Computer Vision

Images in the Frequency Domain

Prof. Flávio Cardeal – DECOM / CEFET-MG

cardeal@decom.cefetmg.br

Abstract

- This lecture introduces some *mathematical concepts* for describing an image in the *frequency domain*.

Fourier Transform

- The *Fourier transform* defines a traditional way for processing signals.

- The *2D Fourier transform* maps an image from its spatial domain into the frequency domain, providing a totally different representation.

Fourier Transform - Overview

- The analysis of data in the frequency domain provides insights into the given image I.
- Changes in the frequency domain of an image I are based on *Fourier filter operations*.
- Using the inverse DFT, we then map the modified Fourier transform back into the modified image.

Fourier Filtering

Fourier Filtering

- The *Fourier filtering* process allows us, for example, to do contrast enhancement, noise removal, or smoothing of images.

- In the context of the Fourier transform we assume that the image coordinates run from 0 to $N_{cols} - 1$ for x and from 0 to $N_{rows} - 1$ for y.

2D DFT

- Formally, the 2D DFT is defined as follows:

$$\mathbf{I}(u, v) = \frac{1}{N_{cols} \cdot N_{rows}} \sum_{x=0}^{N_{cols}-1} \sum_{y=0}^{N_{rows}-1} I(x, y) \cdot \exp\left[-i2\pi\left(\frac{xu}{N_{cols}} + \frac{yv}{N_{rows}}\right)\right]$$

for $u = 0,1,\ldots,N_{cols} - 1$ and $v = 0,1,\ldots,N_{rows} - 1$. The letter $i = \sqrt{-1}$ is the imaginary unit of complex numbers.

Eulerian Formula

- For any real α, the *Eulerian* formula:

$$\exp(i\alpha) = e^{i\alpha} = \cos\alpha + i \cdot \sin\alpha$$

demonstrates that the Fourier transform is actually a *weighted sum of sine and cosine functions*, but in the complex plane.

Inverse 2D DFT

- The inverse 2D DFT transforms a Fourier transform $\mathbf{I}$ back into the spatial domain:

$$I(x, y) = \sum_{u=0}^{N_{cols}-1} \sum_{v=0}^{N_{rows}-1} \mathbf{I}(u, v) \exp\left[i2\pi \left(\frac{xu}{N_{cols}} + \frac{yv}{N_{rows}} \right) \right]$$

- Note: definitions of DFT and inverse DFT may have small variations.

Basis Functions

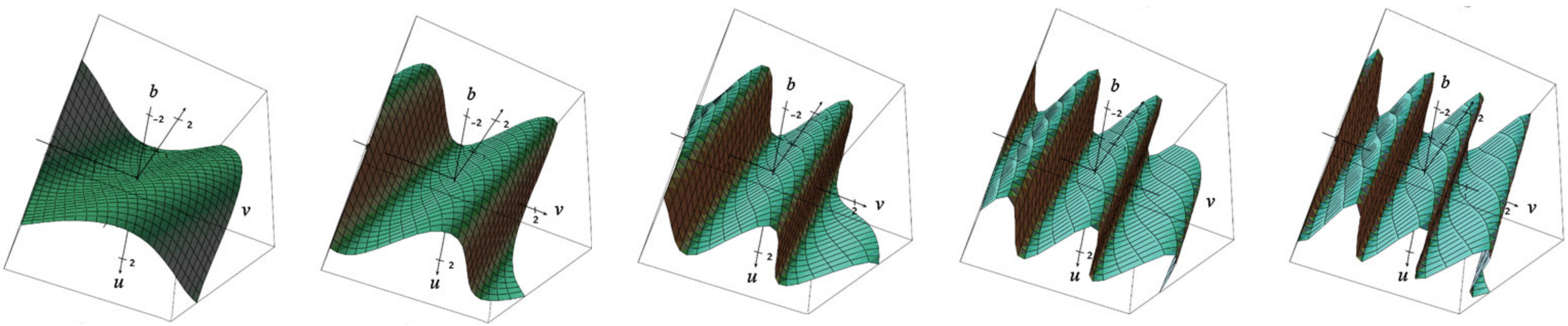
- The equation below shows that we represent the image I now as a weighted sum of basis functions $\exp(i\alpha) = \cos\alpha + i \cdot \sin\alpha$:

$$I(x, y) = \sum_{u=0}^{N_{cols}-1} \sum_{v=0}^{N_{rows}-1} \mathbf{I}(u, v) \exp\left[i2\pi \left(\frac{xu}{N_{cols}} + \frac{yv}{N_{rows}} \right) \right]$$

Basis Functions

- The figure below illustrates five basis functions $\sin(u+nv)$ for the imaginary part b of complex values $a+ib$ in the uv frequency domain.

Source: R. Klette



n=1 *n*=2 *n*=3 *n*=4 *n*=5

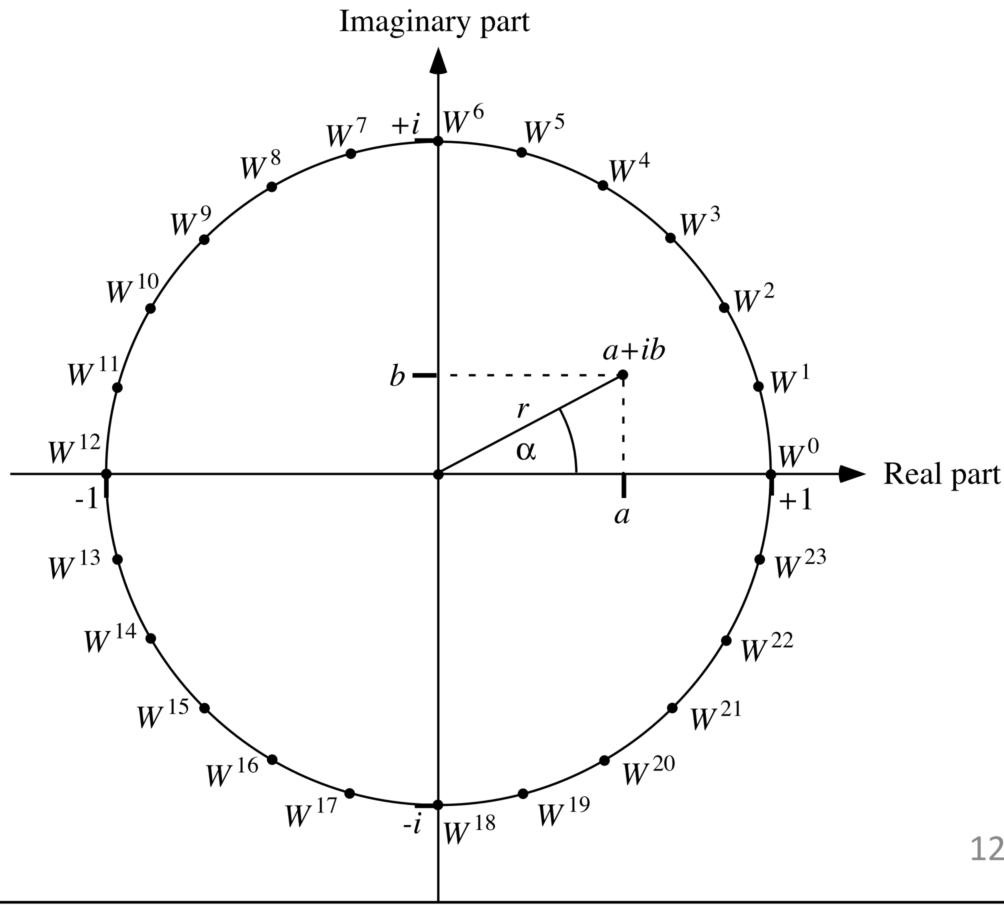
Basis Functions

- The values $\mathbf{I}(u,v)$ of the Fourier transform of I are called *Fourier coefficients* and are the weights to the basis functions $\exp(i\alpha)$.

- For example, *point noise* or *edges* require large coefficients for high frequency components, to be properly represented in this weighted sum.

The Complex Plane

- It is common practice to visualize complex numbers $a+ib$ as points (a,b) or vectors $[a,b]^{\mathrm{T}}$ in the plane, called *complex plane*.



Calculus of Complex Numbers

- Let $z_1 = a_1 + ib_1$ and $z_2 = a_2 + ib_2$ be two complex numbers, with $i = \sqrt{-1}$. So, we have:

Sum: $z_1 + z_2 = (a_1 + a_2) + i \cdot (b_1 + b_2)$

Product: $z_1 \cdot z_2 = (a_1 a_2 - b_1 b_2) + i \cdot (a_1 b_2 + a_2 b_1)$

Norm: $\|z\|_2 = \sqrt{a^2 + b^2}$

Calculus of Complex Numbers

- The conjugate z^* of a complex number $z = a + ib$ is $z = a - ib$, where $(z^*)^* = z$ and $(z_1 \cdot z_2)^* = z_1^* \cdot z_2^*$.

Division: $$\frac{z_1}{z_2} = \frac{z_1}{z_2} \cdot \frac{z_2^*}{z_2^*}$$

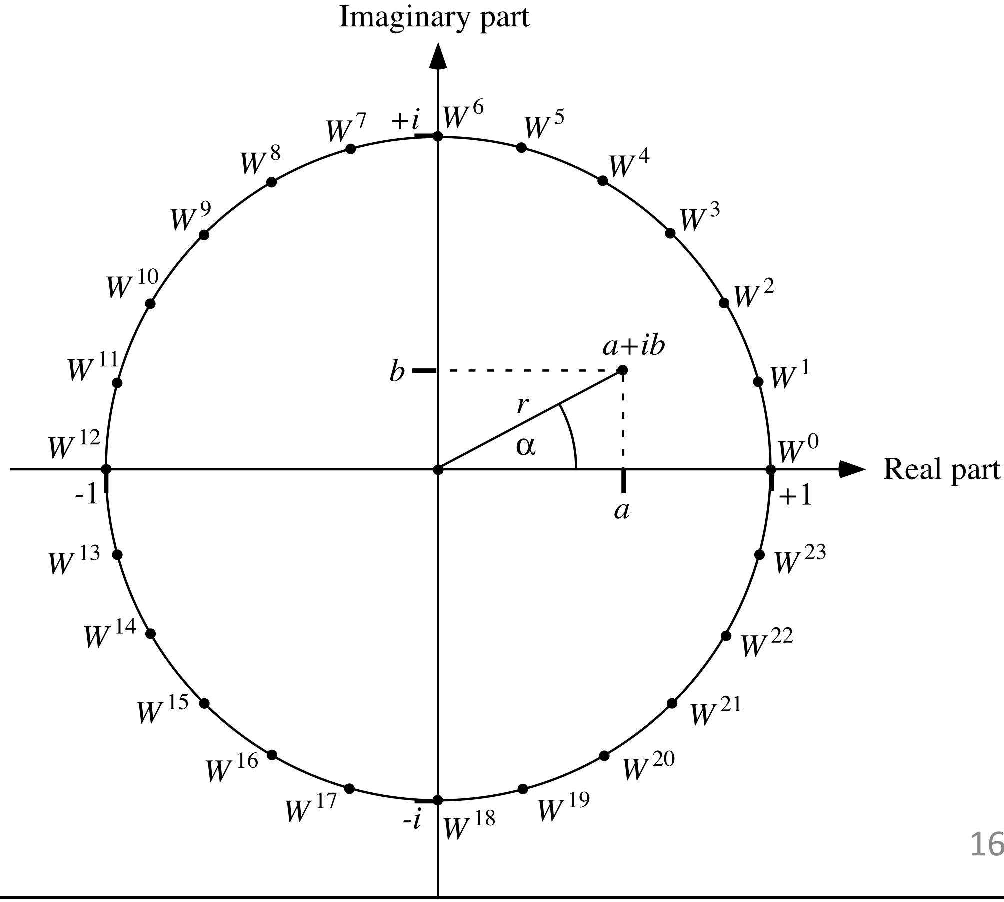
Polar Coordinates

- A complex number z can also be written in the form $z = r \cdot e^{i\alpha}$, with $r = \|z\|_2$ and α is the angle made with the real axis.
- This maps the numbers into polar coordinates.
- A rotation of a vector $[c,d]^{\mathrm{T}}$ about an angle α is the vector $[a,b]^{\mathrm{T}}$, with $a + i \cdot b = e^{i\alpha} \cdot (c + i \cdot d)$.

Roots of Unity

- The complex number $W_M = \exp\left[i2\pi / M\right]$ defines the *M*th root of unity.

- This figure shows all the powers of the *24th* root of unity.



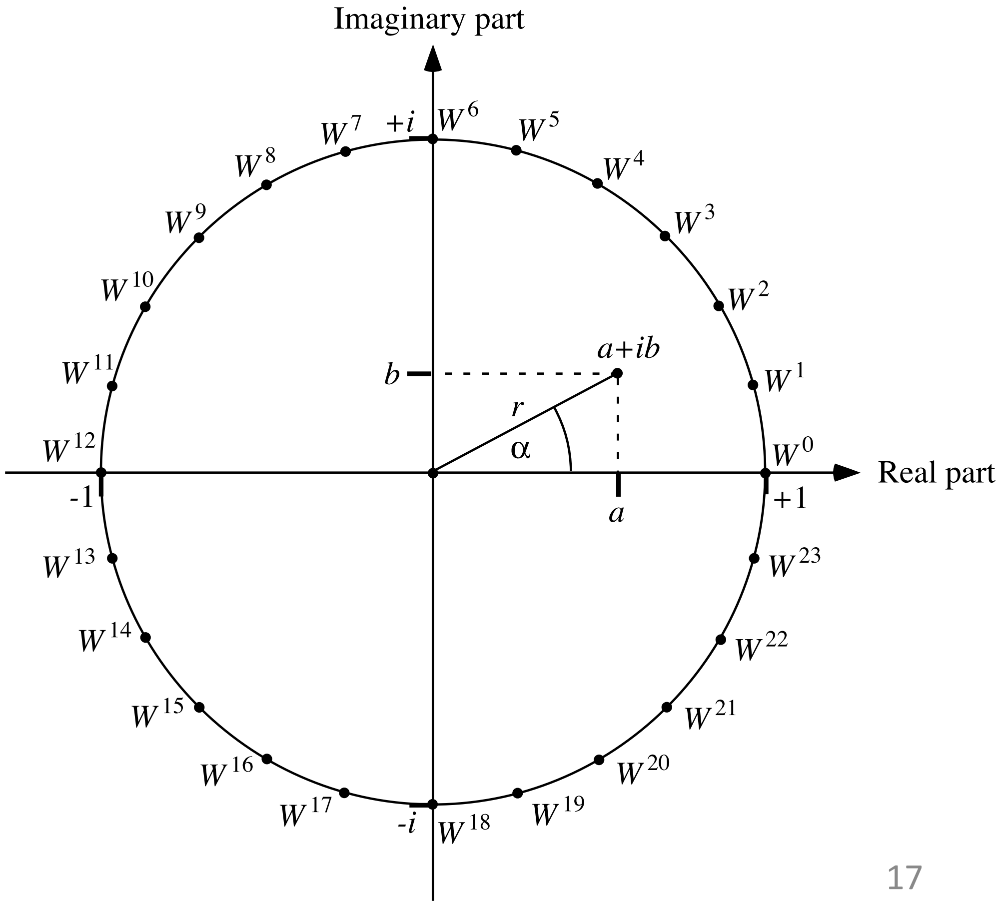
Roots of Unity

- For example:

$$W_{24}^{0} = e^{0} = 1$$

$$W_{24}^{1} = e^{i2\pi/24} = \cos\frac{\pi}{12} + i\sin\frac{\pi}{12}$$

$$W_{24}^{6} = e^{i12\pi/24} = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2} = i$$



Roots of Unity

- The equation for the DFT can be simplified by using the notion of roots of unity:

$$\mathbf{I}(u,v) = \frac{1}{N_{cols} \cdot N_{rows}} \sum_{x=0}^{N_{cols}-1} \sum_{y=0}^{N_{rows}-1} I(x,y) \cdot \exp\left[-i2\pi\left(\frac{xu}{N_{cols}} + \frac{yv}{N_{rows}}\right)\right]$$

$$\Downarrow$$

$$\mathbf{I}(u,v) = \frac{1}{N_{cols} \cdot N_{rows}} \sum_{x=0}^{N_{cols}-1} \sum_{y=0}^{N_{rows}-1} I(x,y) \cdot W_{N_{cols}}^{-xu} \cdot W_{N_{rows}}^{-yv}$$

Frequency Domain

- The complex values of the 2D DFT are defined in the uv frequency domain.
- The values for *low frequencies* u or v (~ to 0) represent long wavelengths of sine or cosine components.
- Values for *large frequencies* u or v (away from zero) represent short wavelengths.

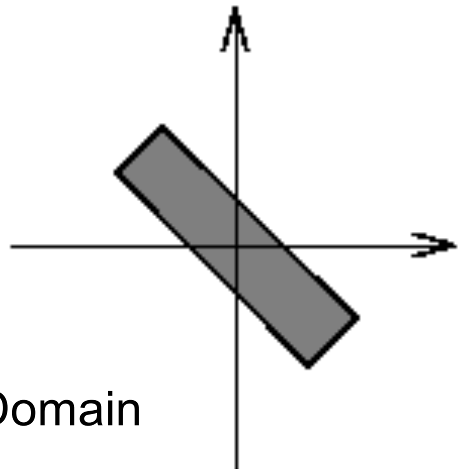
Interpretation of Matrix I

- *Low frequencies* represent *long wavelengths* and thus homogeneous additive contributions to the input image I.

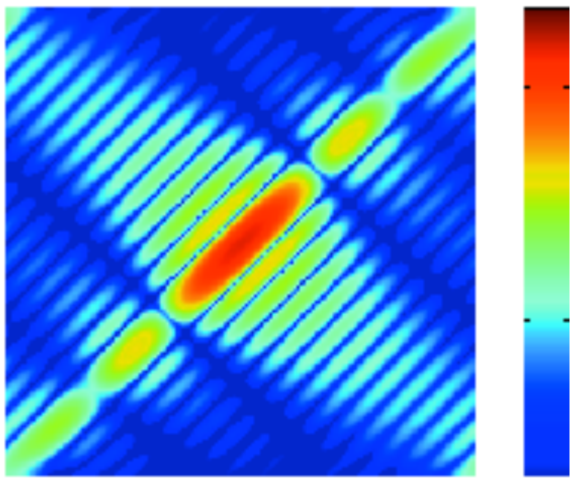
- *High frequencies* represent *short wavelengths* (and thus local discontinuities in I such as edges or intensity outliers).

Interpretation of Matrix I

- *Directional patterns* in I, for example lines into direction β, create value distributions in $\mathbf{I}$ in the orthogonal direction (i.e., in direction $\beta+\pi/2$).



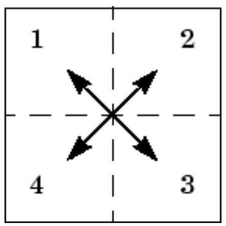
Spatial Domain



Frequency Domain

Interpretation of Matrix I

- In images we have the origin at the upper left corner (left-hand coordinate system).
- If you want the *origin in the center* location of the Fourier transform, then this can be achieved by *swapping the four quadrants* of the matrix.



For example:

I33	I43	I53	I03	I13	I23
I34	I44	I54	I04	I14	I24
I35	I45	I55	I05	I15	I25
I30	I40	I50	I00	I10	I20
I31	I41	I51	I01	I11	I21
I32	I42	I52	I02	I12	I22

Properties of the DFT

- The element $\mathbf{I}(0,0)$ is the mean of I:

$$\mathbf{I}(0,0) = \frac{1}{N_{cols} \cdot N_{rows}} \sum_{x=0}^{N_{cols}-1} \sum_{y=0}^{N_{rows}-1} I(x, y)$$

- It is usually called *DC component* of I .
- Because I has only real values, the imaginary part of $\mathbf{I}(0,0)$ is always equal to zero.

Properties of the DFT

- Parseval's theorem:

$$\frac{1}{|\Omega|}\sum_{\Omega}\left|I(x,y)\right|^2=\sum_{\Omega}\left|\mathbf{I}(u,v)\right|^2$$

- It states identities in total sums of absolute values for the input image I and the Fourier transform $\mathbf{I}$.

Properties of the DFT

- Symmetry Property:

$$\mathbf{I}(-u,-v) = \mathbf{I}(u,v)^*$$

Observation: consider the zero-frequency component shifted to the center.

or equivalently:

$$\mathbf{I}(N_{cols}-1-u, N_{rows}-1-v) = \mathbf{I}(u,v)^*$$

Observation: consider the zero-frequency component at the upper left corner.

Properties of the DFT

- Symmetry Property - Example:

```
I   = [0.0318     0.0971     0.3171
       0.2769     0.8235     0.9502
       0.0462     0.6948     0.0344]

F   = [3.2721 + 0.0000i  -1.1037 - 0.2716i  -1.1037 + 0.2716i
      -0.9670 - 1.1043i   0.8793 + 0.6740i  -0.3015 - 0.1692i
      -0.9670 + 1.1043i  -0.3015 + 0.1692i   0.8793 - 0.6740i]

Fs  = [0.8793 - 0.6740i  -0.9670 + 1.1043i  -0.3015 + 0.1692i
      -1.1037 + 0.2716i   3.2721 + 0.0000i  -1.1037 - 0.2716i
      -0.3015 - 0.1692i  -0.9670 - 1.1043i   0.8793 + 0.6740i]
```

Spectrum and Phase

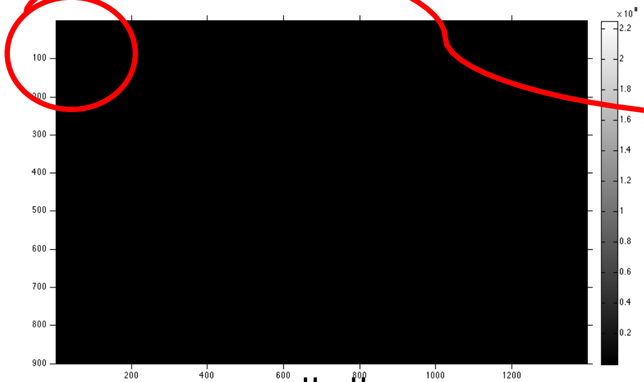
- The *magnitude* or *amplitude* $\|z\|_2 = r = \sqrt{a^2 + b^2}$ and the *phase* $\alpha = \arctan(b/a)$ define complex numbers $z = a + i \cdot b$ in polar coordinates (r, α).

- The magnitude provides a convenient way of representing the complex-valued matrix $\mathbf{I}$ in the form of the *spectrum* $\|\mathbf{I}\|$ or $\|\mathbf{I}(u, v)\|_2$.

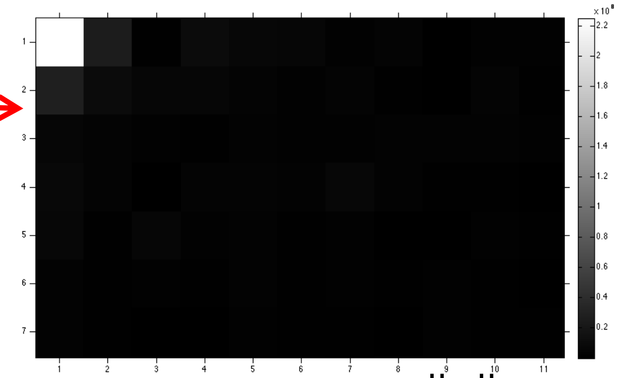
Spectrum

- Usually, when visualizing the spectrum $\|\mathbf{I}\|$ as a grey-level image, it is basically black, just with a bright dot at the origin (mean).
- This is because all values in $\|\mathbf{I}\|$ are typically rather small. For better visibility, the spectrum is log-transformed into $\log_{10}(1+\|\mathbf{I}(u,v)\|_2)$.

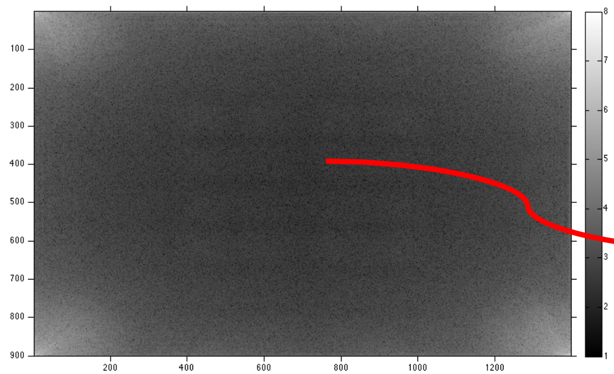
Spectrum



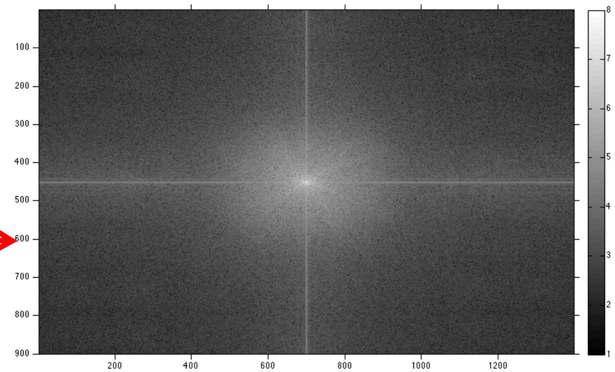
$\|\mathbf{I}\|$



Zoom in of $\|\mathbf{I}\|$



$\log_{10}(1+\|\mathbf{I}(u,v)\|_2)$



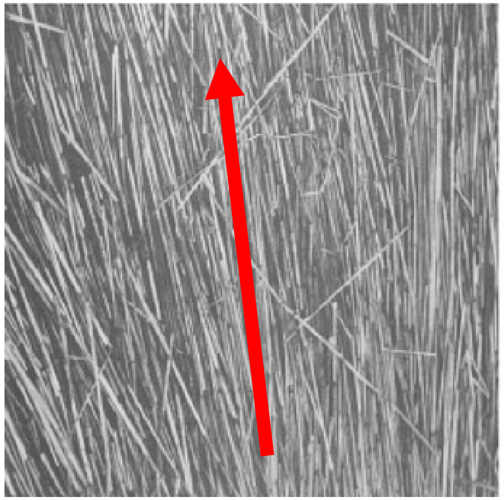
Origin at the center

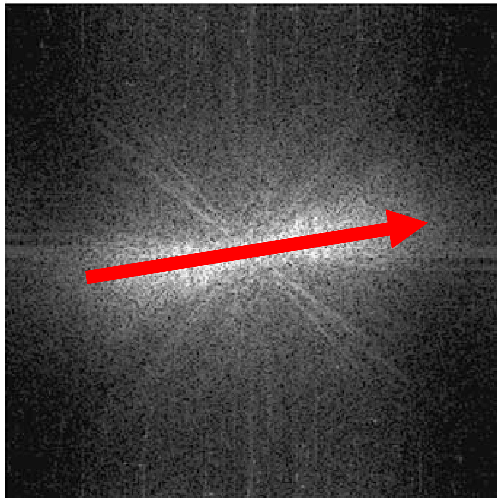
Phase

- Visualizations of the phase components of $\|\mathbf{I}\|$ are not so common.
- In fact, the visual representation of the phase components is not relevant to understand the information present in an image.

Directional Patterns

- *Directional patterns* in the spatial domain are rotated by $\pi/2$ in the frequency domain:





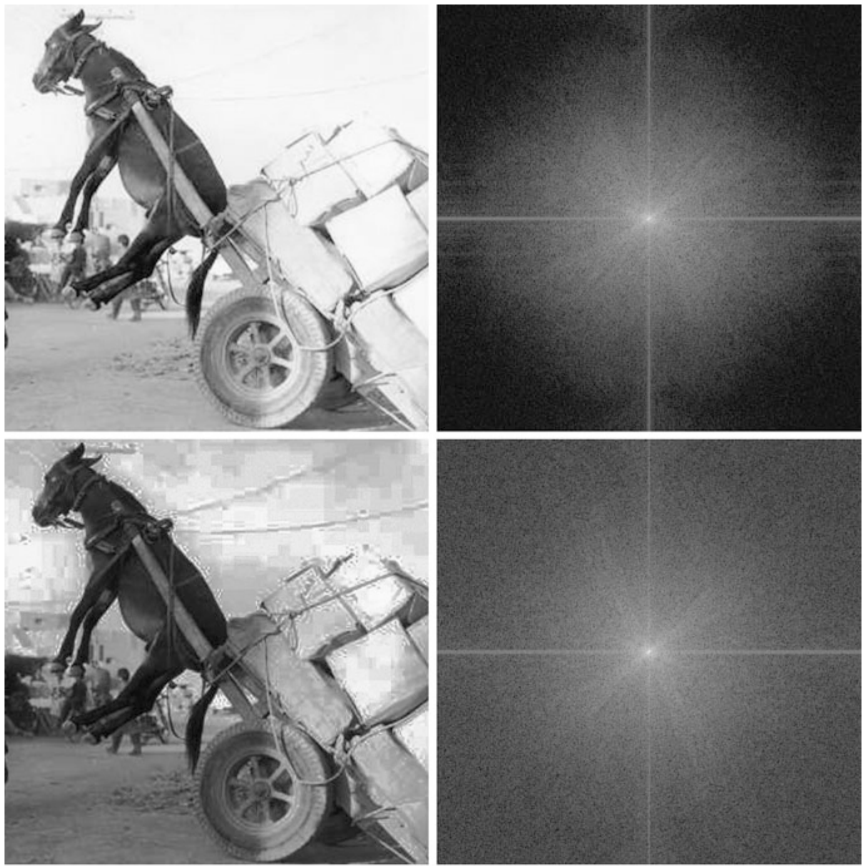
Source: R. Klette

Uniform Transforms

- *Uniform transforms* of the input image do not change the basic value distribution pattern in the spectrum.

- Examples of uniform transforms:
  - Adding a constant to each pixel value;
  - Histogram equalization.

Uniform Transforms



Original image

After histogram
equalization

Source: R. Klette

Fourier Pairs

- An input image and its Fourier transform define a *Fourier pair*.
- Here are some properties of Fourier transform:

$$I * G(x, y) \Leftrightarrow \mathbf{I} \circ \mathbf{G}(u, v)$$

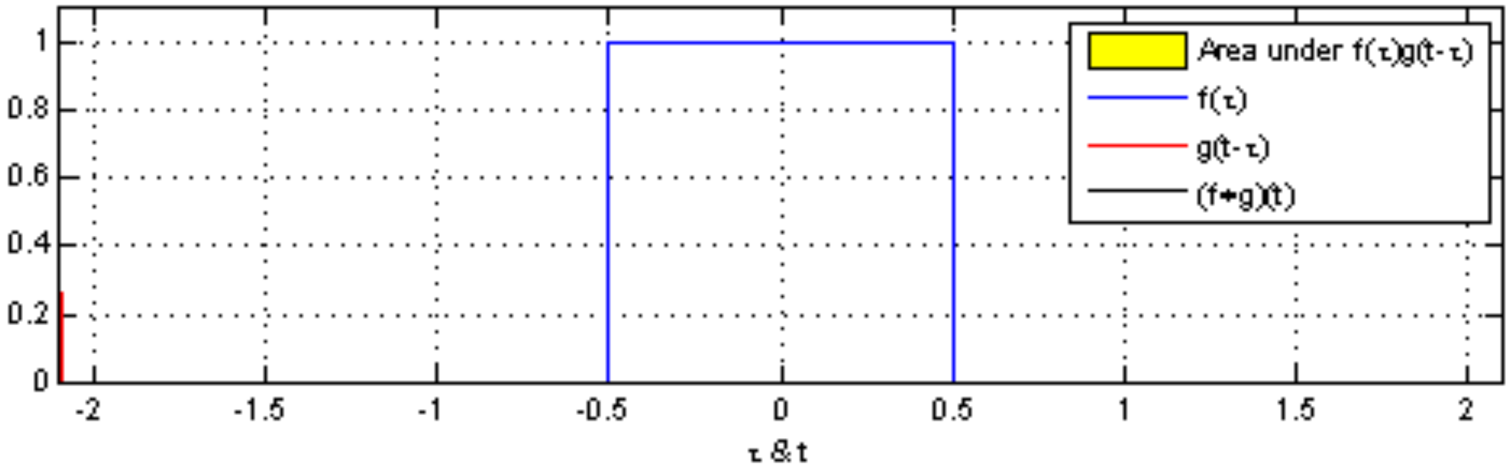
The Fourier transform of a convolution of I with a filter kernel G equals a point-by-point product of values in the Fourier transforms of I and G.

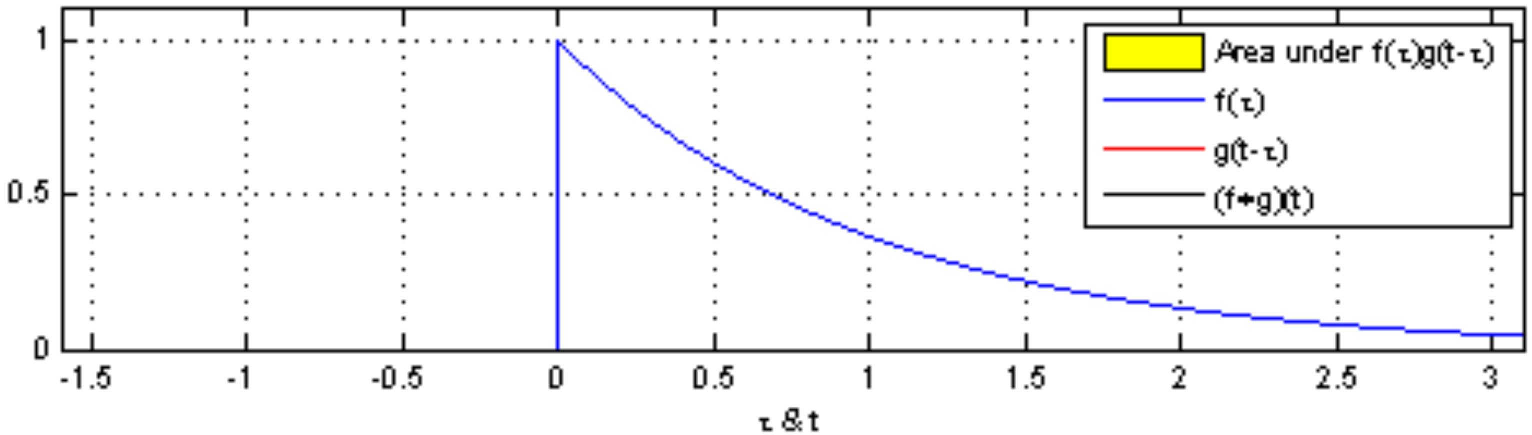
Convolution

- Convolution is an operation on two functions f and g, producing a third one that is viewed as a modified version of one of the original functions.

$$(f * g)(t) = h(t) = \int_{-\infty}^{+\infty} f(\tau) g(t-\tau) d\tau$$

Convolution – Visual Explanation





1D Discrete Convolution

- Analogously, the discrete convolution between two one-dimensional signals f and g is:

$$(f * g)(k) = h(k) = \sum_{n=-\infty}^{+\infty} f(n) g(k-n)$$

- And what about two-dimensional signals?

2D Discrete Convolution

- Analogously, the discrete convolution between two two-dimensional signals I and G is:

$$(I * G)(x,y) = H(x,y) = \sum_{n=0}^{N_{cols}-1} \sum_{m=0}^{N_{rows}-1} I(n,m)G(x-n, y-m)$$

Origin in the Center Location

- If the input image is multiplied by a chessboard pattern of +1 and −1, the zero-frequency component is shifted to the central position:

$$I(x, y) \cdot (-1)^{x+y} \Leftrightarrow \mathbf{I}\left(u + \frac{N_{cols}}{2}, v + \frac{N_{rows}}{2}\right)$$

Linear Transformation

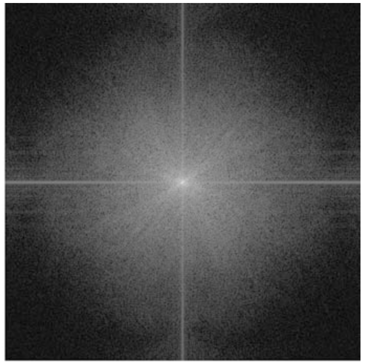
- The Fourier transform is a linear transformation:

$$a \cdot I(x, y) + b \cdot J(x, y) \Leftrightarrow a \cdot \mathbf{I}(u, v) + b \cdot \mathbf{J}(u, v)$$

Linear Transformation

- The Fourier transform is a linear transformation:

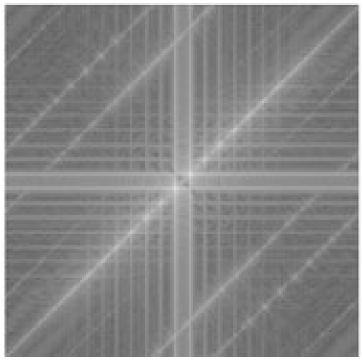
$I(x,y)$



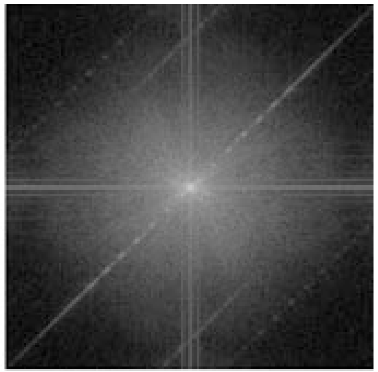
$\mathbf{I}(u,v)$

$I(x,y)+J(x,y)$

$J(x,y)$



$\mathbf{J}(u,v)$



$\mathbf{I}(u,v)+\mathbf{J}(u,v)$

Phase-Congruency

- A dimensionless measure that quantifies the correspondence of *phase values* in an image window defined by a reference point p.
- It is invariant to illumination and contrast changes.
- This mechanism is in accordance with the human visual system that demonstrates good invariance to lighting conditions.

Phase-Congruency

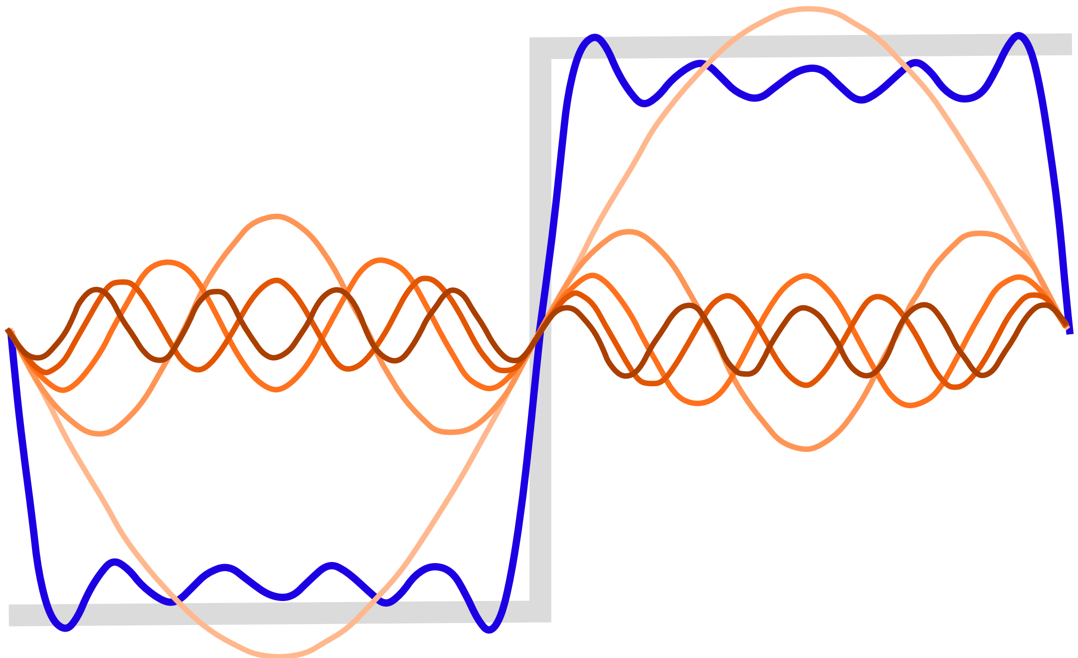
- Consider again the inverse 2D DFT formula:

$$I(x, y) = \sum_{u=0}^{N_{cols}-1} \sum_{v=0}^{N_{rows}-1} \mathbf{I}(u, v) \exp\left[i2\pi \left(\frac{xu}{N_{cols}} + \frac{yv}{N_{rows}} \right) \right]$$

- Note that it represents the image I as a weighted sum of basis functions (sine and cosine waves).

Phase-Congruency

- Let's illustrate this situation for a 1D signal (step):



Note that at the position of the step, all those curves are in the same phase, i.e. their individual zero crossings coincide.

Phase-Congruency

- We also saw that by using the Fourier transform, a real-valued image is mapped into a complex-valued Fourier transform.

- Each complex number z is defined in polar coordinates by the amplitude $\|z\| = r = \sqrt{a^2 + b^2}$ and phase $\alpha = \arctan(b/a)$.

Phase-Congruency

- Note: at the origin in the frequency domain, we saw that $b = 0$, i.e. the phase $\alpha = 0$ for the DC component (mean of all values of I).

- Consider now a *local Fourier transform*, centered at a pixel location $p = (x, y)$ in image I using a $(2k+1)\times(2k+1)$ kernel of Fourier basis functions.

Phase-Congruency

- That is:

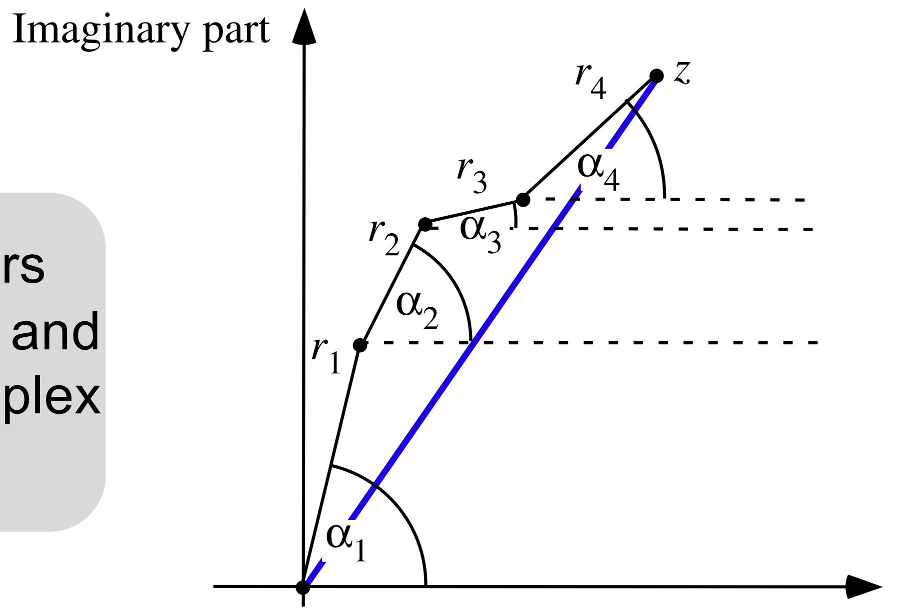
$$\mathbf{J}(u,v) = \frac{1}{(2k+1)^2}\sum_{i=-k}^{k}\sum_{j=-k}^{k} I(x+i,y+j)\cdot W_{2k+1}^{-iu}\cdot W_{2k+1}^{-jv}$$

- Ignoring the DC component, the resulting Fourier transform $\mathbf{J}$ is composed of $n=(2k+1)^2-1$ complex numbers.

Phase-Congruency

- Let those complex numbers be denoted by z_h, for $1 \le h \le n$, each defined by the amplitude $r_h = \|z_h\|_2$ and phase α_h.

Addition of four complex numbers represented by the amplitudes and phases, resulting in a complex number *z*.



Phase-Congruency

- The four complex numbers (r_h, α_h) are *roughly in phase*, meaning that the phase angles α_h do not differ much (i.e. have a small variance only).
- Such an approximate identity defines a *phase congruency*, defined by the property measure:

$$\mathscr{P}_{ideal\_phase}(p) = \frac{\|z\|_2}{\sum_{h=1}^{n} r_h}$$

Phase-Congruency

- <u>Observation</u>: *local phase congruency* identifies *features* in images, such as lines, corners and edges.

Source: R. Klette

Next Lecture

- Color and Color Images

Color Definitions. Color Perception, Visual Deficiencies and Grey Levels. Color Representations.

- Suggested reading

Section 1.3 of textbook.