

Om färginterpolation och sampling

I kap. 22.7 i fotokompndiet¹ utreds hur höga ortsfrekvenser som kan registreras i de tre färgbanden RGB för en sensor med Bayer mosaik, Fig. 1. Vi ska i denna lilla skrift utvidga det hela lite genom att titta på vad detta medför rent praktiskt för färgbilder där man utfört färginterpolation. Vi ska också kort beröra olika strategier för färginterpolation, samt illustrera det hela med autentiska bildresultat. Den som vill skippa ”mellansnacket” kan gå direkt till slutsatsen som finns i slutet.

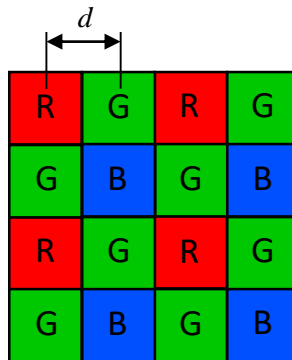


Fig. 1. Färgfiltrens placering i en sensor med Bayer mosaik.

Layouten av de färgade pixlarna i Fig. 1 är densamma för rött och blått, men skiljer sig för de gröna pixlarna. Det medför att den högsta ortsfrekvensen som kan registreras utan moiré skiljer sig för det gröna färgbandet jämfört med de blå och röda. Dessutom kommer mönsterorienteringen att spela in. Dessa saker finns utredda i kompendiets kap. 22.7, och summeras i tabellen nedan.

Tabell 1.

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	$\frac{1}{2d}$	$\frac{1}{2\sqrt{2}d}$

Låt oss göra ett litet experiment för att se hur väl detta tycks stämma med verkligheten. En bild av streckmönster med olika orientering fotograferades med en Sony A7R (sensor med 7392 x 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 x 2460 pixlar. En detalj av denna bild visas i Fig. 2a. Sedan interpolerades denna rödbild upp till dubbel bredd och höjd (7392 x 4920) med hjälp av bikubisk interpolation (vilket är en hyggligt bra metod). En detalj av den interpolerade bilden visas i Fig. 2b.

¹ <https://www.kth.se/social/files/5e7c77215dc0ec59bced1be8/fotokompndium-sk1140-vt20.pdf>

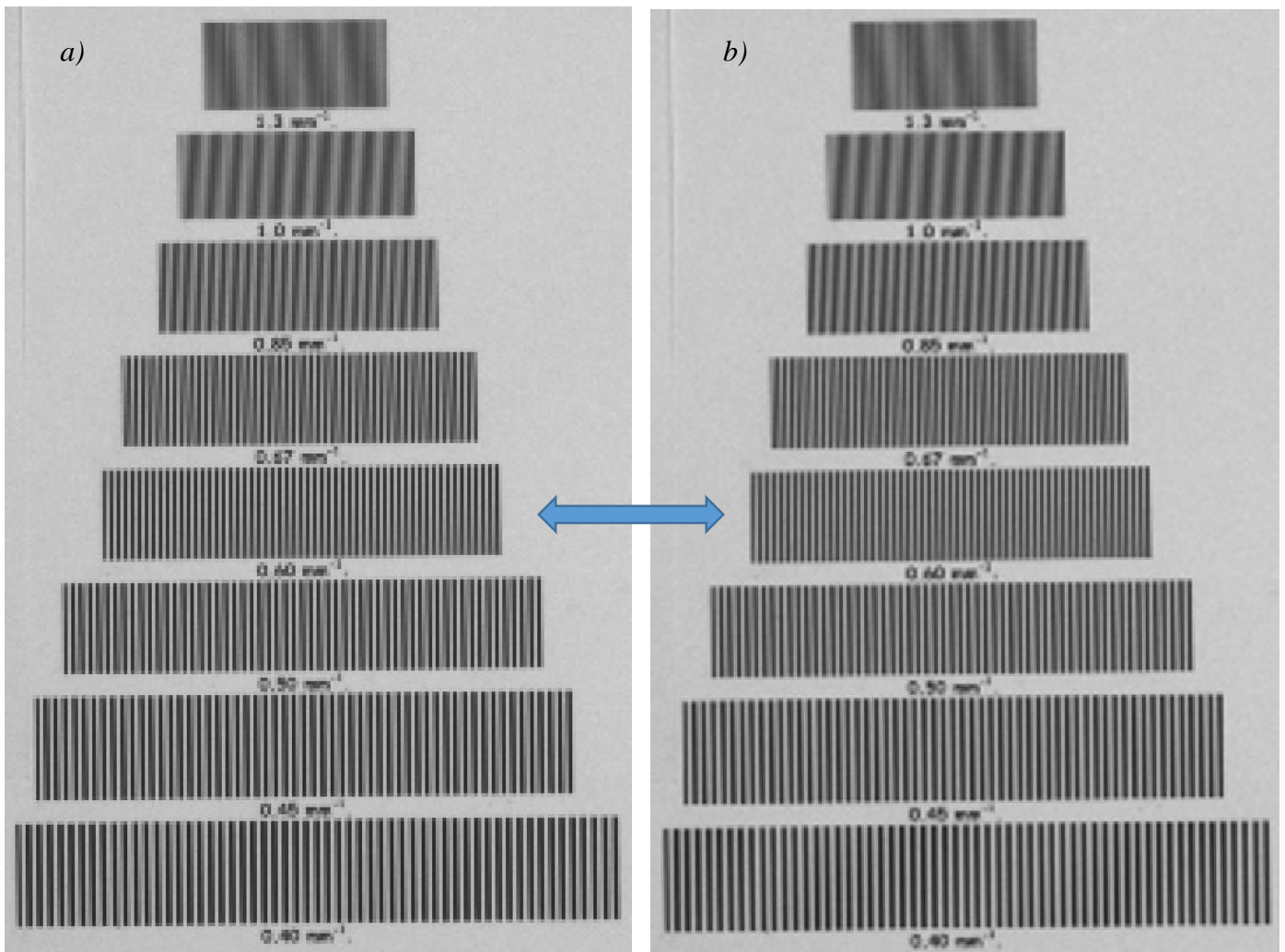


Fig. 2. Bilder uppbyggda av enbart röda pixlar extraherade ur RAW-filen. a) Detalj direkt ur R-bilden. b) Samma bild som a), men genom bikubisk interpolation har pixelantalet dubblerats både på bredden och höjden.

Anledningen till att RAW-data användes vid framställning av bilden i Fig. 2, 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 1), 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. 2 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 i fotokompendiet).

Låt oss nu titta på mönster med 45 graders orientering. Fig. 3 visar en detalj ur samma originalbild som Fig. 2b. 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 Tabell 1 som säger att man klarar $\sqrt{2}$ gånger så hög frekvens i 45 graders riktning jämfört med horisontell/vertikal orientering ($\sqrt{2} \times 0.60 = 0.85$).

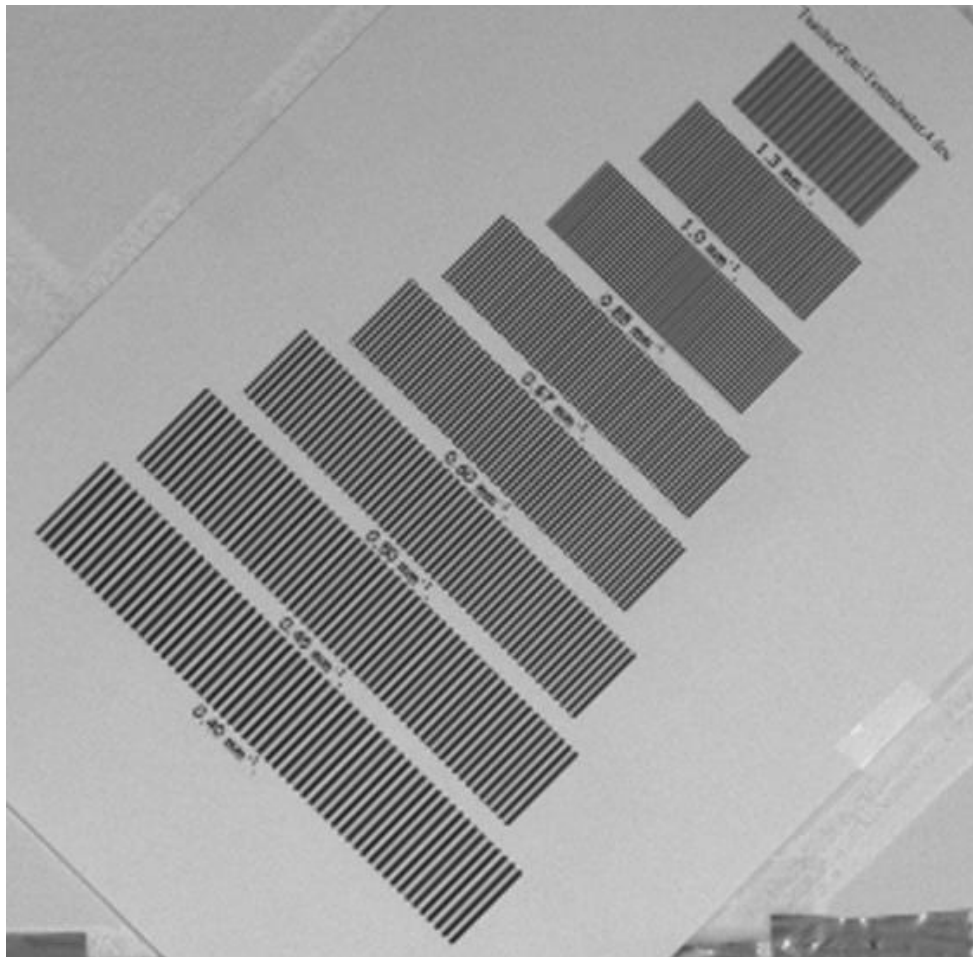


Fig. 3. Detalj ur samma interpolerade bild som Fig. 2b.

Så här långt stämmer alltså allting precis enligt läroboken, 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 pixlar². 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. 4.

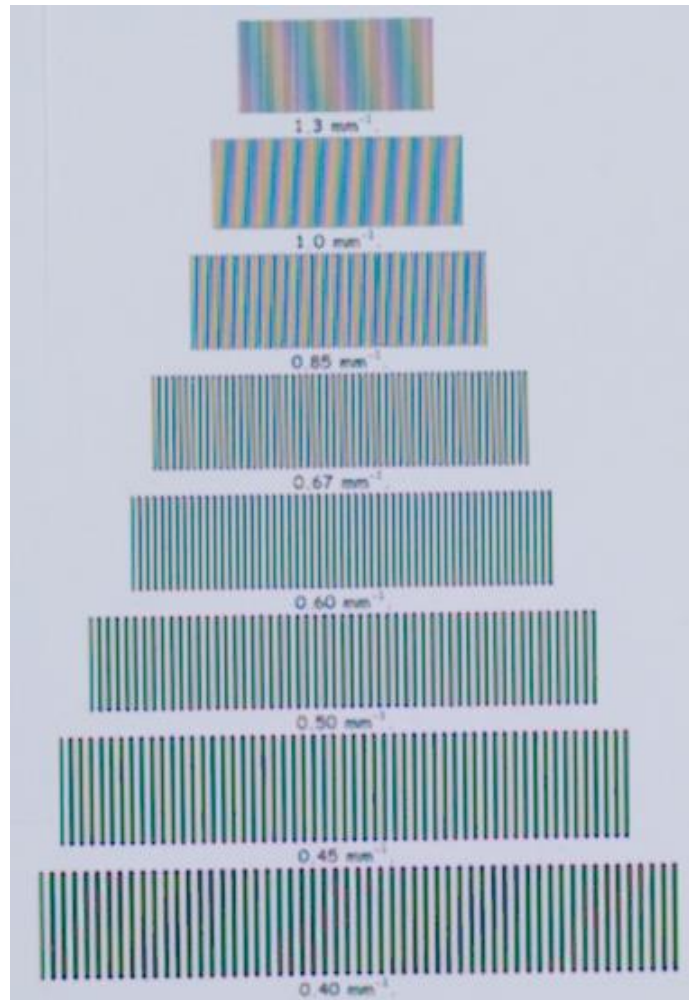


Fig. 4. Detalj ur färgbild visande samma utsnitt som i Fig. 2. Bilden framställdes ur framinterpolerade RGB-delbilder från en RAW-fil. (Demosaicing algorithm VNG4)

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 (Tabell 1). Om vi däremot betraktar mönster i 45 graders riktning, Fig. 5, så ser vi knappt några störande färgmönster alls, och vi klarar 0.85 mm^{-1} utan moiré (precis som i R-

² 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.

delbilden). Detta resultat är också naturligt, eftersom man för 45 graders orientering klarar lika höga ortsfrekvenser i alla tre färgbanden (Tabell 1).

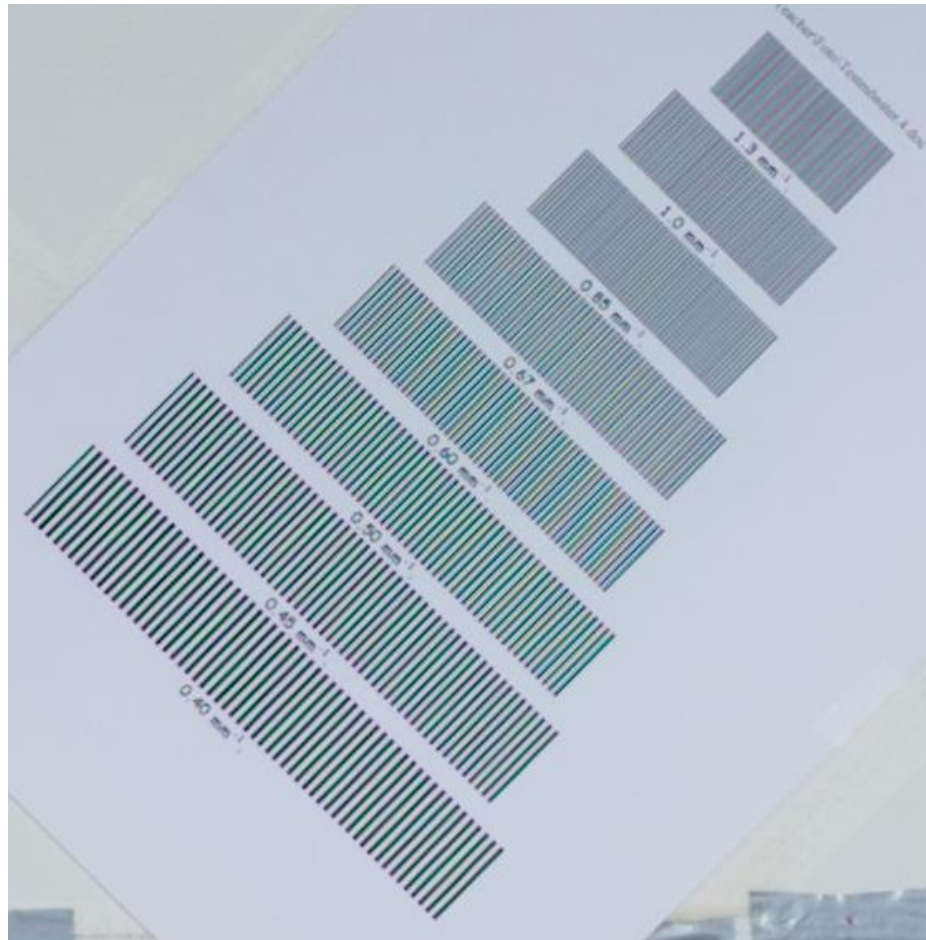


Fig. 5. Detalj ur färgbild visande samma utsnitt som i Fig. 3. Bilden framställdes ur framinterpolerade RGB-delbilder på samma sätt som Fig. 4.

Så här långt har vi inte stött på några överraskningar, allt verkar stämma enligt läroboken. Låt oss nu titta på jpeg-filen som kameran levererade av samma motiv, Fig. 6. Denna jpeg-bild har alltså av kameran räknats fram från samma RAW-data som vi har använt i Fig. 2-5. Jämfört med Fig. 2b och Fig. 4, så klarar jpeg-bilden i Fig. 6 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. 6b (R-kanalen i jpeg-bilden) korrekt återge betydligt högre ortsfrekvenser än Fig. 2b (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 ljusnivåer³.

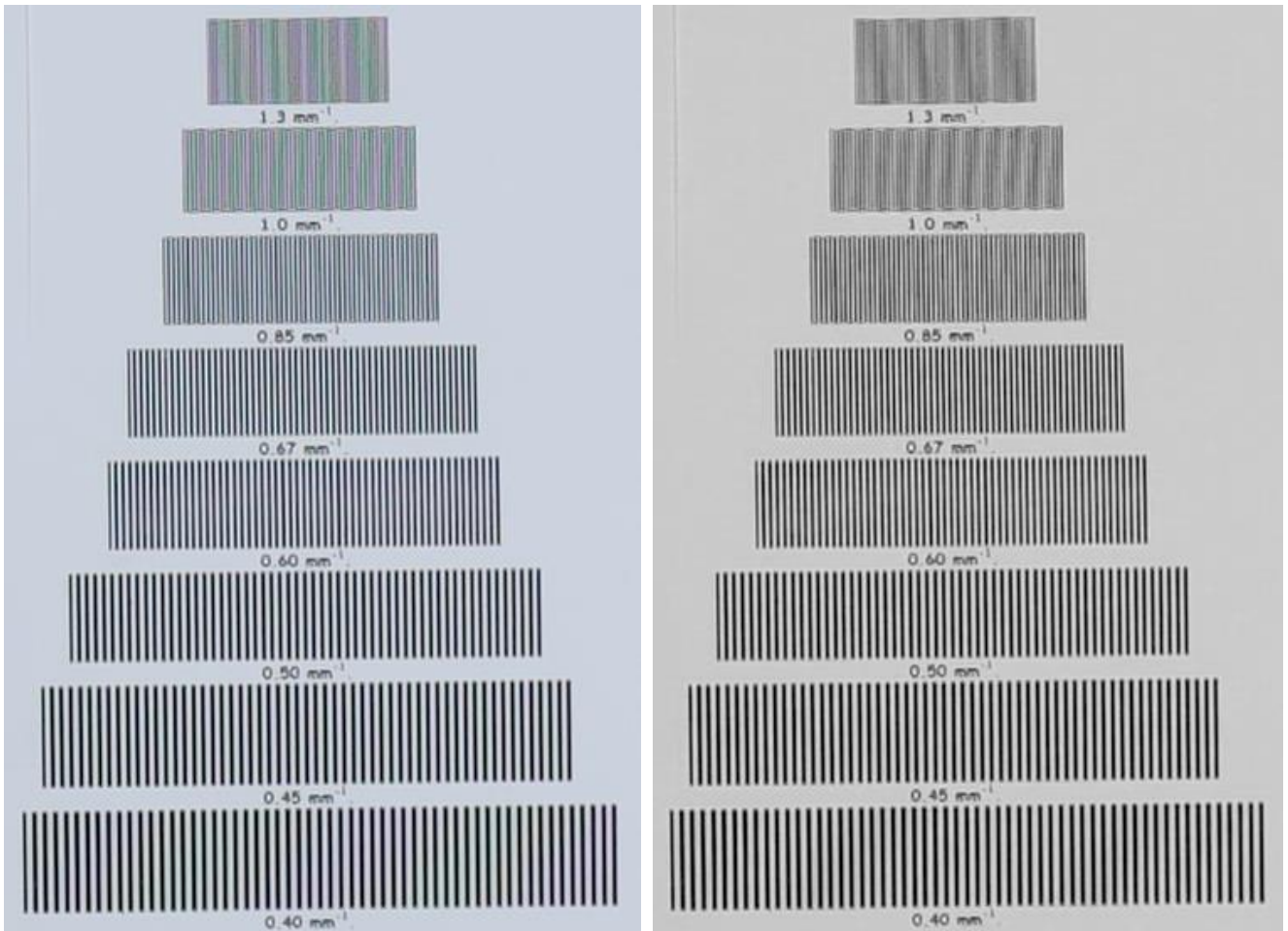


Fig. 6. Detalj ur jpeg-bilden direkt från kameran. a) Färgbild (motsvarar Fig. 4). b) Enbart rödkanalen (motsvarar Fig. 2b).

För 45 graders mönsterorientering uppvisar samma jpeg-bild som i Fig. 6 ett utseende som visas i Fig. 7. Precis som i Fig. 6 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. 3 & 5 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 olikfärgade

³ <https://www.kth.se/social/files/5c88c38056be5bb29067e3b0/Tonreproduktion.Sony%20A7.pdf>

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 ”demosaiicing” som det kallas på engelska). Detta gör att jpeg-bilderna kan få lite olika utseende. Genom att testa lite olika demosaiicing-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.

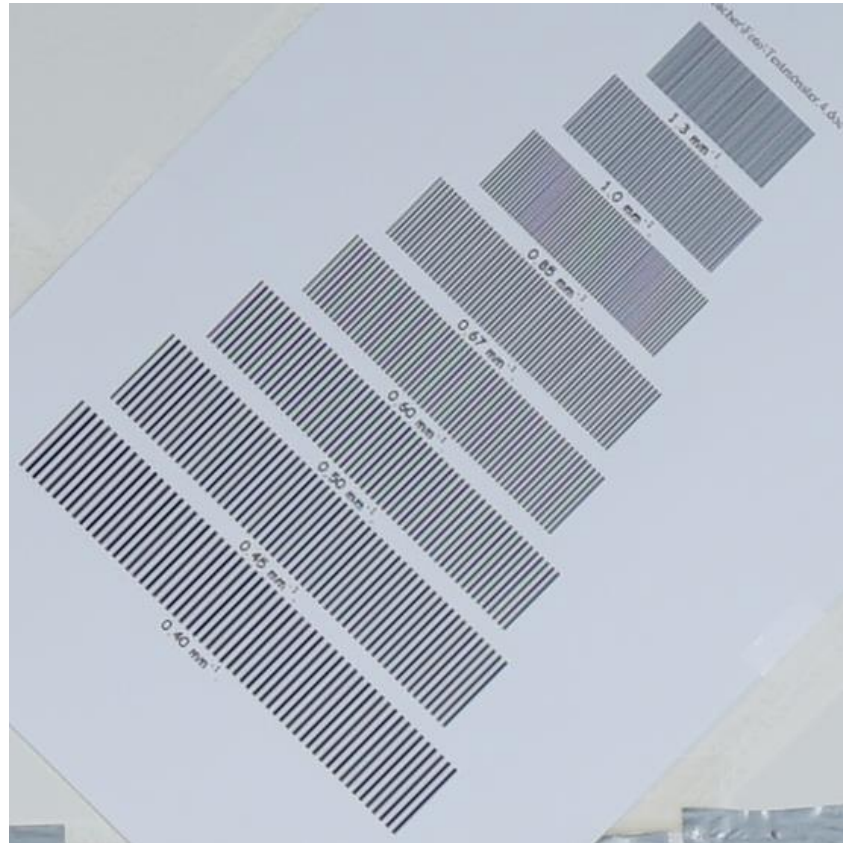


Fig. 7. Ett annat utsnitt ur samma jpeg-bild som visas i Fig. 6a.

Vi kan sammanfatta resultaten av undersökningarna på följande sätt:

1. Om färginterpolation (demosaiicing) utförs enbart mellan pixlar i samma färg, erhålls ett resultat som stämmer bra med de fysikaliska gränser som ges i Tabell 1. Det betyder att horisontella/vertikala streckmönster återges korrekt upp till ortsfrekvensen $\frac{1}{4d}$, där d = centrum-tillcentrum avståndet mellan pixlar (se Fig. 1). 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 (Fig. 1).

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.