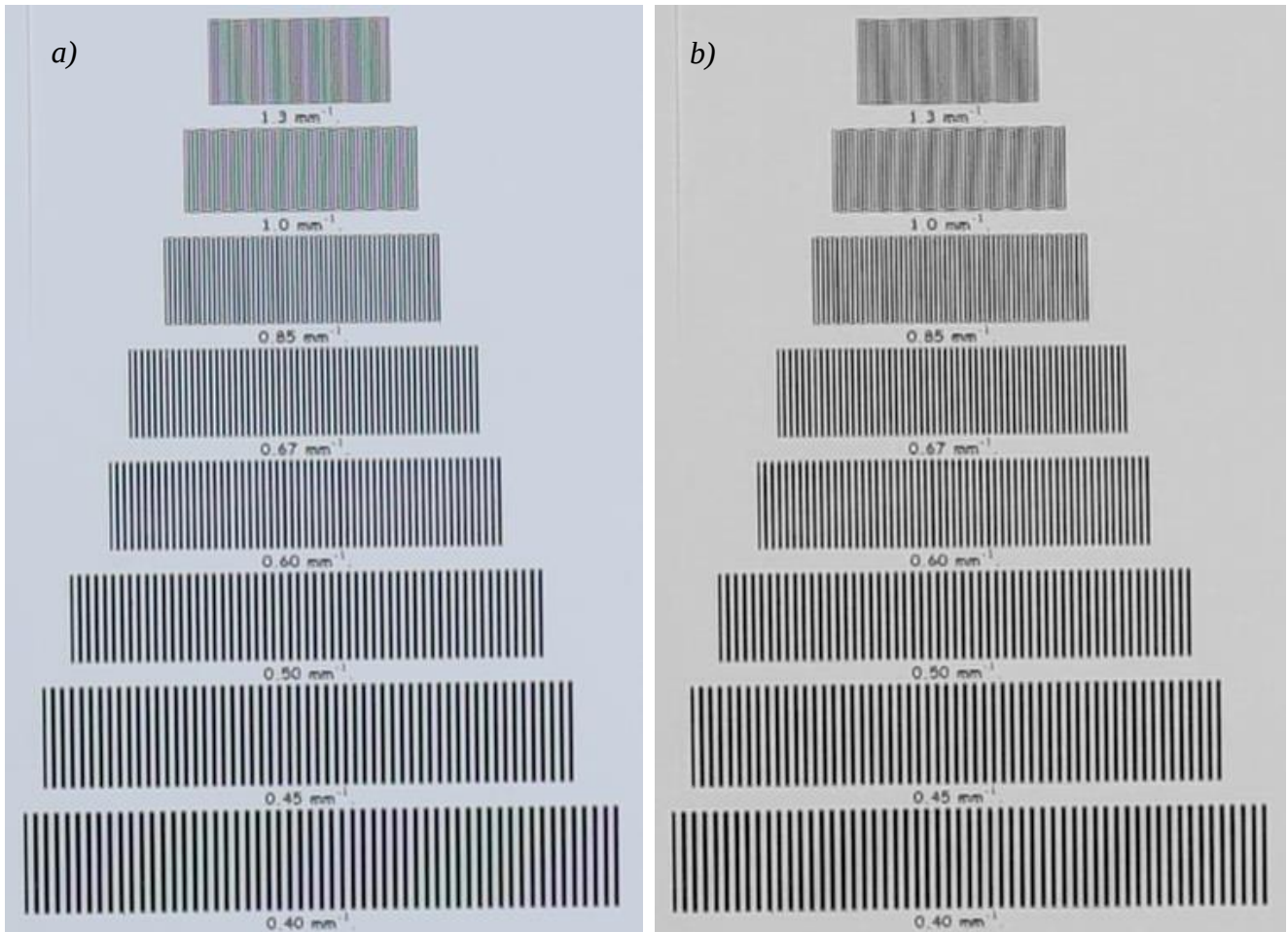


Fig. 22.18. Detalj ur jpeg-bilden direkt från kameran. a) Färgbild (motsvarar Fig. 22.16). b) Enbart rödkanalen (motsvarar Fig. 22.14b).

För 45 graders mönsterorientering uppvisar samma jpeg-bild som i Fig. 22.18 ett utseende som visas i Fig. 22.19. Precis som i Fig. 22.18 återges här frekvenser upp till 0.85 mm^{-1} på ett korrekt sätt, medan 1.0 och högre ger upphov till moiré. För 45 graders orientering får vi alltså samma prestanda som i Fig. 22.15 & 22.17 där bara interpolation inom samma färgband utnyttjades. Det verkar alltså som om man i 45 graders fallet inte tjänar något på att utnyttja antaganden om korrelation mellan

² <https://www.kth.se/social/files/5c88c38056be5bb29067e3b0/Tonreproduktion.Sony%20A7.pdf>

olikfärgade pixlar. Givetvis kan olika kameror utföra interpolationsräkningarna på lite olika sätt, och dessutom kan man med programvara på en yttre dator (s.k. RAW-converter) omvandla en RAW-fil till jpeg. I det sistnämnda fallet kan man vanligen välja mellan ett antal olika rutiner för att göra färginterpolation (eller ”demosaicing” som det kallas på engelska). Detta gör att jpeg-bilderna kan få lite olika utseende. Genom att testa lite olika demosaicing-rutiner med en RAW-converter, så kunde dock konstateras att resultatet skiljde sig ganska lite, i varje fall för de streckmönster som vi här har tittat på. Så de tidigare angivna slutsatserna verkar hålla ganska bra.



Fig. 22.19. Ett annat utsnitt ur samma jpeg-bild som visas i Fig. 22.18a.

Vi kan sammanfatta resultaten av undersökningarna på följande sätt:

1. Om färginterpolation (demosaicing) utförs enbart mellan pixlar i samma färg, erhålls ett resultat som stämmer bra med de fysikaliska gränser som ges i tabellen i Fig. 22.13. Det betyder att horisontella/vertikala streckmönster återges korrekt upp till ortsfrekvensen $\frac{1}{4d}$, där d = centrum-tillcentrum avståndet mellan pixlar (se Fig. 22.13). För mönster i 45 graders-riktningen klaras en frekvens som är $\sqrt{2}$ gånger högre.
2. Om färginterpolation utförs med antagande om att det finns en korrelation mellan pixelvärdena för pixlar med olika färg, så klarar man typiskt $\sqrt{2}$

gångar högre frekvenser för horisontella/vertikala mönster än i fall 1, medan resultatet för 45 graders-mönster är ungefär detsamma som i fall 1.

I praktiken utförs färginterpolation nästan alltid enligt fall 2. Det innebär att man ganska oberoende av mönsterorienteringen kan förvänta sig att korrekt återge frekvenser upp till $\frac{1}{2\sqrt{2}d}$, där d = centrum-till-centrum avståndet mellan pixlarna.

Som nämndes ovan så ger olika algoritmer tillgängliga i en RAW-converter ganska likartade resultat (i denna studie användes RawTherapee). Men vi ska komma ihåg att vi bara har studerat ett väldigt speciellt motiv, nämligen svart/vita streck. De olika algoritmerna är nämligen olika bra på att undertrycka olika artefakter som kan uppstå vid betraktande av andra motivstrukturer. Men det ska vi inte gå in på här, den intresserade kan hitta information om olika algoritmers för- och nackdelar på webben.

22.8. Mikrostegning genom förflyttning av sensorn

Man kan praktiskt taget eliminera aliasingproblemet genom att registrera flera bilder av samma motiv. Om man därvid gör mycket små förskjutningar av sensorn i sid- och höjddled mellan bilderna, kommer man att uppnå en högre samplingfrekvens. Låt oss ta sensorn i Fig. 22.8 som exempel. Antag att vi tar totalt fyra bilder av samma motiv, och förskjuter sensorn mekaniskt sträckan $\frac{d}{2}$ i sid- och höjddled mellan bilderna. Detta innebär att vi har fördubblat samplingfrekvensen. I Fig. 22.6 skulle det innebära att Nyquistfrekvensen hamnar på 1 period/pixel. Vi kan då se att MTF-värdet för en typisk kamera är praktiskt taget lika med noll. Ingen risk för moiré alltså, och vi behöver inte använda något anti-aliasing filter som suddar till bilden. Nackdelarna är naturligtvis att kameran blir mer avancerad och dyr, att man får fyra gånger så stor datamängd, samt att motivet måste vara stillastående. Metoden används därför mest i avancerade specialkameror som används för t.ex. dokumentation av föremål i muséer och liknande. Men den förekommer också i ett par kameror från Olympus och Sony ämnade för konsumentmarknaden. Förflyttningen av sensorn sker vanligen med piezo-elektriska kristaller som tillåter förflyttning med submikrometer-precision. I kamerorna från Olympus och Sony utnyttjas samma piezo-element både till mikrostegning och bildstabilisering (kap. 24.4).

Förutom att kraftigt reducera aliasing-problemen, så kan mikroförflyttningar av sensorn undvika problem som uppstår genom färginterpolationen som görs för sensorer med Bayer-mosaik (kap 14.8). Låt oss ta sensorn i Fig. 22.12 som exempel. Genom att ta fyra bilder av samma motiv, och förskjuta sensorn sträckan d i sid- och höjddled mellan bilderna, så har vi ökat samplingfrekvensen med en faktor två i röda och blå våglängdsbanden. Då har vi alltså lika tät sampling i alla tre våglängdsbanden (plus att vi samplar gröna värden två gånger för ökad noggrannhet). Ingen färginterpolation behöver göras, vilket förbättrar både färgåtergivningen och skärpan.

Slutligen kan man kombinera ovanstående tekniker, så att man både kan höja Nyquistfrekvensen i samtliga färgband till 1 period/pixel, och samtidigt slippa färginterpolation. Detta åstadkoms genom att registrera totalt 16 bilder av samma motiv, med fyra stycken $\frac{d}{2}$ – förflyttningar i sidled och lika många i höjddled. Kamerafabrikanten Sinar tillverkar kameror som medger denna typ av bildregistrering. Typiskt utnyttjas då en ”full frame transfer” CCD (kap. 14.2) i storlek 36 mm x 48 mm med 48 megapixlar och Bayer-mosaik. Centrum-till-centrum avståndet mellan pixlarna är 6.0 μm , och vid mikrostegning flyttas alltså sensorn 3.0 μm i sid- och höjddled. En kamera av denna typ används på Kungliga Biblioteket i Stockholm för att dokumentera gamla skrifter.

22.9. Rekonstruktionsproblemet

Även om man uppfyller samplingteoremet $v_s > 2v_{\text{max}}$, och därmed undviker aliasing, så ser man ändå ofta taggigheter och andra skönhetsfel i digitala bilder. Vad beror detta på? Om man följer samplingteoremet så har man ju registrerat all information om originalet, och borde därför få en bild som helt överensstämmer med originalet, eller hur? Nej, tyvärr är det inte fullt så enkelt i praktiken. Såsom tidigare sagts kan man ur samplade data helt och hållet återskapa originalet om man har följt samplingteoremet. Detta kräver emellertid att man ur samplade data gör en korrekt rekonstruktion. Detta görs t.ex. med data från en musik-CD när man spelar upp den på CD-spelaren. Utsignalen är i detta fall en analog signal som rekonstruerats från digitala data. Resultatet blir utomordentligt bra, och man störs inte av några digitaliseringseffekter. I fallet digitala bilder är det annorlunda. Här görs i allmänhet ingen rekonstruktion alls, utan man visar helt enkelt upp samplade data direkt på t.ex. en datorskärm (man gör inte ens en linjärinterpolation enligt Fig. 22.2!). Det är just avsaknaden av en korrekt rekonstruktion som gör att man får taggigheter och andra felaktigheter i den digitala bilden.

Varför gör man då ingen korrekt rekonstruktion och slipper dessa felaktigheter? Skillnaden mot fallet med musik-CD, är att de digitala bilderna praktiskt taget aldrig omvandlas till analoga bilder. Slutprodukten är alltså inte i form av en kontinuerligt varierande funktion, utan består av diskreta pixlar. Detta innebär att vi för att verkligen kunna visa upp resultatet av en korrekt bildrekonstruktion ur samplade värden, skulle behöva öka på pixelantalet mycket kraftigt när vi presenterar slutbilden. En bild från en 15 megapixel-kamera skulle kanske behöva presenteras som en 300 megapixelbild på datorskärmen för att verkligen ge rättvisa åt den rekonstruerade bilden. Detta är naturligtvis svårt av praktiska skäl, och resultatet blir att man oftast sitter och tittar direkt på de råa samplade bilderna utan att göra någon rekonstruktion. Så länge man följer samplingteoremet kommer man emellertid inte att få några moiré-effekter av den typ som visas i Fig. 22.5, utan felaktigheterna blir av subtilare art (t.ex. en viss taggighet i konturer).

Om det rör sig om ett mindre delområde av en bild, kan man emellertid kosta på sig att göra en rekonstruktion och visa resultatet som en digital bild med många fler pixlar än originalet. Ett exempel på detta visas i Fig. 22.20, där man genom interpolationsräkningar har ökat på antalet pixlar med en faktor 6 i både x- och y-riktningarna. Interpolationsmetoden som använts kallas bikubisk interpolation, och är en approximation av den teoretiskt korrekta rekonstruktionsmetoden för

samplade data. Observera att ingen ny information erhålls på detta sätt, men informationen synliggörs på ett bättre sätt i b-bilden än i a-bilden.

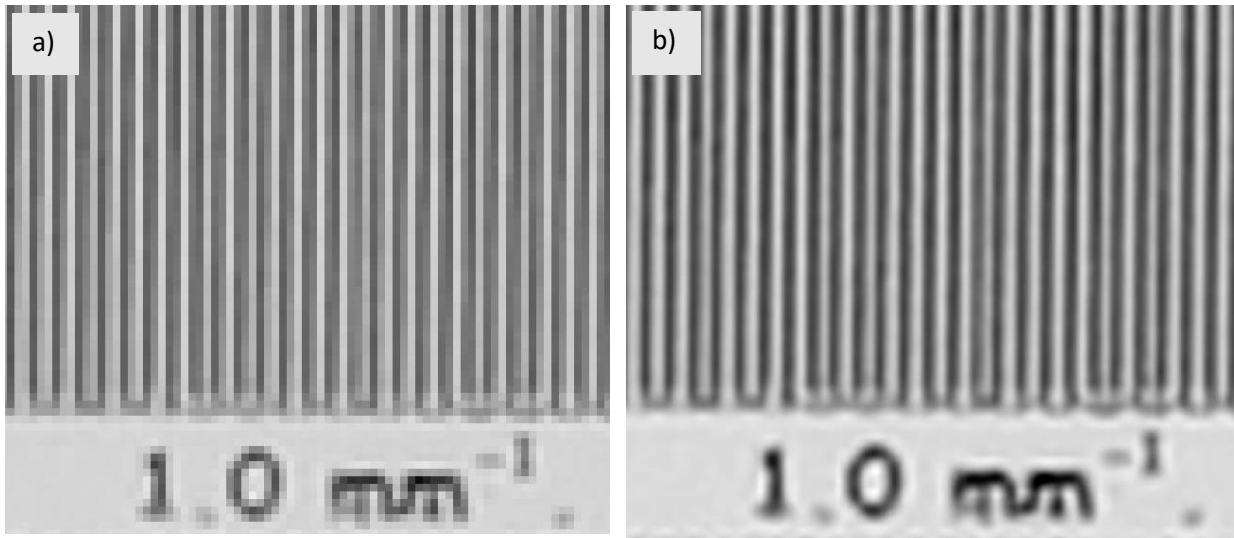


Fig. 22.20. a) Delområde i originalbild från digitalkamera. b) Samma delområde efter bikubisk interpolation.

23. Lagring och utskrift av digitalbilder

23.1. Inledning

Vid digitalfotografering lagras bilderna upp på ett minneskort i kameran. Ofta görs därvid en komprimering för att minska storleken på filerna. Vanligtvis för man sedan över bildfilerna till en dator för mer permanent lagring, samt eventuell bearbetning och utskrift. Vi ska i detta kapitel titta lite på hur sådana utskrifter går till, nödvändiga prestanda mm. Den i amatorsammanhang vanligaste skrivaren är bläckstråleskrivare, som här kommer att beskrivas lite mer detaljerat. Övriga skrivartyper kommer att beröras mer översiktligt. Först ska vi dock titta lite på bildlagring och kompression. Färgbehandling (color management) kommer inte att avhandlas här, men har berörts i kap. 12.6.

23.2 Bildlagring och kompression

I detta avsnitt ska vi bara mycket kort beröra lagring och kompression av digitala bilder. Lagringen sker på ett bildminneskort som sitter instuckat i kameran, se Fig. 14.1. Minneskortet är lätt utbytbara, så när man har fyllt ett kan man byta ut det mot ett nytt tomt och fotografera vidare. Flera olika standarder finns för minneskort, och några vanliga typer går under beteckningarna Compact Flash och Memory Stick. Lagringskapaciteten på bildminnen har successivt ökat, och i dagsläget (2015) ligger den typiskt på 16-64 gigabyte. Detta innebär i praktiken att man ofta får plats med flera tusen bilder på ett minneskort.

För att minska storleken på bildfilerna är det mycket vanligt att man gör en komprimering av data. Den vanligaste komprimerade filtypen betecknas jpeg (alt. jpg), vilket är en förkortning av Joint Photographic Experts Group. Man kan för jpeg-formatet välja olika kraftig komprimering, men typiskt väljer man att komprimera så att man får ner datamängden med cirka en tiopotens. Trots att man alltså kastar bort 90% av ursprungsinformationen, så görs detta på ett så intelligent sätt att man med blotta ögat inte brukar kunna skilja den komprimerade bilden från originalbilden. Det finns mycket utförliga beskrivningar av jpeg-komprimering i t.ex. olika ISO-normer. Vi ska här bara kort nämna de grundläggande principerna.

I originalbilden representeras vanligen varje pixel av tre 8-bitars tal (ett för vardera R-, G- och B-värdet). I första steget till jpeg-komprimering omvandlas bilden till en annan färgrymd, där varje pixel representeras av ett ljushetsvärde och två färgvärden. Detta är bara en koordinattransformation och ger ingen komprimering. Men det smarta är att man kraftigt kan komprimera färgkanalerna, och ändå få en bild som ser bra ut, men man ska undvika att alltför kraftigt komprimera ljushetskanalen. Den komprimerade bilden kan då liknas vid en skarp svartvit bild överlagrad med lite suddigare färginformation. Detta tolkar vår hjärna som en skarp färgbild. Samma typ av kompression används för övrigt inom TV-teknik. Själva kompressionen i jpeg sker genom att först reducera antalet pixlar i de två färgkanalerna, så att man helt enkelt får en grövre pixellering där. I nästa steg, som görs på alla kanalerna, indelas bilden i delbilder om 8 x 8 pixlar som behandlas var för sig. I varje sådan delbild delas informationen upp i sitt ortsfrekvensinnehåll (jmf.

Kap. 17.3). Detta innebär i princip samma sak som när man delar upp t.ex. en akustisk signal i olika deltoner. Nu visar det sig att det mänskliga ögat är betydligt sämre på att särskilja små nivåskillnader i ljusvärde om vi har en hög ortsfrekvens än om vi har en låg. Detta innebär att informationen om de höga ortsfrekvenserna kan lagras med färre databitar. Vissa ortsfrekvenser som knappast förekommer alls i en bild kan till och med slopas helt. Men en undersökning av ortsfrekvensinnehållet i den aktuella bilden måste alltså göras innan man kastar bort information. I Fig. 23.1 visas två exempel på jpeg-komprimerade bilder med olika kompressionsgrad.



Fig. 23.1. Upp till en faktor tio i kompression är vanligen möjlig i jpeg-formatet utan synbar kvalitetsförlust, vänstra bilden. Vid mycket kraftig komprimering blir dock förlusten av information synlig som i den högra bilden. Den grövre pixelleringen som syns överlagrad i högra bilden beror på den indelning i delbilder om 8 x 8 pixlar som görs i jpeg.

jpeg-filer innehåller bildinformation som har processats kraftigt av en i kameran inbyggd processor. Sålunda har interpolationsräkningar gjorts för att få fram RGB-värden i varje pixel. Detta är nödvändigt därför att varje pixel i sensorn bara registrerar en färgkomponent (se kap. 14.8). Vidare har gammakorrektion gjorts (kap. 14.6), samt en datormässig uppskärpning av kantinformation i bilderna. I lite mer professionella sammanhang vill man ofta kunna utföra dessa bildberäkningar på en extern, mer kraftfull dator. På så sätt kan man på ett helt annat sätt styra hela processen och få ett bättre slutresultat. Man kan t.ex. utföra färginterpolation på ett mer avancerat sätt. För att kunna utföra denna bildbehandling på en extern dator måste man emellertid hämta bildinformationen från kameran i så kallat RAW-format. Som namnet antyder är då bilddata helt opåverkade, och utgör helt enkelt den rådata som kommer direkt från sensorn. Förutom detta skiljer sig RAW-formatet genom att man har digitaliserat pixelvärdena med fler bitar, vanligen

någonstans mellan 12 och 16 istället för 8. Detta ger mindre kvantiseringfel, vilket innebär att man bättre kan utnyttja kamerans dynamiska vidd. Särskilt i motivets mörka partier kan man ofta få fram en bättre detaljåtergivning genom att fotografera i RAW-format. Dessutom kan man ofta exponera sensorn lite kraftigare vid fotografering i RAW-format (kap. 14.10), vilket också förbättrar återgivningen av mörka partier. Ett visst problem är dock att olika kamerafabrikanter använder olika RAW-format, och att det behövs speciell mjukvara för att hantera dessa format.

23.3. Bläckstråleskrivarens princip

Om man betraktar en bläckstråleutskrift i hög förstoring, så ser man att den byggs upp av ett stort antal punkter, Fig. 23.2. Den liknar därmed de rasterade bilder som trycks i böcker och tidningar. Precis som i alla andra färgsammanhang behövs i en bläckstråleskrivare tre grundfärger för att bygga upp en bild. Man använder samma färger som när man trycker bilder i färg, nämligen gult, magenta och cyan. Dessutom använder man ett fjärde färgämne, svart, för att uppnå tillräcklig svärta i mörka partier (samt för att spara bläck). I enklare skrivare brukar det finnas två bläckpatroner, en med svart bläck, och en innehållande tre behållare med gult, magenta och cyan bläck.



(Foto: Nils Åslund)

Fig. 23.2. a) En bläckstrålebild sedd på stort avstånd uppvisar grånyanser och färger på ungefär samma sätt som ett gammaldags analogt fotografi. b) I hög förstoring syns att bilden byggs upp av många små bläckfläckar (ca. 50 μm stora).

De bläckstråleskrivare som är avsedda för utskrift av bilder (fotoskrivare) brukar ha fler än fyra bläcksorter. Förutom mättade gul-, magenta- och cyanbläck, har man här lagt till omättat bläck i olika färger. Anledningen till detta är att man annars får en störande prickighet vid utskrift av bildpartier med låg färgmättnad. Som exempel

på hur denna prickighet uppstår, kan vi tänka oss att man vill skriva ut en mycket omättad gul färgton. I en skrivare med bara en sorts gult bläck måste detta ha en mättad färgton för att man ska kunna skriva ut kraftigt gula färger. Om man då vill skriva ut en omättad färg, måste skrivaren lägga ut ett glest mönster med starkt gula prickar. Sett från stort avstånd kommer dessa små prickar, med stora vita pappersytor emellan, att uppfattas som en blek, dvs omättad, gul färg. Ju mindre mättnad färgen ska ha, desto glesare måste prickmönstret vara. Men till slut innebär detta att prickarna ligger så glest att de syns mot den vita pappersytan även på normalt betraktningsavstånd. För att undvika denna effekt använder man i en fotoskrivare ett omättat gult bläck som kan läggas i ett tätare prickmönster, och där varje prick inte kontrasterar så kraftigt mot papperet.

Den principiella funktionen hos en bläckstråleskrivare visas i Fig. 23.3. Ett papper matas stegvis framåt, under det att skrivarhuvudet rör sig fram och tillbaka vinkelrätt mot papperets rörelseriktning. I skrivarhuvudet finns ett stort antal munstycken genom vilka bläckdroppar sprutas ut enligt ett datorstyrt mönster som ger de önskade färg- och gråtonerna. Det finns skrivarhuvuden som arbetar efter två olika principer, termo-principen och piezo-principen. Dessa kommer att beskrivas i nästa avsnitt.

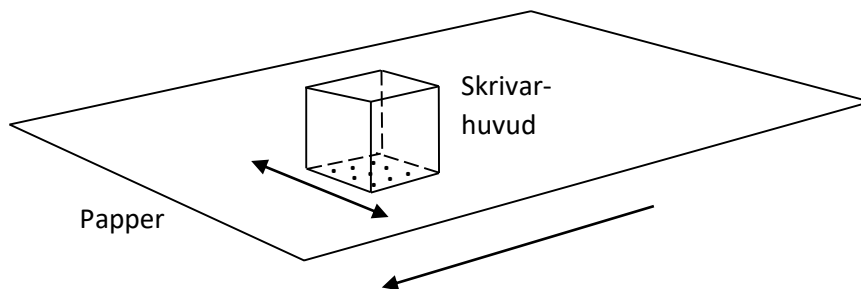


Fig. 23.3. Bläckstråleskrivarens princip. Ett papper förflyttas stegvis i en riktning, medan skrivarhuvudet rör sig fram och tillbaka vinkelrätt mot papperets rörelseriktning. I skrivarhuvudet finns ett stort antal munstycken ur vilka bläckdroppar sprutas ut. Minst fyra olika bläckfärger, gult, magenta, cyan och svart, används för att skriva ut bilder i färg.

23.4. Skrivarhuvud enligt termo- och piezo-principerna.

I alla bläckstråleskrivare gäller att man mycket snabbt och med god noggrannhet måste kunna spruta ut små bläckdroppar. En metod för att göra detta, termo-principen, används av bl.a. Hewlett-Packard och illustreras i Fig. 23.4.

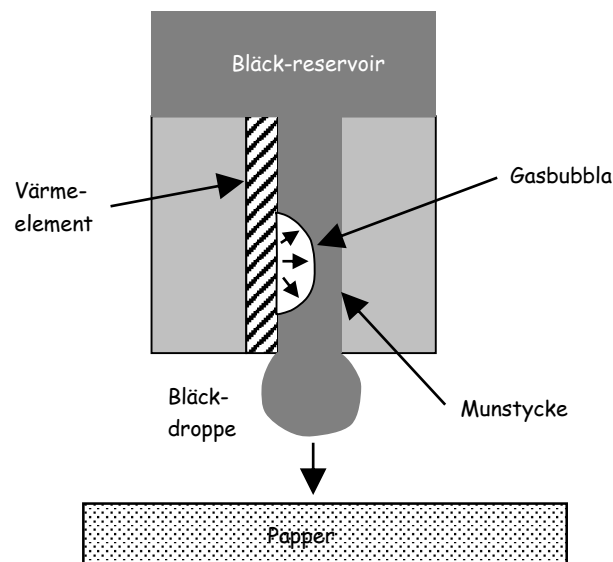


Fig. 23.4. Tvärsnitt i hög förstoring genom ett bläckmunstycke i ett skrivarhuvud som arbetar enligt termo-principen.

Från bläckreservoiren leder smala munstycken ner mot undersidan av skrivarhuvudet. På grund av kapillärkraften rinner inte bläcket ut ur munstyckena, utan det måste på något vis pressas ut. I skrivare som arbetar efter termo-principen sker detta genom att upphetta ett litet värmeelement längs sidan på kanalen. Då kommer bläcket närmast elementet att förångas, och den expanderande gasbubblan pressar ut en droppe bläck ur munstycket. När värmeelementet därefter svalnar försvinner gasbubblan, och nytt bläck från reservoiren fyller på kanalen. Sedan kan processen upprepas igen. Hela förloppet måste ske mycket snabbt, vilket är möjligt eftersom värmeelementet är mycket litet så att det kan värmas upp och kylas av mycket snabbt.

Piezo-principen, som används av Epson, illustreras i Fig. 23.5. Skillnaden gentemot den termiska principen gäller endast sättet på vilket dropparna pressas ut ur munstyckena. I detta fall sker det mekaniskt genom att en så kallad piezo-elektrisk kristall sitter monterad alldeles ovanför bläckkanalen. När man lägger en spänning på kristallen kommer dess tjocklek att snabbt förändras. Genom att välja lämplig polaritet kan man få den att expandera och därigenom pressa ut en bläckdroppe ur

munstycket. När spänningen sedan tas bort återgår kristallen till sin ursprungliga tjocklek, varvid nytt bläck fylls på i kanalen.

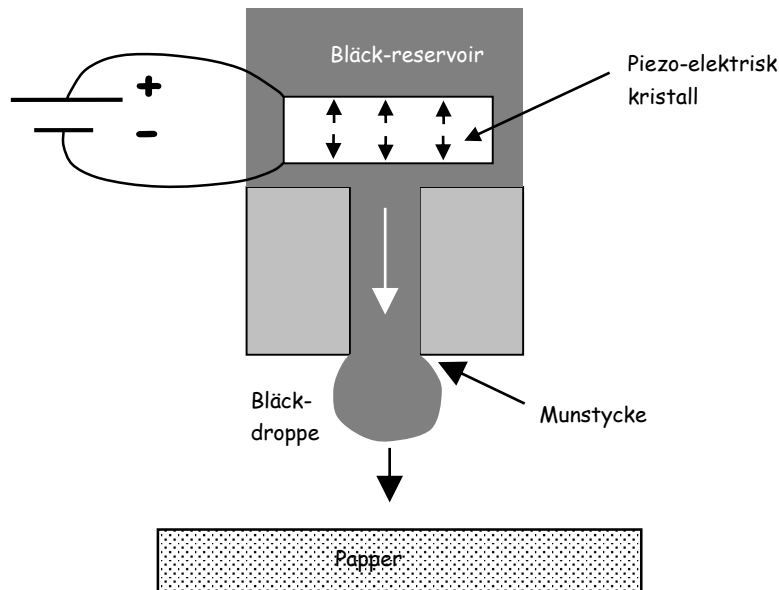


Fig. 23.5. Tvärsnitt i hög förstoring genom ett bläckmunstycke i skrivarhuvud som arbetar enligt piezo-principen.

23.5. Prestanda för bläckstråleskrivare och synsinnet

Ett prestandatal som ofta anges för bläckstråleskrivare är ”upplösning” uttryckt som dpi (dots per inch, dvs. bläckpunkter per tum^{*}), och en typisk siffra kan vara 1200 dpi. Man ska dock inte tro att skrivaren kan trycka denna täthet av bildpixlar. Anledningen är att man i varje bläckpunkt är ganska begränsad vad gäller antalet färger och grånivåer som kan åstadkommas. För att få tillräckligt omfång i återgivningen av färger och grånivåer, krävs därför att varje bildpixel byggs upp av flera bläckpunkter bredvid varandra. Hur många bläckpunkter som krävs för varje pixel kan variera eftersom olika skrivare använder lite olika system. En grov tumregel säger dock att man ska dividera dpi-talet med fyra för att få en uppfattning om hur många pixlar den kan skriva ut per tum. Om således dpi-talet är 1200 kan man förvänta sig en verklig pixeltäthet av cirka 300 ppi (pixels per inch = pixlar per tum). Detta innebär att om man har en digital bild med tre miljoner färgpixlar, så måste den skrivas ut i ett format av minst 120 mm × 180 mm om skrivaren ska kunna återge alla bildens pixlar.

I detta sammanhang kan det också vara på sin plats att tänka lite på synsinnets förmåga att urskilja enskilda pixlar i den bild som skrivits ut. Räcker tätheten 300 ppi för att inte pixelleringen ska begränsa den upplevda bildkvaliteten vid ett betraktavstånd av, säg, 25 cm? Här går åsikterna starkt isär om man tittar i

* En tum är 25.4 mm.

litteraturen. Vissa säger att 200 ppi räcker bra, medan andra hävdar att det kan behövas upp till 4 gånger högre täthet. Frågeställningen blir inte lättare av att ögats egenskaper ändras med betraktningssituationen, t.ex. belysningsnivån. Men låt oss ändå försöka göra en grov uppskattning av vad vi kan förvänta oss. Grovt kan ögats upplösningssförmåga sättas till cirka en bågminut ($1/60$ grad). Detta innebär att man nätt och jämnt kan se ett linjemönster med omväxlande svarta och vita linjer då en period (en svart plus en vit linje) upptar en synvinkel av $1/60$ grad. För att maximalt utnyttja ögats upplösning, krävs då två pixlar inom en synvinkel av $1/60$ grad. Detta ger en pixeltäthet av ca. 700 ppi vid betraktning på 25 cm avstånd. Överslagsräkningen styrker alltså den högre siffran ovan vad gäller behövlig pixeltäthet i en papperskopia. Praktiska försök, utförda i normala betraktningssituationer, brukar dock visa att de flesta personer har svårt att se någon nämnvärd förbättring i bildkvalitet när pixeltätheten överstiger ca. 300 ppi. Denna gräns är dock motivberoende, och man ska inte utesluta att man i vissa fall (t.ex. mycket detalj- och kontrastrika motiv) kan uppleva en viss förbättring genom att öka detta till kanske 600 ppi.

Med utgångspunkt från ovanstående kan vi uppskatta hur många pixlar som krävs i en utskrift av visst format för att vi inte ska störas av pixelleringen vid betraktandet. 300 ppi räcker som vi har sett oftast bra vid betraktning på 25 cm avstånd. Enkel matematik ger oss då att för en $10\text{ cm} \times 15\text{ cm}$ pappersbild behövs ca. 2 Mpixel. Även 200 pixlar per tum ger ”nära fotokvalitet”, vilket skulle innebära att även 1 Mpixel ofta är fullt acceptabelt. Detta visar att det oftast inte behövs så många pixlar i en bild för att vi inte ska störas av pixelleringen. Viktigare är ofta storleken på den sensor som använts vid fotograferingen. Sensorstorleken har nämligen en avgörande betydelse för hur skarpa och brusfria bilder man kan få.

Sammanfattningsvis kan man säga att även ganska enkla bläckstråleskrivare brukar ge en tillfredsställande kvalitet för ”normala” bilder. För utskrift av bilder med ”fotografisk” kvalitet på blankt papper bör man använda en skrivare med både mättade och omättade bläckfärger. Även sådana skrivare kan köpas förhållandevis billigt (< 2000:-). Den blanka pappersytan gör att bläcket inte sugas in i papperet och sprids ut åt sidorna. Detta har två fördelar. Dels blir bläckfläckarna mindre, vilket förbättrar bildskärpan. Dels kommer färgmättnaden att bli högre eftersom bläcket inte ”blandas ut” med vita pappersfibrer. Vid betraktning på normalt avstånd (ca. 25 cm) kan bläckstrålebilder mycket väl kan mäta sig med kvalitén hos de pappersbilder som fås med klassiskt fotografiskt papper (ljuskänslig emulsion på pappersbas). Om man däremot betraktar bilderna på nära håll, t.ex. med ett förstoringsglas, så ser man att bläckstrålebilderna innehåller betydligt färre detaljer än den klassiska fotografiska papperskopian.

23.6. Andra typer av skrivare för fotografiska bilder

Förutom bläckstråleskrivare finns även andra typer av skrivare för fotografiska bilder. Två typer av laserskrivare ska här kort beskrivas.

Laserskrivare (eng. laser printer) används som beteckning på två olika typer av skrivare. Den vanligaste (och billigaste) fungerar ungefär som en laserskrivare för

svartvita datorutskrifter. Men istället för bara svart toner (pigmentpulver), finns i laserfärgskrivaren också toner i färgerna gult, magenta och cyan. Den grundläggande funktionsprincipen för dessa laserskrivare är densamma som i en fotokopiator. En laserstråle sveper radvis över en trumma uppladdad med statisk elektricitet. På de ställen som träffas av laserstrålen sker en urladdning, och på dessa ställen kommer sedan toner att fästa. Genom att ett vitt papper rullas mot trumman överförs tonermönstret till papperet. Slutligen smälts tonern fast i papperet med hjälp av ett värmeelement. Tonerpulvret består av små färgpartiklar inbakade i plastpartiklar. I färgfallet kommer den beskrivna processen att ske i fyra steg, ett för varje färgkomponent. Bildkvaliteten är i allmänhet något lägre än för de övriga skrivare som beskrivs i detta kapitel. Detta beror både på att dpi-talet (märk väl inte ppi-talet) är ganska lågt, vanligen 600, och att bara mättade färger kan tryckas. Därför ser bilderna vid närmare betraktande ganska prickiga ut, och gråtonsåtergivningen är inte den bästa. För många typer av färgbilder (t.ex. med mättade färger) kan detta vara acceptabelt, men för fotografiska utskrifter räcker det i nuläget inte riktigt till.

Den andra typen av laserskrivare är en betydligt större och dyrare apparat, som ofta används av professionella fotolaboratorier. I denna typ av skrivare ritas en bild upp rad för rad på ett fotografiskt papper innehållande tre ljuskänsliga emulsionsskikt. Detta sker genom att laserstrålar i de tre delfärgerna rött, grönt och blått snabbt sveper över fotopapperet. För varje bildpunkt som exponeras, kan man variera styrkan på de tre laserstrålarna så att färg- och gråtoner blir de önskade. Efter exponeringssteget framkallas fotopapperet kemiskt i en automatisk framkallningsmaskin som är integrerad med skrivaren. Typiska prestanda för denna typ av laserskrivare är en pixeltäthet av 400 ppi, och en utskriftshastighet av 15 kopior per minut. Kvaliteten på kopiorna är mycket hög, och helt i klass med vad man kan åstadkomma med klassisk ”analog” fotografi. Jämfört med bläckstrålebilder, ser man vid förstoring en bättre upplösning och mindre påtaglig pixellering.

24. Automatik

24.1. Inledning

På den ”gamla goda” tiden fanns vare sig exponeringsmätare eller någon exponerings- eller fokuseringsautomatik inbyggd i kamerorna. Ett bra resultat krävde därför erfarenhet hos fotografen. Men under svåra eller ovanliga fotograferingssituationer kanske inte ens en erfaren fotografs kunskaper räckte till, och i sådana lägen var man tvungen att ta ett stort antal exponeringar med olika inställningar och hoppas att någon blev bra. För att minska osäkerheten vad gäller exponeringen började man använda lösa, handhållna exponeringsmätare. Detta underlättade situationen, speciellt när det blev allt populärare att fotografera på färgdiafilm vilken har mindre slarvmarginal för felexponering än negativfilm (på grund av högre kontrast). När kompakta småbildskameror (typ ”turistkamera”) och spegelreflexkameror blev alltmer populära började man snart bygga in exponeringsmätaren i kamerahuset. I fallet spegelreflexkamera gick man också över till strålgångsmätning genom objektivet. Exponeringsmätaren gör fotograferandet lättare, men fortfarande tar det tid att göra inställning av bländare och exponeringstid enligt mätarens rekommendation. Därför kan det hända att man missar bilder när det gäller att snabbt få fram kameran och knäppa (t.ex. om ett flygande tefat plötsligt dyker upp för att sedan snabbt försvinna igen). Många ovana fotografer förstår inte heller vad tids- och bländarinställning innebär, och skulle helst vilja slippa göra några inställningar över huvud taget. För att underlätta fotograferandet har det därför under årens lopp blivit allt vanligare med automatiska kameror. Först infördes exponeringsautomatik, och senare även autofokus. Digitalkameror är naturligtvis i hög grad automatiserade både vad gäller exponering och fokusering. Det är dock värt att komma ihåg att automatiken inte fungerar bra i alla situationer. Därför är det fördelaktigt om man manuellt kan göra alla inställningar för att få full kontroll på fotograferandet. I detta kapitel ska vi titta lite på olika typer av automatik som förekommer, samt den grundläggande funktionen hos dessa system. Vi börjar med exponeringsautomatiken.

24.2. Exponeringsautomatik

Kameror med exponeringsautomatik ställer automatiskt in exponeringstid och/eller bländartal med utgångspunkt från vad exponeringsmätaren mäter upp. Man kan särskilja tre grundläggande typer, och ofta kan man i en och samma kamera välja vilken typ man vill använda.

Första typen kallas **bländarförval** (eng: aperture-priority mode). Den går ut på att man manuellt ställer in önskat bländartal, varefter kameran automatiskt väljer lämplig exponeringstid. Andra typen är **tidsförval** (eng: shutter-priority mode). Här ställer man in önskad exponeringstid, varefter kameran väljer lämpligt bländartal. Tredje typen är **helautomatik** (eng: programmed mode). Här väljer kameran automatiskt både bländarinställning och exponeringstid enligt någon (förhoppningsvis) intelligent regel. Många kameror har i helautomatisk mod flera olika program att välja på, t.ex. ”depth”, ”action” och ”portrait”. I dessa olika program optimeras val av inställningar så att man får stort skärpedjup (hög

bländartal), liten rörelseoskärpa (kort slutartid) eller litet skärpedjup (lågt bländartal). Den helautomatiska moden är naturligtvis bekväm att använda, men om man vill ha god kontroll över skärpedjup eller rörelseoskärpa bör man välja någon av de två första moderna.

Det man ska komma ihåg vid all användning av exponeringsautomatik, är att kameran utnyttjar exponeringsmätarens värden för att göra inställningar av bländartal och/eller exponeringstid. Detta innebär att man kan råka ut för felexponerade bilder om motivet är mycket ljust eller mörkt (se kap. 11.2). Vill man ha fullständig kontroll över exponeringen bör man använda histogram (kap. 11.3), men då är man tvungen att arbeta i manuell mod.

24.3. Autofokus

Förutom rätt exponering så krävs korrekt fokusering för att få en bra bild. För att förenkla fotograferandet är därför många kameror försedda med såväl exponeringsautomatik som autofocus. Vad gäller digitalkameror så är autofocus standard, men man kan på lite mer avancerade kameror även använda manuell fokusering om så önskas. Genom kombinationen autofocus och exponeringsautomatik kan man förverkliga ”point-and-shoot” målsättningen, dvs. det enda fotografen behöver göra är att välja motiv och trycka på avtryckaren (ev. kan man behöva välja lämplig zoom-inställning på objektivet också). Precis som för exponeringsautomatiken så finns det olika principer för autofocus. Man kan dela in autofocus i två olika grupper, **kontrastmätning** och **fasdetektion**, beroende på vilken princip som används.

Den första tekniken, kontrastmätning, används i digitalkameror som inte är av spegelreflex typ. Principen illustreras i Fig. 24.1. Belysningsfördelningen längs en linje i sensorplanet registreras. Pixlarnas ljusvärden överförs till en mikroprocessor som analyserar hur mycket intilliggande pixelvärden avviker från varandra. Ju större genomsnittlig avvikelse vi har mellan grannpixlar, desto skarpare bild kan vi förvänta oss. Genom trial-and-error söker kameran upp den avståndsställning som ger maximal avvikelse. För att göra avståndsställningen säkrare, analyseras flera pixelrader utplacerade på olika ställen i sensorplanet och med olika orienteringar.

Ofta kan man ställa in ett makroläge för autofokuseringen. I detta läge utför kameran sökning efter maximal avvikelse mellan grannpixlar inom ett större avståndsområde än i standardläget. Detta gör att kameran hittar korrekt fokus även vid närbildsfotografering. Nackdelen med makroläget är att fokusinställningen kan ta längre tid, eftersom kameran kommer att avsöka ett större område.

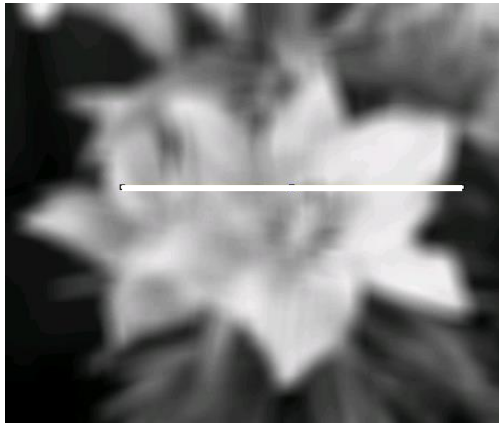
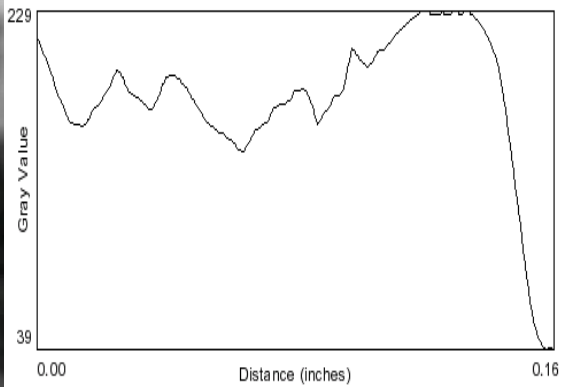
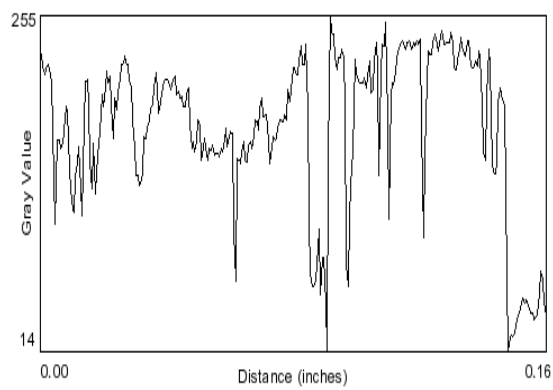
**Felaktig fokusinställning****Intensitetsprofil längs vita linjen****Korrekt fokusinställning****Intensitetsprofil längs vita linjen**

Fig. 24.1. Autofokusering genom kontrastmätning. Som optimalt fokus väljs den avståndsställning som maximerar genomsnittliga intensitetsavvikelsen mellan grannpixlarna längs en linje.

Den andra tekniken för autofokusering, fasdetektion, används i spegelreflexkameror och vissa andra kameror. En speciell autofokusenhet sitter placerad mellan den rörliga spegeln och ridåslutaren, se Fig. 24.2. En del av infallande ljuset passerar genom huvudspeglens halvgenomskinliga centraldel, och reflekteras ner mot kamerahusets botten där autofokusenheten är placerad. I den mycket förenklade principskissen i Fig. 24.2 b) består autofokusenheten av två små linser symmetriskt belägna på ömse sidor om optiska axeln. Till varje lins hör en liten elektronisk sensor som typiskt bara är en eller några få pixlar bred, och som har en blygsam längd. På grund av smålinsernas storlek och placering, kommer varje lins bara att träffas av ljus som har passerat en liten yta i kameraobjektivets periferi. Smålinserna avbildar kamerans korrekta bildplan på autofokus-sensorernas plan. Beroende på om kameraobjektivet avbildar motivet i korrekt bildplan eller inte, så kommer bilderna på autofokus-sensorerna hamna på lite olika ställen (heldragen respektive streckad strålgång). Genom att jämföra bildernas lägen på de bägge sensorerna, kan man därför avgöra om fokuseringen är korrekt eller ej. Om den inte är korrekt, kan man även avgöra åt vilket håll objektivet ska omfokuseras och hur mycket. Detta gör fasdetektionsmetoden

snabbare än kontrastmetoden, som ju arbetar efter metoden trial-and-error tills optimal fokusering har hittats.

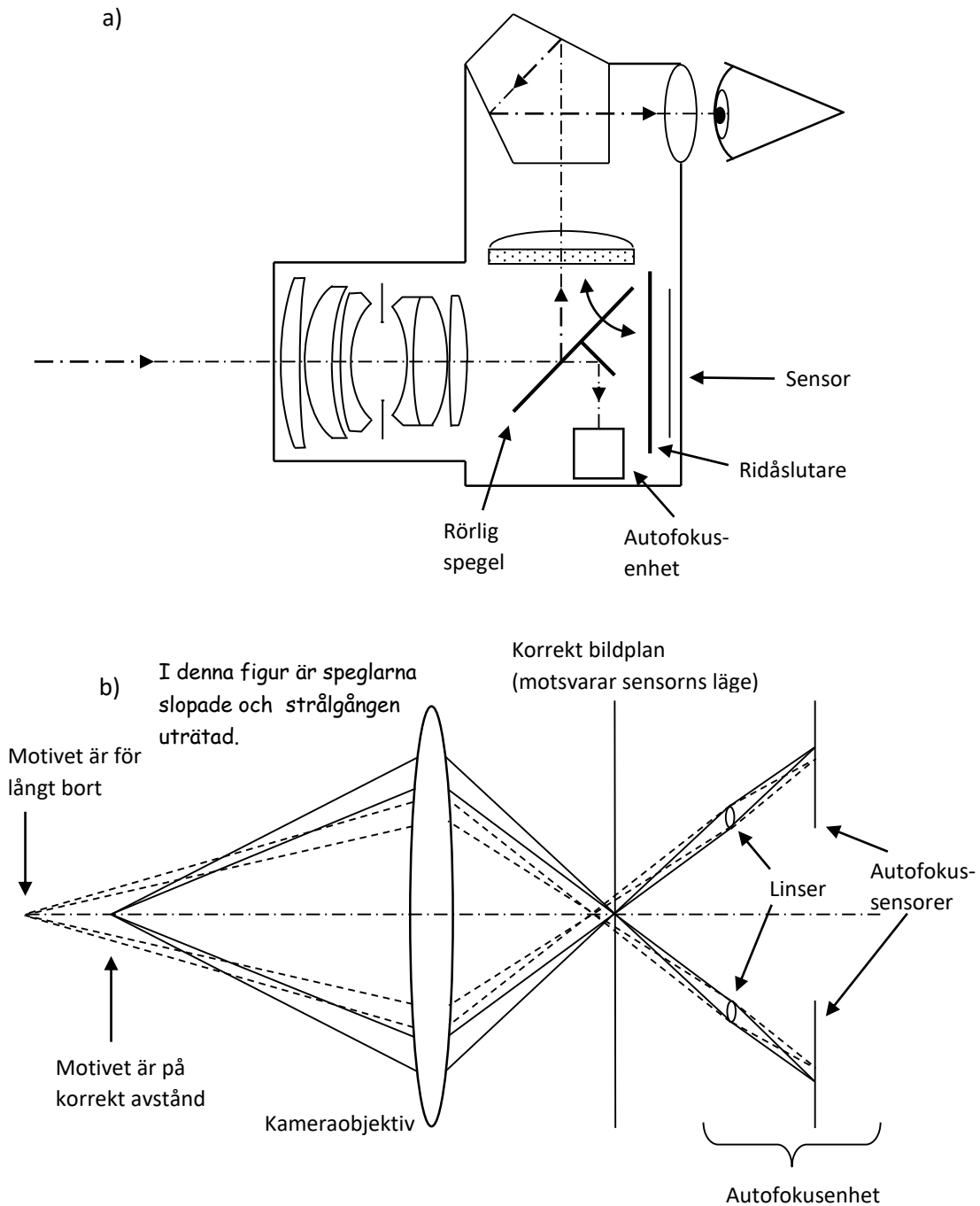


Fig. 24.2. Autofokus i spegelreflexkameror. a) Den rörliga spegeln är halvgenomskinlig i centrum, och har en extra spegel på baksidan som reflekterar ljus mot en autofocusenhet i kamerahusets nederdel. b) Autofokusenheten innehåller två linser och två elektroniska sensorer. Beroende på om motivet befinner sig i fokus eller ej, kommer bilderna på de två autofocus-sensorerna att hamna på olika ställen. Linserna har stort skärpedjup, vilket gör att bilderna på autofocus-sensorerna blir tillräckligt skarpa även när kameraobjektivet är felfokuserat. Se texten för detaljer.

Anledningen till att man behöver ett symmetriskt system med två smålinser och två sensorer, är att motivet i verkligheten inte består av en punktkälla på optiska axeln såsom i figuren. Detta innebär att man måste tillämpa en lite mer sofistikerad metod för att bestämma om bildens position längs autofokus-sensorn är korrekt. Det som görs i verkligheten är en beräkning av autokorrelationsfunktionen mellan de bägge sensorernas bilder, där man utgår från en nollpunkt längs sensorerna som svarar mot korrekt fokus. Autokorrelationsfunktionen blir maximal då de två bilderna överlappar perfekt, vilket svarar mot korrekt fokus. Ur denna autokorrelationsfunktion kan man få upplysning om hur kameraobjektivet ska omfokuseras för att slutbilden ska bli skarp. Tilläggas kan att den beskrivning som gjorts av fasdetektionsmetoden är förenklad jämfört med ett verkligt system, som innehåller betydligt flera komponenter. Man har t.ex. flera autofokus-sensorer med olika orienteringar och olika placeringar i bildfältet. Detta för att man ska klara av många olika typer av motiv.

Autofokus har svårt att klara av motiv med mycket låg kontrast, t.ex. en enfärgad vägg eller en nästan jämngrå himmel. En annan nackdel är att den kräver att motivet sänder ut ljus för att göra fokusinställningen. Om man till exempel vill ta en blixtbild i dålig omgivningsbelysning, klarar kameran inte av att göra fokusinställningen innan blixten brinner av. Det finns dock kameror som sänder ut ett randigt ljus mot motivet för att ge autofokusmekanismen något att fokusera på i sådana situationer. Alternativt kan den inbyggda blixten ge ett antal provblixtar innan bilden tas för att möjliggöra autofokusering.

Generellt kan man säga att autofokussystemen har blivit allt mer sofistikerade med tiden. Resultatet är att i flertalet fall fungerar de alldeles utmärkt, och de låter inte lura sig så lätt. Detsamma gäller för exponeringsautomatiken. Men man måste komma ihåg att dessa system aldrig kan förstå vad fotografen vill framhäva i sin bild. Exakt vilka motivdetaljer ska ligga i fokus? Hur stort ska skärpedjupet vara? Vilka motivdelar ska vara optimalt exponerade, etc.? Genom fullständig manuell kontroll av fokus, bländare och exponeringstid, kan en erfaren fotograf erhålla resultat som inte går att få med automatiska inställningar. För en kamera som ska användas till annat än standardmässiga bilder, är det därför viktigt att kräva att den ska gå att använda i manuell mod där man har full kontroll över dessa inställningsmöjligheter.

24.4. Bildstabilisering

Vid långa exponeringstider kommer man vid frihandsfotografering att få problem med skakningsoskärpa i bilderna. Det finns en gammal regel som säger att man på fri hand inte bör använda en längre exponeringstid i sekunder än reciproka värdet på brännvidden i millimeter. För exempelvis ett 60 mm objektiv innebär detta att man inte bör använda längre exponeringstid än 1/60 sekund vid fotografering på fri hand. Denna regel gäller för fullformatskameror (24 mm x 36 mm sensor). För kameror med andra sensorstorlekar behöver man räkna om den verkliga brännvidden till "135 mm equivalent" (se kap. 6.6).

För att tillåta frihandsfotografering vid längre exponeringstider, är vissa digitalkameror utrustade med en finess som kallas bildstabilisering. Man kan då

använda tider som i bästa fall är upp till en tiopotens längre än enligt ovanstående regel. Detta är naturligtvis mycket användbart i situationer då det är ont om ljus. Det finns olika metoder att åstadkomma bildstabilisering, men inom stillbildsfoto brukar man antingen använda en rörlig sensor eller rörliga linselement i objektivet. För att kunna utföra en skakningskompensation måste man kunna mäta upp den momentana vibrationsrörelsen. Detta kan ske antingen genom rörelsedetektorer i kameran, eller genom att analysera bilden som faller på sensorn. Både metoden med rörliga linser och rörlig sensor fungerar bra vad gäller bildstabilisering. En fördel med rörlig sensor är att bildstabiliseringen är inbyggd i kamerahuset. Det ställs alltså inga speciella krav på vilka objektiv som kan användas. I andra fallet måste varje objektiv som används vara utrustat med bildstabiliseringsutrustning. Om man vill kunna använda ett flertal objektiv på kameran är alltså rörlig sensor det mest ekonomiska alternativet.

Man bör komma ihåg att bildstabilisering kompenserar enbart för att hela kameran skakar. Den kompenserar inte för den rörelseoskärpa som uppkommer därför att motivet rör sig.

25. Termovision (värmekamera)

Det är i många sammanhang värdefullt att kunna få en bild av temperaturfördelningen på olika föremål. Om föremålet har en hög temperatur, t.ex. $\approx 1000^\circ\text{C}$, så kommer det att sända ut tillräckligt med elektromagnetisk strålning i det synliga och nära infraröda området för att kunna registreras med fotografiska metoder. Om föremålet däremot har ungefär rumstemperatur (300 K) så kommer praktiskt taget ingen strålning att sändas ut inom våglängdsområden som är användbara för fotografi. Vi kan se i svartkroppskurvorna i Fig. 13.1 att vi då behöver registrera strålning med våglängder någonstans i området 5-25 μm . Detta är inte möjligt med fotografisk film eller med de matrissensorer som används i digital fotografi. En ytterligare svårighet med att jobba i detta långvågiga IR-område är att vanliga glaslinser inte släpper igenom denna strålning.

Under årens lopp har olika varianter av termovisionskameror sett dagens ljus. I början användes kameror med ett enda detektorelement, vilket kylades till mycket låga temperaturer (t.ex. med flytande kväve) för att öka känsligheten. För att få en bild så utfördes en mekanisk scanning. Istället för glasoptik så användes IR-reflekterande speglar eller linser av germanium (vilket har god IR-transmission). Dessa kameror var tunga och klumpiga, samt ömtåliga och dyra. Dessutom hade de nackdelen att de måste förses med flytande kväve eller något liknande köldmedium. På senare år har det skett en snabb utveckling på termovisionsområdet. Precis som för den digitala fotografien, har man nu lyckats framställa matrissensorer som är känsliga inom det långvågiga IR-området (även om antalet pixlar är betydligt lägre). Detta gör att man inte längre behöver använda någon optomekanisk scanning för att läsa in bilder. Kamerorna kan därigenom göras både mindre och billigare.



Image courtesy of SPI Corp, www.x20.org

Fig. 25.1. Termovisionsbild av en hund. Olika temperaturer återges med olika gråtoner enligt skalan till höger ($100^\circ\text{F} = 38^\circ\text{C}$). Notera den höga temperaturen på ögon och tunga. Nosen är däremot kall.

Matrissensorer för termiska IR-området kan fungera på två helt olika sätt:

Variant 1: Denna variant bygger på samma grundprincip som matrissensorer för synliga området (se kap. 14). Sensorn detekterar alltså enskilda fotoner och därvid frigörs elektroner som kan samlas upp och mätas. Eftersom långvågiga IR-fotoner är mycket energifattigare än synliga fotoner, måste detektormaterialet (t.ex. HgCdTe) innehålla elektroner som har betydligt lägre bindningsenergi än vad som är fallet för detektorer för synligt ljus. En konsekvens av detta är att IR-detektorn blir värmekänslig, eftersom termiska vibrationer i materialet lätt kan slå loss elektroner. Denna typ av detektor kräver därför kylning ner till ca. -200°C , vilket naturligtvis är en nackdel. En fördel är att denna detektorvariant är mycket känslig, vilket innebär att den kan detektera små temperaturdifferenser. Kameror med denna typ av detektorer är dyra, och de används därför mest i krävande tillämpningar (t.ex. inom militären).

Variant 2: Denna variant detekterar inte enskilda fotoner, utan känner av den temperaturökning som förorsakas av att termisk strålning absorberas av detektormaterialet. Dessa så kallade mikrobolometer-matriser används ofta i termovisionskameror för enklare övervakning. Det kan gälla t.ex. att detektera värmeläckage i husväggar, inbrottstjuvar (de fungerar i komplett mörker), överhettning i kraftledningar mm. Dessa kameror kan typiskt ha matriser som består av 640×480 detektorelement, vardera med en storlek av $20 \mu\text{m} \times 20 \mu\text{m}$. Materialet i detektorelementen är ofta vanadiumoxid, vilket har en hög temperaturkoefficient för resistansen ($2\text{--}3\%/K$). Genom att känna av de individuella elementens resistans får man en bild av temperaturskillnader i det avbildade motivet. Dessa detektorer kräver ingen kylning, men det behövs korrektion för variationer i detektorns omgivningstemperatur. Kamerorna är billigare än variant 1, och känsligheten är lägre. Dock kan objekttemperaturer i området -20 till $+250$ grader Celsius typiskt registreras med en upplösning av 0.1°C , vilket ofta är tillräckligt. Typiskt kan bilder registreras med videofrekvens.



Image courtesy of SPI Corp, www.x20.org

Fig. 25.2. Hudens temperatur styrs i hög grad av blodgenomströmningen. I t.ex. handen syns blodkärlen tydligt på grund av att de har en högre temperatur.

Eftersom glaslinser inte transmitterar våglängder över ca. 2-3 μm så är man tvungen att använda andra optiska material i termovisionskameror. Såsom redan nämnts är halvledarmaterialet germanium ett alternativ. Zinkselenid är ett annat användbart alternativ. Dessa material har god transmission i det långvågiga IR-området, och de används ofta som linsmaterial. Brytningsindex är mycket högt, ca. 4, och dispersionen är låg, vilket är gynnsamt för objektivtillverkning. Det höga brytningsindex gör dock att ytreflexerna blir mycket kraftiga (> 50% i en enkel lins). Därför är det mycket viktigt med antireflexbehandling för denna typ av optik. För att få god känslighet för termovisionskameran krävs också att objektivet är ljusstarkt, ett typiskt värde kan vara ca. 1.0. Detta kräver asfäriska ytor för att inte sfäriska aberrationen ska bli för stor, se kap. 15. Tillverkningen av sådan optik blir emellertid dyr, eftersom man använder diamantsvarvning för att få rätt form på linsytorna. Även med små avbildningsfel kommer emellertid upplösningen att bli betydligt sämre än för synliga våglängder. Detta beror på att diffraktionen blir stor på grund av den långa våglängden. Därför har termovisionsbilder ett ”softat” utseende, vilket gör att detektormatriser med relativt få element är tillfyllest.

Det finns många användningsområden för termovision. Bland de vanligaste kan nämnas:

- *Upptäckt av ”hot spots” i maskiner och motorer, vilket kan vara ett förebud om fel och kommande haverier.*
- *Upptäckt av värmeläckage genom husfasader på grund av bristfällig isolering.*
- *Upptäckt av människor och djur, t.ex. om någon gått vilse i terrängen eller vid inbrottsövervakning. (Polishelikoptrar är utrustade med värmekamera.)*
- *Inom medicinen för att se hudområden med avvikande temperatur.*

Referenser

Allen, E. & Triantaphillidou, S. Eds., "Manual of Photography", Focal Press, 10:th ed. 2009

Gernsheim, H. & Gernsheim, A., "The History of Photography", Oxford University Press, 1955.

Ray, S.F., "Applied Photographic Optics", Focal Press, London, 1988.

Carlsson, K., "Imaging physics", KTH, 2016, se
<https://www.kth.se/social/files/57d11da0f276542d9216413a/Comp.ImagingPhysics.new.pdf>

www.dpreview.com Website som listar kameror, tillbehör etc. tillsammans med beskrivningar, tester, forum mm.

www.dxomark.com Website som publicerar kvalitetstester på foto- och videoutrustning.

<http://www.photonstophos.net/> Website som listar en stor mängd mätningar gjorda på digitalkameror.

<https://www.kth.se/social/files/542d2d2df276546ca71dffaa/Pinhole.pdf> redovisar datorsimuleringar av avbildningskvaliteten för hålkameror med olika håldiametrar.

<https://www.kth.se/social/files/5c8675e856be5bdc4e5c4033/Pinhole.camera.pdf> innehåller praktiska mätningar av bilkvaliteten på hålkameror, samt anvisningar för bygge av en digital hålkamera.

<https://www.kth.se/social/files/5c86765456be5bdfb084187e/MTFm%C3%A4tningar.pdf>. innehåller information om freeware för mätning av MTF på digitalkameror, samt resultat av mätningar.

<https://www.kth.se/social/files/5c88c38056be5bb29067e3b0/Tonreproduktion.Sony%20A7.pdf> redovisar mätningar av tonreproduktionen (RAW + jpeg) för digitalkameran Sony A7.

Index

A

Aberrationer, 157
 ADC, 128
 Additiv färgsyntes, 80
 AdobeRGB, 90
 Airy-fläcken, 166
 Akromatobjektiv, 159
 Aliasing, 250
 Analog/digital-omvandling, 110, 128
 Anti-aliasing filter, 252
 Antireflexbehandling, 33
 Antireflexskikt, 63
 APS-sensor, 104
 Asfäriska linsytor, 157, 285
 Astigmatism, 161
 Ateljékamera, 18
 Atmosfärisk inverkan, 95
 Autofokus, 278
 Automatik, 277
 Avbildningsfel, 157
 Avbildningsskala, 25
 Avfältning, 211
 Avståndsdisk, 226

B

Back illumination, 124
 Barriärfilter, 232
 Bayer mosaik, 136, 256
 Belysning, 53
 Bildavstånd, 24
 Bildfältskrökning, 161
 Bildförstärkare, 243
 Bildkompression, 269
 Bildkvalitet
 MTF, 175
 upplösningförmåga, 167
 Bildminneskort, 269
 Bildskärpa, 167
 Bildstabilisering, 281
 Bildvinkel, 28, 31
 Binning, 152
 Blixtaggregat, 100
 Blixtfotografering, 22, 101
 Blooming, 107
 Bläckstråleskrivare, 103, 271

Bländare, 19
 Bländarförval, 277
 Bländarsteg, 19, 56
 Bländartal, 19, 56
 Bokeh, 74
 Bolometer, 284
 Bromsstrålning (röntgen), 234
 Brus, 119
 Brytningsfel, 157
 Brännvidd, 24, 26

C

Camera obscura, 7
 Candela, 52
 CCD, 105
 Centralslutare, 20
 Centrumvågning, 76
 Charge transfer efficiency, 110
 Charge-coupled device, 105
 CIE-systemet, 85
 CMOS, 112
 Color rendering index, 97
 Contrast Transfer Function, 205
 $\cos^4$ -avfallet, 58
 CRI, 97
 CRT (Cathode Ray Tube), 125
 CTF, 205

D

Daguerrotyp, 9
 Deltafunktion, 177
 Diffraktion, 157, 165
 Diffraktionsbegränsat objektiv, 165
 Diod, 105
 Distorsion, 163
 Djupintryck, 41
 dpi (dots per inch), 274
 Dual gain, 114, 120
 Dynamisk vidd, 121, 129, 140

E

Effektiv exponeringstid, 20
 Ekvivalent brännvidd, 34
 Elektron/hål par, 106
 Elektronisk slutare, 23, 115
 Excitationsfilter, 232

Exponering, 53
 Exponeringsautomatik, 277
 Exponeringstid, 20

F

Fill factor, 105
 Fish-eye objektiv, 60
 Fixed pattern noise, 114
 Fjärrgräns, 67
 Flatpanelsdetektor, 237
 Fluorescensfotografi, 231
 Flygfotografering, 226
 Fotodiod, 111
 Fotografi (ordet), 7
 Fotometri, 50
 Fotomultiplikator, 244
 Fouriertransform, 177
 Foveon, 136
 Full frame transfer CCD, 110
 Fullformat-sensor, 104
 Färganalys, 83
 Färginterpolering, 136, 146, 208
 Färglära, 79
 Färgmättnad, 81
 Färgregistrering (sensor), 135
 Färgrymd, 90
 Färgseende, 79
 Färgtemperatur, 95
 Färgton, 84
 Färgtriangel, 84

G

Gammakorrektion, 125, 132
 Global shutter, 116
 Gränsfrekvens, 181

H

HDR, 122, 140
 Histogram, 77
 Histogram, färg, 101
 Huvudplan, 26, 160
 Hyperfokalinställning, 67
 Hål (laddningsbärare), 106
 Hålkamera, 10
 Höghastighetsfotografi, 240
 Höghastighetsvideo, 240

I

Infrarödfotografi, 225

Ingångspupill, 56
 Inkoherent, 177
 Integralmätning, 76
 Interlace, 151
 Irisbländare, 19
 Irradians, 51
 ISO-tal, 148

J

JPEG-format, 126, 128, 269

K

Kamera, 15
 Kamerahus, 15
 Kameramobil, 17
 Koherent, 177
 Koma, 163
 Kompaktkamera, 20
 Komplementfärg, 83
 Korrelerad färgtemperatur, 96
 Kromatisk aberration, 157
 Kvantverkningsgrad, 123
 Kvicksilverlampa, 231

L

Lambertreflektor, 54
 Lateral kromatisk aberration, 160
 LCD-skärm, 15
 LED-lampa, 98
 Linearitet (sensor), 125
 Linjespridningsfunktion, 190
 Linjetäthet, 171
 Linsformeln, 24
 Ljusflöde, 52
 Ljuskällor, 94
 Ljusstyrka, 32, 52, 57
 Ljusutbyte, 98
 Longitudinell kromatisk aberration, 158
 Lumen, 52
 Luminans, 52
 Lux, 53

M

Makrofotografering, 201
 Makroläge (autofokusering), 278
 Makroobjektiv, 32
 Mattskiva, 16
 MCP, 243
 Micro-channel plate, 243

Mikroskopi, 203, 231
 Mjukteckning, 221
 Mobilkamera, 17
 Modulation Transfer Function, 175
 Modulationsgrad, 179
 Moiré-effekt, 152, 251
 Motivavstånd, 24
 MPEG-format, 155
 MTF, 175
 Multiplikationsregeln (MTF), 185
 Mörkersignal, 121
 Mörkerström, 114

N

Naturliga ljusavfallet, 58
 Normalobjektiv, 30
 NTSC, 151
 Numerisk apertur, 62
 Nyquistfrekvens, 252
 Närbildsfoto, 32
 Närgräns, 67
 Näthinna, 79

O

Objektiv, 24, 30
 Objektivets ljusstyrka. *Se Ljusstyrka*
 Optisk avbildning, 24
 Optisk axel, 24
 Optiska avbildningsfel, 157
 astigmatism, 161
 bildfältskrökning, 161
 distorsion, 163
 koma, 163
 kromatisk längdaberration, 158
 lateral kromatisk aberration, 160
 sfärisk aberration, 157
 Optiskt lågpasfilter, 252
 Ortsfrekvens, 178
 Oskärpecirkel, 66
 Oskärpefläck, 167
 Oversampling, 152

P

PAL, 151
 Perspektiv, 41
 Photogate, 105
 Pixel, 105
 Plancks strålningslag, 50
 Posterization, 132

psf (punktspridningsfunktion), 177
 Punktspridningsfunktion, 177

Q

Quad Bayer, 139

R

Radians, 51
 Radiell ortsfrekvens, 182
 Radiometri, 50
 RAW-format, 102, 136, 270
 Rayleighkriteriet, 167
 Rayleighspridning, 226
 Retrofokusobjektiv, 31
 RGB-filter, 80, 136
 RGB-histogram, 101
 Ridåslutare, 21
 Rolling shutter, 115
 Rymdvinkel, 50
 Röntgenfotografi, 234
 Röntgenrör, 234

S

Sampling, 247
 Samplingfrekvens, 249
 Samplingteoremet, 249
 Scheimpflug-villkoret, 37
 SECAM, 151
 Seriebildsexponering, 150
 Sfärisk aberration, 157
 Sharpening, 195, 214
 Signal/brus förhållande, 122
 Skakningsoskärpa, 281
 Skärmbildsfotografering, 236
 Skärpa, 157, 167
 Skärpedjup, 65, 105
 Skärpedjupsskala, 67, 71
 Slanted edge (MTF), 190
 Slutare, 19
 Snedperspektiv, 37
 SNR (Signal-to-noise ratio), 122
 Softfokus, 221
 Spegelreflexkamera, 15
 Spektral densitet, 83
 Spektralfördelning (ljuskällor), 95
 Spotmätning, 76
 sRGB, 90
 Stereobas, 46
 Stereofotografi, 46

Stereoseende, 46
 Strobofoto, 240
 Stroboskop, 243
 Strålgångskonstruktion, 24
 Ströljus, 63
 Störtande linjer, 35
 Subtraktiv färgsyntes, 82
 Svartkroppsstrålare, 50
 Synvinkel, 41
 Systemkamera, 17, 30

T

Tangentiell ortsfrekvens, 182
 Teleconverter, 38
 Teleförlängare, 38
 Teleobjektiv, 31
 Temperaturstrålare, 94
 Termovision, 283
 Testmönster (upplösning), 169
 Tidsförval, 277
 TIFF-format, 128
 Tilt/shift-objektiv, 34
 Tripletobjektiv, 162
 Tunn lins, 24

U

Ultraviolettfotografi, 230
 Upplösningsförmåga, 167
 Utarmningsområde, 106
 UV-fotografi, 231

V,W

Well capacity, 107, 118
 Video, 150
 Vidvinkelobjektiv, 31
 Vikning, 250
 Vinjettering, 60
 Vitbalans, 95, 101
 Värmekamera, 283

Y

YMC-filter, 82

Z

Zoomobjektiv, 32

Ö

Ögat (spektral känslighet), 52
 Ögonbas, 46