Linear Operators and Fourier Transform

DD2423 Image Analysis and Computer Vision

Mårten Björkman

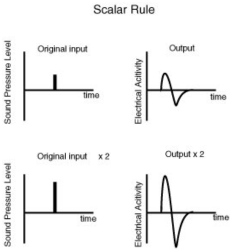
Computational Vision and Active Perception
School of Computer Science and Communication

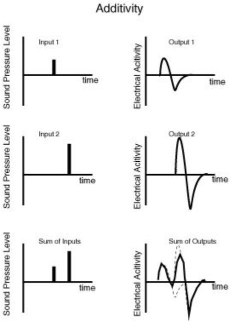
November 12, 2013

Linear Systems

Image processing operations can be modeled by utilizing linear systems theory. A linear system obeys the principle of superposition:

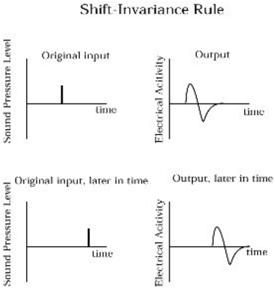
- Homogeneity (scalar rule): an increase in strength of the input, increases the output/response for the same amount.
- Additivity: if the input consists of two signals, the output/response is equal to the sum of the individual responses.

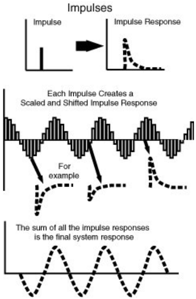




Additional properties:

- Shift-invariance: If a system is given two impulses with a time delay, the response remains the same except for time difference.
- Signals can be represented as sums of impulses of different strenghts (image intensities), shifted in time (image space).
- If we know how system responds to an impulse, we know how it reacts to combination of impulses: **impulse-response function**.





Notation

- Assume f and f' are 2D images, then $f \overset{\mathcal{L}}{\Longrightarrow} f' = \mathcal{L}(f)$,
  where $\mathcal{L}$ is an operator that "converts" the input f into the output f'.

Linear operator $\mathcal{L}$ satisfies

- Homogeneity: $\mathcal{L}(\alpha f(x,y)) = \alpha\mathcal{L}(f(x,y))$; $\alpha \in \mathbb{R}$
- Additivity: $\mathcal{L}(f(x,y)+g(x,y)) = \mathcal{L}(f(x,y)) + \mathcal{L}(g(x,y))$; $x, y \in \mathbb{R}$

Given

- $g \rightarrow \boxed{\quad\mathcal{L}\quad} \rightarrow \mathcal{L}(g)$
- $f \rightarrow \boxed{\quad\mathcal{L}\quad} \rightarrow \mathcal{L}(f)$

we have

- $(\alpha f + \beta g) \rightarrow \boxed{\quad\mathcal{L}\quad} \rightarrow \alpha\mathcal{L}(f) + \beta\mathcal{L}(g)$

Linear Shift Invariant Systems

$\mathcal{L}$ is called shift-invariant, if and only if a shift (translation) of the input causes the same shift of the output:

$$f(x,y) \rightarrow \boxed{\quad \mathcal{L} \quad} \rightarrow \mathcal{L}(f(x,y))$$

$$f(x-x_0, y-y_0) \rightarrow \boxed{\quad \mathcal{L} \quad} \rightarrow \mathcal{L}(f(x-x_0, y-y_0))$$

Alternative formulation: $\mathcal{L}$ commutes with a shift operator $\mathcal{S}$

$$\rightarrow \boxed{\quad \mathcal{L} \quad} \rightarrow \boxed{\quad \mathcal{S} \quad} \rightarrow \text{ same as } \rightarrow \boxed{\quad \mathcal{S} \quad} \rightarrow \boxed{\quad \mathcal{L} \quad} \rightarrow$$

$$\begin{tikzcd}
\mathbf{S(f)} \arrow[r, "\boldsymbol{L}"] & \mathbf{same!} \\
\mathbf{f} \arrow[u, "\mathbf{S}"] \arrow[r, "\boldsymbol{L}"'] & \boldsymbol{L}\mathbf{(f)} \arrow[u, "\mathbf{S}"']
\end{tikzcd}$$

Linear filtering in image processing

Using digital linear filters to modify pixel values based on some pixel neighborhoods. Linear means linear combination of neighbors.

- Linear methods simplest.
- Can combine linear methods in any order to achieve same result.
- May be easier to invert.

Useful to:

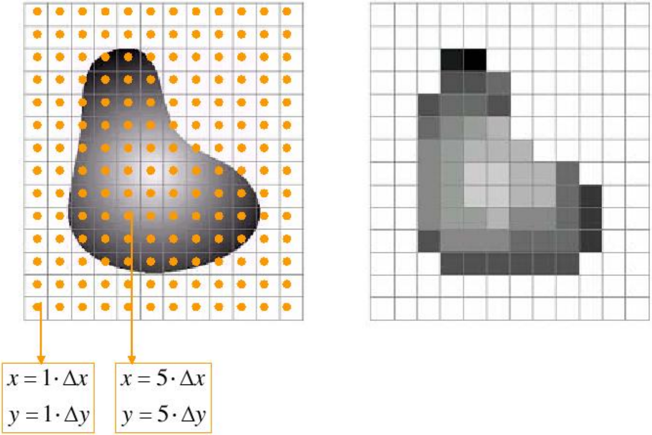
- Integrate information over constant regions.
- Scale changes of image intensities.
- Detect changes (edge detection).

Sampling

Our signals (images) are not in a continuous domain, but in a discrete.

- A continuous function $f(x, y)$ (an image) can be sampled using a discrete grid of sampling points.
- The image is sampled at points $(j\Delta x, k\Delta y)$, with $j = 1, ..., M$ and $k = 1, ..., N$, where is (M, N) is the size of the image in pixels.
- Here Δx and Δy are called the sampling interval.

Sampling



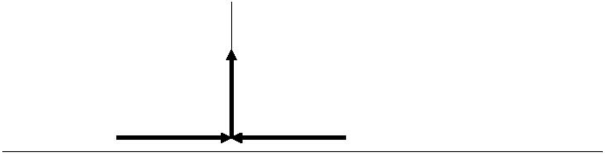
Sampling

Dirac (continuous domain) and Kronecker (discrete) delta functions.

Ideal impulse defined using ***Dirac distribution***

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\delta(x,y)\,dx\,dy = 1$$

and $\delta(x,y) = 0$ *for all* $x, y \neq 0$



Sampling

The 'sifting property' of the dirac distribution provides a value of the function f(x,y) at point a,b

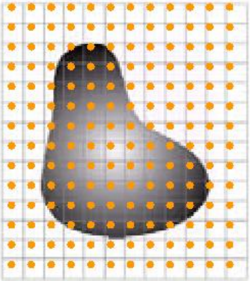
$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\delta(x-a,y-b)dxdy = f(a,b)$$

The sifting equation can be used to describe the sampling process of a continuous image function f(x,y)

Sampling

The ideal sampling s(x,y) in the regular grid can be represented using a collection of Dirac distributions δ

$$s(x,y) = \sum_{j=1}^{M} \sum_{k=1}^{N} \delta(x - j\Delta x, y - k\Delta y)$$

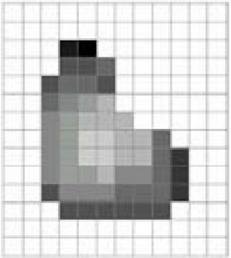


Sampling

The sampled image $f_s(x,y)$ is the product of the continuous image $f(x,y)$ and the sampling function $s(x,y)$

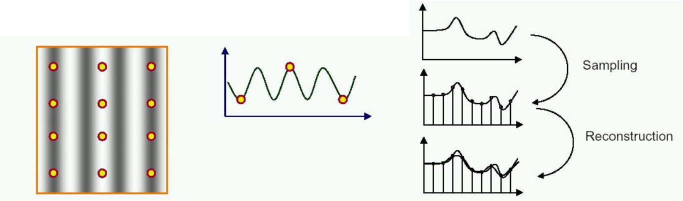
$$f_s(x,y) = f(x,y)s(x,y)$$

$$= f(x,y)\sum_{j=1}^{M}\sum_{k=1}^{N}\delta(x - j\Delta x, y - k\Delta y)$$



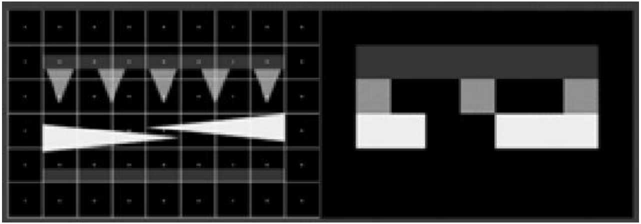
Sampling

- Sources of error during sampling:
  - Intensity quantization (not enough intensity resolution).
  - Spatial aliasing (not enough spatial resolution).
  - Temporal aliasing (not enough temporal resolution).
- Sampling Theorem answers (more later):
  - How many samples are required to describe the given signal without loss of information? (more later)
  - What signal can be reconstructed given the current sampling rate?

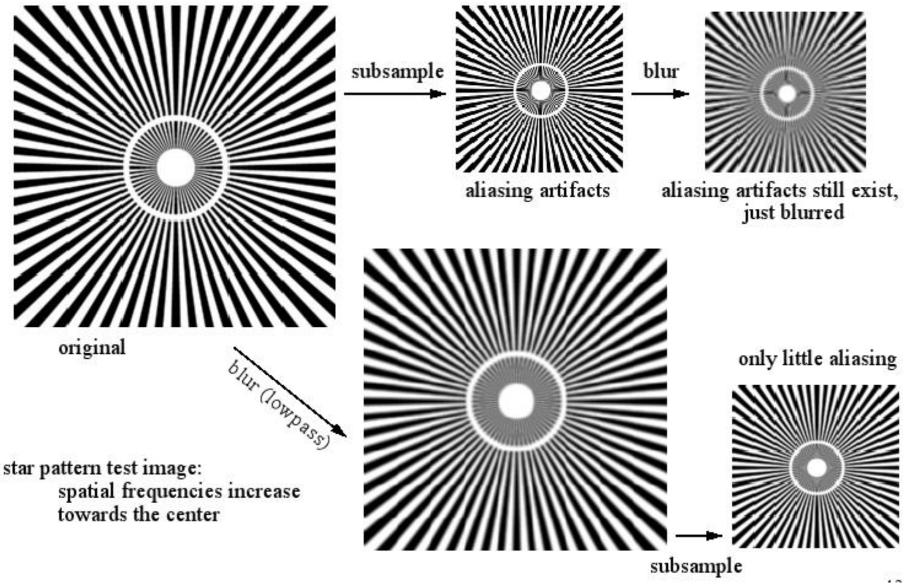


Aliasing and anti-aliasing

- Artifacts produced by under-sampling or poor reconstruction.
  Fine structures disappear and distort coarser structure.
- Spatial and temporal aliasing.
- Anti-aliasing: sample at higher rate or prefiltering.
  Tools: Fourier transform, convolution and sampling theory.

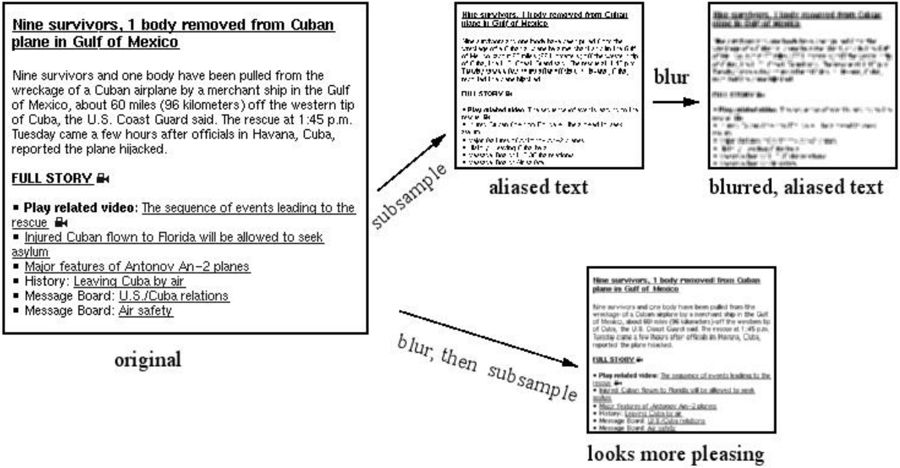


Example: Aliasing



Example: Aliasing

Low pass filtering (bluring) important!



Linear image filtering

- Estimate an output image by modifying pixels in the input image using a function of a local pixel neighborhood.
- The neighborhood and the corresponding linear weights per pixel is called a convolution kernel.

$$\begin{array}{|c|c|c|}\hline 9 & 5 & 3 \\ \hline 4 & 5 & 1 \\ \hline 1 & 1 & 7 \\ \hline \end{array} \quad \overset{\text{some function}}{\Longrightarrow} \quad \begin{array}{|c|c|c|}\hline & & \\ \hline & 9 & \\ \hline & & \\ \hline \end{array}$$

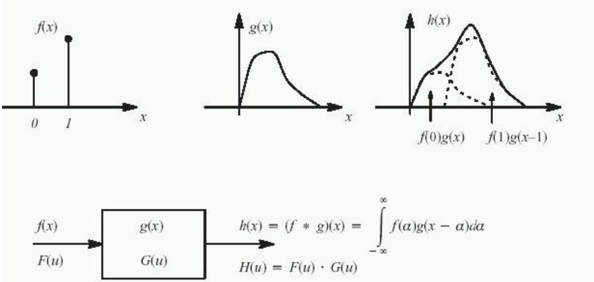
$$\begin{array}{|c|c|c|}\hline 9 & 5 & 3 \\ \hline 4 & 5 & 1 \\ \hline 1 & 1 & 7 \\ \hline \end{array} * \begin{array}{|c|c|c|}\hline 0 & -1 & 0 \\ \hline -1 & 4 & -1 \\ \hline 0 & -1 & 0 \\ \hline \end{array} = \begin{array}{|c|c|c|}\hline & & \\ \hline & 9 & \\ \hline & & \\ \hline \end{array}$$

Convolution

- Convolution is a tool to build linear shift invariant (LSI) filters.
- Mathematically, a convolution is defined as the integral over space of one function at α, times another function at $x-\alpha$.

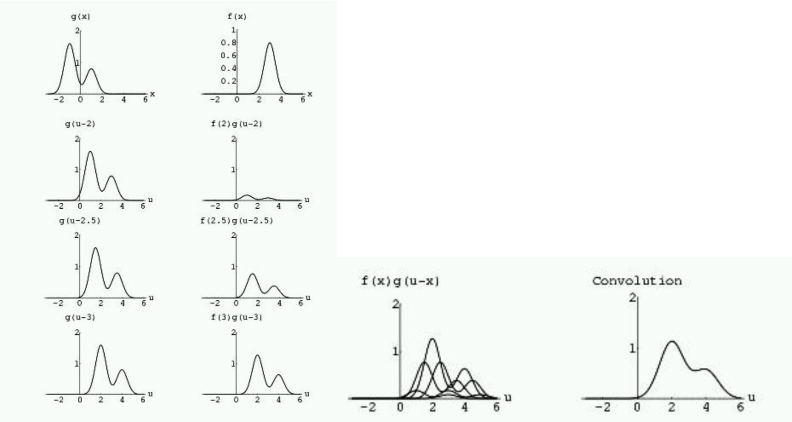
$$f(x) * g(x) = \int_{\alpha \in \mathbb{R}^n} f(\alpha) g(x-\alpha) d\alpha = g(x) * f(x) = \int_{\alpha \in \mathbb{R}^n} g(\alpha) f(x-\alpha) d\alpha$$

- Convolution operation is commutative!



Convolutions as weighted sums

Way of considering convolution: weighted sum of shifted copies of one function, with weights given by the function value of the second function at the shift vector.



Theorem 1

Every shift invariant linear operator can be written as a convolution

$$\mathcal{L}(\mathrm{f}) = \mathrm{g} * \mathrm{f}$$

- Continuous case

$$\mathcal{L}(\mathrm{f}(x)) = \int_{\alpha \in \mathbb{R}^n} \mathrm{g}(\alpha)\mathrm{f}(x-\alpha)d\alpha$$

- Discrete case

$$\mathcal{L}(\mathrm{f}(x)) = \sum_{\alpha \in \mathbb{R}^n} \mathrm{g}(\alpha)\mathrm{f}(x-\alpha)$$

Convolution (discrete case)

- The convolution of an image $f(x,y)$ with a kernel h(x,y) is

$$g(x,y) = h(x,y) * f(x,y) = \sum_{m=-M}^{M} \sum_{n=-N}^{N} h(m,n)\, f(x-m, y-n)$$

- Convolution kernel $h(x,y)$ represented as a matrix and is also called:
  - impulse response,
  - point spread function,
  - filter kernel,
  - filter mask,
  - template...

Convolution (filtering)

- Frame mask over image - multiply mask values by image values and sum up the results - a sliding dot product.

- For mathematical correctness: From the definition, the kernel first has to be flipped x-wise and y-wise. People are sloppy though.

Convolution: 1D example

If

$$F_1 = [1 \quad 2 \quad 3 \quad 4 \quad 5]$$
$$F_2 = [1 \quad 2 \quad 1 \quad 2 \quad 1]$$
$$G_1 = [-1 \quad 2 \quad -1]$$
$$G_2 = [1 \quad 2 \quad 3]$$

then

$$F_1 * G_1 = [-1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 6 \quad -5]$$
$$F_2 * G_1 = [-1 \quad 0 \quad 2 \quad -2 \quad 2 \quad 0 \quad -1]$$
$$F_1 * G_2 = [1 \quad 4 \quad 10 \quad 16 \quad 22 \quad 22 \quad 15]$$
$$F_2 * G_2 = [1 \quad 4 \quad 8 \quad 10 \quad 8 \quad 8 \quad 3]$$

Note1: outside the windows, values are assumed to be zero.
Note2: normally you assume $x = 0$ at center of filter kernel.

Convolution: 1D example

$F_1 = [1 \quad 2 \quad 3 \quad 4 \quad 5]$

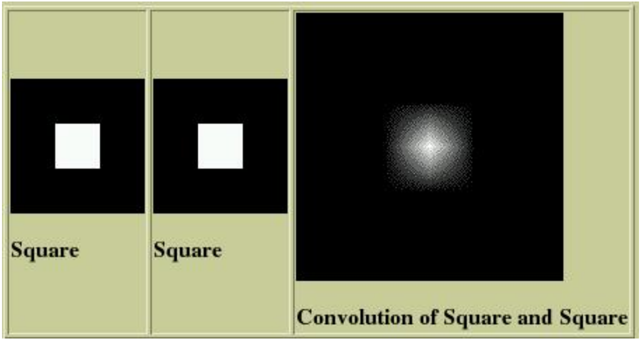
$G_2 = [1 \quad 2 \quad 3]$

$$
\begin{array}{cccccccc}
 & & & 1 & 2 & 3 & 4 & 5 \\
* & & & & & 1 & 2 & 3 \\
\hline
 & & & 3 & 6 & 9 & 12 & 15 \\
 & & 2 & 4 & 6 & 8 & 10 & \\
+ & 1 & 2 & 3 & 4 & 5 & & \\
\hline
 & 1 & 4 & 10 & 16 & 22 & 22 & 15
\end{array}
$$

An easier way of doing it! Almost like regular multiplication.

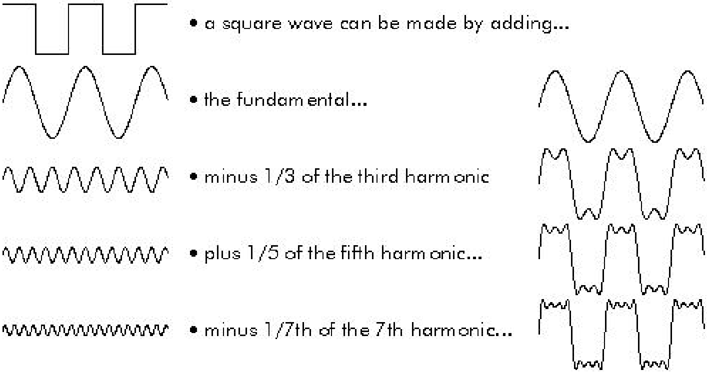
Convolution: 2D example

- Convolution of two images: since the squares have the same image and size, their convolution creates a gradient with the brightest spot in the center.



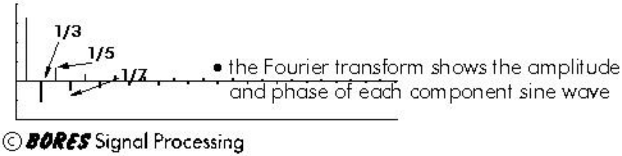
Signal decomposition

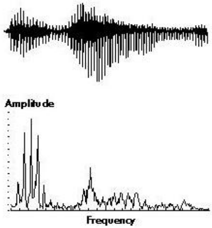
- In 1807 Jean Baptiste Fourier showed that any periodic signal could be represented by a series/sum of sine waves with appropriate amplitude, frequency and phase.



The Fourier transform

- The Fourier transform is an equation to calculate the frequency, amplitude and phase of each sine wave needed to make up any given signal.
- Using the Fourier transform we can study frequency components of signals.
- The Fourier transform *converts* between the spatial and frequency domain representations.



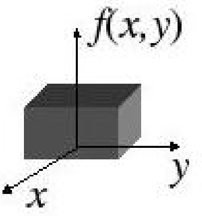


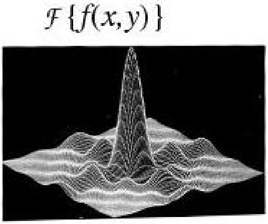
The Fourier transform

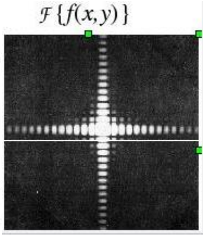
- An important image processing tool which is used to decompose an image into its sine and cosine components.
- The output of the transformation represents the image in the Fourier or frequency space, while the input image is the real space equivalent.
- In the Fourier space image, each point represents a particular frequency contained in the real domain image.
- The Fourier Transform is used in a wide range of applications, such as image analysis, image filtering, image reconstruction and image compression

Images and spatial frequency

- The spatial frequency of an image refers to the rate at which the pixel intensities change.
- The easiest way to determine the frequency composition of signals is to inspect that signal in the frequency domain.
- The frequency domain shows the magnitude of different frequency components.







Change of basis functions

- An image can be viewed as a spatial array of gray level values, but can also thought of as a spatially varying function.
- Decompose the image into a set of orthogonal basis functions.
- When basis functions are combined (linearly) the original function will be reconstructed.
- Spatial domain: basis consists of shifted Dirac functions.
  Fourier domain: basis consists of complex exponential functions.
- The Fourier transform is “just” a change of basis functions.

Change of basis functions

- Assume you have a vector

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = 1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + 4 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

- This may be expressed with another basis (e.g. Haar wavelet).

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = 2.5 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - 1 \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} - 0.5 \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} - 0.5 \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}$$

- The only condition is that the basis vector are orthogonal.

$$\mathcal{F}(\mathrm{f}(x)) = \int_{x\in\mathbb{R}^n} \mathrm{f}(x) e^{-i\omega^T x} dx = \hat{\mathrm{f}}(\omega)$$

$$\mathcal{F}^{-1}(\hat{\mathrm{f}}(\omega)) = \frac{1}{(2\pi)^n} \int_{\omega\in\mathbb{R}^n} \hat{\mathrm{f}}(\omega) e^{i\omega^T x} d\omega$$

$$e^{i\omega^T x} = \cos \omega^T x + i \sin \omega^T x$$

Terminology:

Frequency spectrum : $\hat{\mathrm{f}}(\omega) = R(\omega) + iI(\omega) = |\ \hat{\mathrm{f}}(\omega)\ |\ e^{i\phi(\omega)}$

Fourier spectrum: $|\ \hat{\mathrm{f}}(\omega)\ | = \sqrt{R^2(\omega) + I^2(\omega)}$

Power spectrum: $|\ \hat{\mathrm{f}}(\omega)\ |^2$

Phase angle: $\phi(\omega) = \arg \hat{\mathrm{f}}(\omega) = \tan^{-1} \frac{I(\omega)}{R(\omega)}$

Terminology

- Angular frequency $\omega = (\omega_1 \; \omega_2)^T$

$$\omega_1 = \text{angular frequency in } x \text{ direction}$$
$$\omega_2 = \text{angular frequency in } y \text{ direction}$$

- Frequency

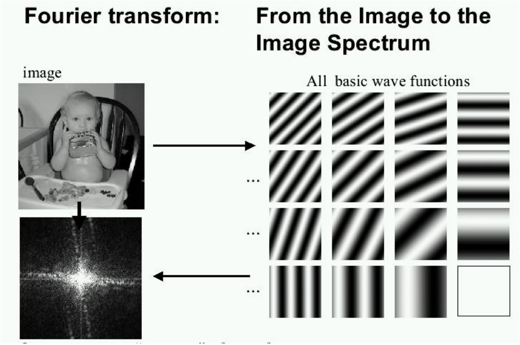
$$f = \frac{\omega}{2\pi}$$

- Wavelength

$$\lambda = \frac{2\pi}{\|\omega\|} = \frac{2\pi}{\sqrt{\omega_1^2 + \omega_2^2}}$$

$$e_\omega(x) = e^{i\omega^T x} = e^{i(\omega_1 x_1 + \omega_2 x_2)} = \cos\omega^T x + i\sin\omega^T x \quad \text{(Euler's formula)}$$

$$Re(e_\omega(x)) = \cos(\omega^T x) \ \text{ and } \ Im(e_\omega(x)) = \sin(\omega^T x), \ e_\omega : \mathbb{R}^2 \rightarrow \mathbb{C}$$



Magnitude and Phase

- The Fourier coefficients $F(u,v)$ are complex numbers, but it is not obvious what the real and imaginary parts represent.
- Another way to represent the data is with phase and magnitude.

  Magnitude:

  $$|F(u,v)| = \sqrt{Re^2(u,v) + Im^2(u,v)}$$

  Phase:

  $$\theta(u,v) = \tan^{-1}\frac{Im(u,v)}{Re(u,v)}$$

Note: Change in notation from $\hat{f}(\omega_1, \omega_2)$ to $F(u,v)$, which is more common, if we know we are working with images in two dimensions.

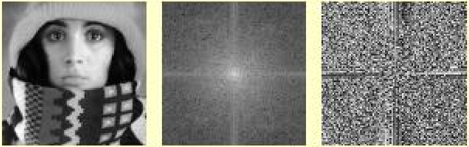


Figure 7 Illustration of an original image, magnitude and phase of the Fourier transform respectively

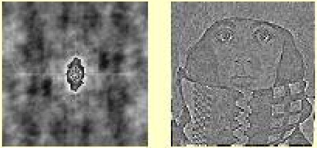
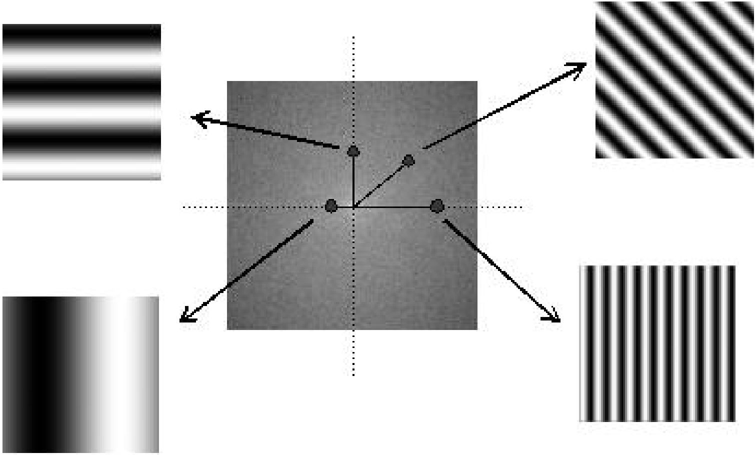
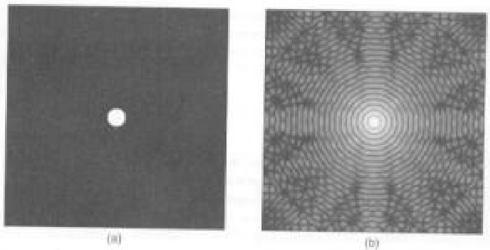


Figure 8 reconstructed images using only magnitude and phase

Example



- An image of a spot (left) and the Fourier Transform (right).



- In the frequency domain, u represents the spatial frequency along the original image's x axis and v represents the spatial frequency along the y axis. The origin is in the center of the image.

Summary of good questions

- What properties does a linear system have?
- What does shift-invarience mean in terms of image filtering?
- How can a Dirac function be used to model sampling?
- How do you define a convolution?
- Why are convolutions important in linear filtering?
- How do you define a 2D Fourier transform?
- If you apply a Fourier transform to an image, what do you get?
- What information does the phase contain? What about the magnitude?
- Will the second lab be as boring as the first?

Readings

- Gonzalez and Woods: Chapter 4
- Szeliski: Chapters 3.2 and 3.4
- Introduction to Lab 2