Imperial College London

Digital Image Processing

Image Transforms

Unitary Transforms and the 2D Discrete Fourier Transform

DR TANIA STATHAKI

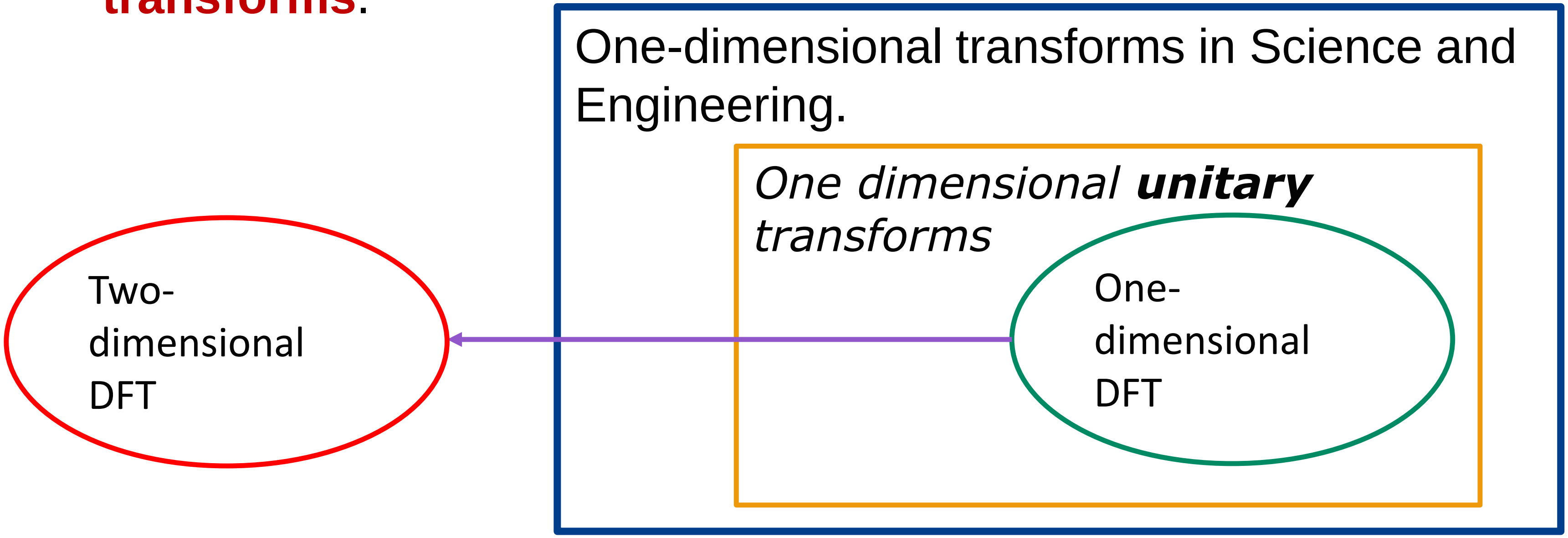
READER (ASSOCIATE PROFFESOR) IN SIGNAL PROCESSING

IMPERIAL COLLEGE LONDON

Imperial College London

What is this lecture about?

- **Welcome back to the Digital Image Processing lecture!**
- In this lecture we will learn about the Discrete Fourier Transform in images. This is a two-dimensional transform (2D DFT).
- We will briefly recall the one-dimensional Discrete Fourier Transform (1D DFT) and show how this transform can be extended into two dimensions.
- The 1D DFT is a member of a large family of transforms called **unitary transforms**.



Imperial College London

The generic form of a one-dimensional signal transform

- Consider a one-dimensional discrete signal with N samples.
- Call the signal $f(x)$, $0 \leq x \leq N-1$.
- A transform of the signal $f(x)$ will "convert" this signal into a new signal which let us call $g(u)$.
- It makes sense that if $f(x)$ has N samples, then $g(u)$ has N samples.
- In engineering and science most of the transforms we deal with have the following generic form:
  $g(u) = \sum_{x=0}^{N-1} T(u,x) f(x)$, $0 \leq u \leq N-1$, $T(u,x)$ is a function of u, x called the **forward transformation kernel**.
- Note that all values of $f(x)$ are required to produce a single value of $g(u)$.
- We can put the signals $f(x)$ and $g(u)$ in column vectors. Denote these with $\underline{f}$ and $\underline{g}$, respectively .
- Therefore, the transform can be written in the form $\underline{g} = T \cdot \underline{f}$.
- The matrix T is of dimension $N \times N$ and contains the values $T(u,x)$ for different (u,x).

Example of a one-dimensional signal transform: DFT

- Recall the generic form:
  $g(u)=\sum_{x=0}^{N-1} T(u,x)f(x)$, $0\le u\le N-1$, $T(u,x)$ is a function of u,x.
- Now recall the one-dimensional Discrete Fourier Transform (DFT)
  $F(u)=\sum_{x=0}^{N-1} f(x)\frac{1}{N}e^{-j2\pi\frac{ux}{N}}$, $0\le u\le N-1$
  Very often x represents time. We can write:

$$F(u)=\frac{1}{N}\begin{bmatrix}1 & e^{-j2\pi\frac{u}{N}} & \dots & e^{-j2\pi\frac{u(N-1)}{N}}\end{bmatrix}\cdot\begin{bmatrix}f(0)\\ f(1)\\ \vdots \\ f(N-1)\end{bmatrix}$$

- We see that for each u the signal is projected (observe the inner product) onto complex sinusoids which depend on u (the frequency of the complex sinusoid).
- Try to observe the analogy with the generic form: for the DFT we have $T(u,x)=\frac{1}{N}e^{-j2\pi\frac{ux}{N}}$.

Imperial College London

The generic form of the inverse of a one-dimensional signal transform

- Transforming a signal $f(x)$ into a signal $g(u)$ might be handy but we must have a method to get back to $f(x)$ from $g(u)$.
- Therefore, we need an inverse transform.
- The inverse transform has the generic form: $f(x) = \sum_{u=0}^{N-1} I(u,x) g(u)$, $0 \le x \le N-1$, $I(u,x)$ is a function of u, x called the inverse transformation kernel.
- Note again that all values of $g(u)$ are required to produce a single value of $f(x)$.
- The inverse transform can be written in the form $\underline{f} = I \cdot \underline{g}$.
- The matrix I is of dimension $N \times N$ and contains the values $I(u,x)$ for different (u,x).
- Note that $\underline{f} = I \cdot \underline{g} = I \cdot T \cdot \underline{f} \Rightarrow I = T^{-1}$. We must invert T to obtain I.
- Requiring to invert a matrix is a tedious task. Think that if $N = 1000$, i.e., $f(x)$ is a signal with 1000 samples, then T is of dimension 1000×1000.
- There must be a way to do this easily, otherwise using a transform would not be efficient.

The one-dimensional inverse DFT (IDFT)

- Recall the generic form:
  $f(x) = \sum_{u=0}^{N-1} I(u,x)g(u)$, $0 \le x \le N-1$, $I(u,x)$ is a function of u, x.
- Now recall the one-dimensional Inverse Discrete Fourier Transform (IDFT).
  $f(x) = \sum_{u=0}^{N-1} e^{j2\pi\frac{ux}{N}} F(u)$, $0 \le x \le N-1$
- Observe that for the IDFT we have
  $I(u,x) = e^{j2\pi\frac{ux}{N}} = N\frac{1}{N} e^{j2\pi\frac{ux}{N}} = N\frac{1}{N}(e^{-j2\pi\frac{ux}{N}})^* = N(\frac{1}{N} e^{-j2\pi\frac{ux}{N}})^* = NT^*(u,x)$.
  Therefore, $I = T^{-1} = NT^*$.
- The above relationship reveals that the inverse of matrix T is simply its complex conjugate multiplied with N, i.e., the size of the signal.
- The inverse of matrix T would be identical to T if we put $\frac{1}{\sqrt{N}}$ in both forward and inverse transform instead of $\frac{1}{N}$ and 1.
- Therefore, we don't really have to calculate any inverse matrix in the case of DFT. This is a universal observation: in all transforms we use in engineering and science, the inverse of the transformation matrix is easily obtained from the forward transformation matrix.

Unitary matrices: some brief revision of linear algebra

- A square matrix T is **unitary** if $T^H = T^{-1}$, where T^H denotes the **Hermitian** (conjugate transpose) of T, $T^H = {T^*}^T$ and T^{-1} is the matrix inverse.
- Unitary matrices leave the length of a complex vector unchanged, i.e., $\|Tx\| = \|x\|$.
- For real matrices, $T^H = T^T = T^{-1}$ and the term unitary matrix becomes the same as the term orthogonal matrix.
- The term orthogonal is quite obvious since $T^T = T^{-1}$ implies that $T^T \cdot T = \mathbf{I}$ which means that the rows/columns of T are orthogonal (and more accurately orthonormal).
- In fact, the rows r_i of a unitary matrix are a unitary basis. That is, each row has length one, and their Hermitian inner product is zero, i.e., $r_i \cdot {r_j^*}^T = 0, i \neq j$ and $r_i \cdot {r_i^*}^T = 1$.
- Similarly, the columns are also a unitary basis.
- Given any unitary basis, the matrix whose rows are that basis is a unitary matrix. It is shown that the columns are another unitary basis.
- The eigenvalues of a unitary matrix have the property $|\lambda| = 1$.

Let us go back to the one-dimensional DFT

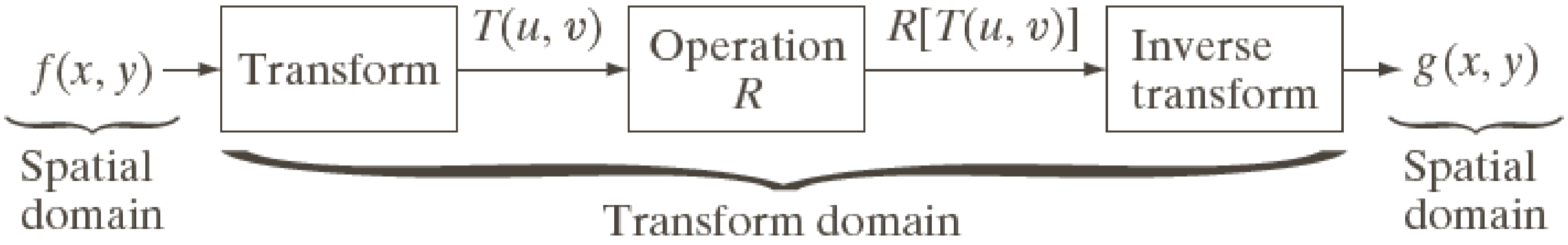
- Now recall again the one-dimensional IDFT

  $f(x) = \sum_{u=0}^{N-1} e^{j2\pi\frac{ux}{N}} F(u),\ 0 \le x \le N-1$

  and the result we have shown, i.e., $I = T^{-1} = NT^{*}$.
- In that case $T^{-1} \cdot T = \mathbf{I} = NT^{*} \cdot T = NT^{*^{T}} \cdot T$ (observe that T and T^{*} are symmetric and therefore, $T^{*} = T^{*^{T}}$.
- Therefore, $T^{*^{T}} \cdot T = \frac{1}{N}\mathbf{I}$. We see that the rows/columns of T are not orthonormal but they are orthogonal. This property is equally good to make the transformation matrix easily invertible; the inverse is simply the conjugate, multiplied by a constant factor.
- However, why are we still discussing one-dimensional DFT (1D-DFT) in an Image Processing course?
- The 1D-DFT is the building block for the implementation of the two-dimensional DFT introduced next.

Why do we use image transforms?

- Before we embark into the two-dimensional DFT (2D DFT) let us comment in general on Image Transforms. Why do we use them?
- Often, image processing tasks are best performed in a domain other than the spatial domain. The key steps are:
  - Transform the image.
  - Carry the task(s) of interest in the transformed domain.
  - Apply inverse transform to return to the spatial domain.



The generic form of a two-dimensional transform

- Consider a two-dimensional signal $f(x,y)$ $x = 0,\dots,M-1$ and $y = 0,\dots,N-1$. We can assume that $f(x,y)$ is an image.
- The generic form of a two-dimensional (image) transform is a straightforward extension from the one-dimensional case. It is defined as follows:

$$g(u,v) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} T(u,x,v,y) f(x,y)$$

- A 2D transform is **separable** if $T(u,x,v,y) = T_1(u,x)\cdot T_2(v,y)$.
- A 2D transform is **symmetric** if $T_1(u,x) = T_2(u,x)$
- The inverse transform is defined as follows:

$$f(x,y) = \sum_{u=0}^{M-1}\sum_{v=0}^{N-1} I(u,x,v,y) g(u,v)$$

- The same definitions for separability and symmetry apply to the inverse transform.

Separable and symmetric image transforms

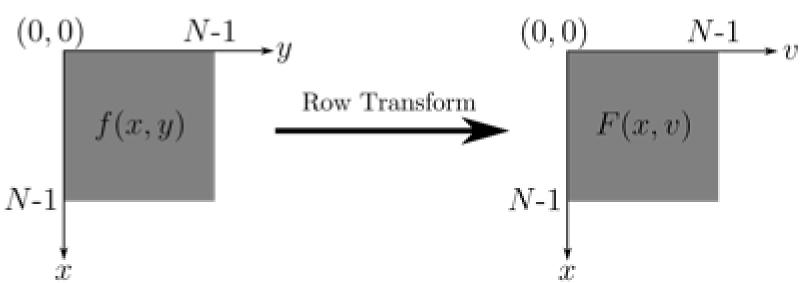
- Consider the generic image transform below. Assume that x indicates row and y indicates column.

$$g(u,v) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} T(u,x,v,y)f(x,y)$$

- Suppose that it is separable and symmetric:

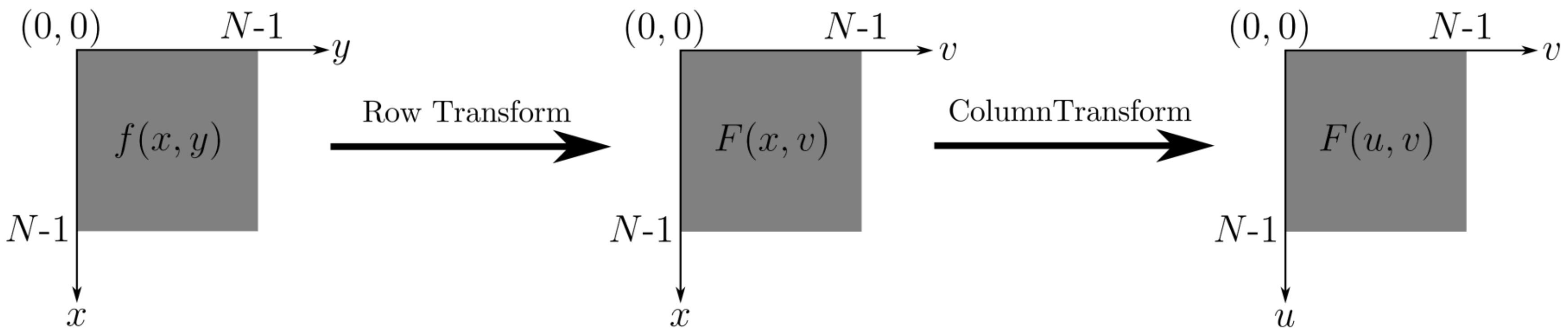
$$g(u,v) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} T(u,x)T(v,y)f(x,y)$$

- We can write $g(u,v) = \sum_{x=0}^{M-1} T(u,x) \sum_{y=0}^{N-1} T(v,y)f(x,y)$
- The sum $\sum_{y=0}^{N-1} T(v,y)f(x,y)$, $v = 0, \dots, N-1$ is basically the corresponding one-dimensional transform applied along the x –th row of the image.
- By applying the one-dimensional transform in each row $x = 0, \dots, M-1$ we obtain an intermediate image $F(x,v)$.



Separable and symmetric image transforms cont.

- We can now write $g(u,v) = \sum_{x=0}^{M-1} T(u,x)F(x,v)$
- The above sum is now the corresponding one-dimensional transform applied along the v −th column of the intermediate image $F(x,v)$.
- We see, therefore, that the implementation of a separable and symmetric transform in an image requires the sequential implementation of the corresponding one-dimensional transform row-by-row and then column-by-column (or the inverse).
- The above process is depicted in the figure below.



Two crucial properties of most signal and image transforms

- **Energy preservation.** A transform possesses the property of energy preservation if both the original and the transformed signal have the same energy, i.e., $\left\|\underline{g}\right\|^2 = \left\|\underline{f}\right\|^2$. For images this property can be written as $\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}|f(x,y)|^2 = \sum_{u=0}^{M-1}\sum_{v=0}^{N-1}|g(u,v)|^2$.
- **Energy compaction.** Most of the energy of the original data is concentrated in only a few transform coefficients, which are placed close to the origin; remaining coefficients have small values.
- The property of energy compaction facilitates the compression of the original image. A generic compression scheme based on this idea can be summarized in the following steps.
  - Transform the image.
  - Keep a small fraction of the transformed image values close to the origin (the $(0,0)$ point). That way you save space.
  - For reconstruction, replace the values that you discarded with zeros and use the inverse transform.

Two-dimensional Discrete Fourier Transform (2D DFT)

- Consider an image $f(x,y)$ of size $M \times N$. The 2D DFT is defined as follows:

$$F(u,v) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\, e^{-j2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)}$$

- The 2D Inverse DFT (2D IDFT) is defined as:

$$f(x,y) = \sum_{u=0}^{M-1}\sum_{v=0}^{N-1} F(u,v)\, e^{j2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)}$$

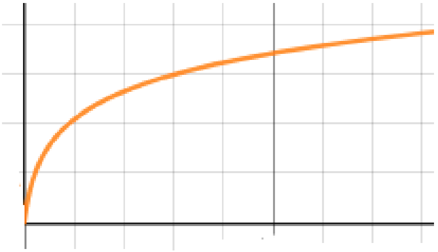
- It is separable and symmetric, with its one-dimensional version being (almost) unitary.
- Therefore, the one-dimensional DFT and more specifically the FFT, can be used for the implementation of the two-dimensional DFT (2D FFT.

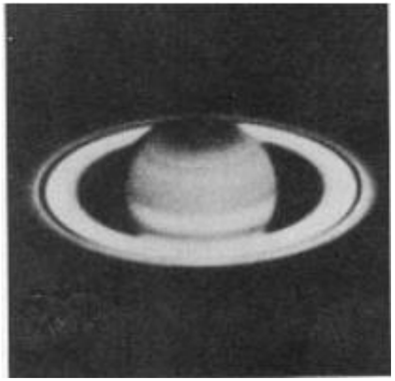
Two-dimensional Discrete Fourier Transform (2D DFT) cont.

- The location of the $\frac{1}{MN}$ is not important.
- Sometimes it is located in front of the inverse transform.
- Other times it is found split into two equal terms of $\frac{1}{\sqrt{MN}}$ multiplying the transform and its inverse.
- In the cases of square images $M = N$.
- Very often in Image Processing we work with square images whose size is a power of 2. Powers of 2 facilitate easier implementation of DSP algorithms as you already know.
- The amplitude spectrum of 2D DFT is $|F(u,v)| = [R^2(u,v) + I^2(u,v)]^{1/2}$, $R(u,v)$ and $I(u,v)$ are the real and imaginary parts of $F(u,v)$, respectively.
- The phase spectrum is $\phi(u,v) = \tan^{-1}\left[\frac{I(u,v)}{R(u,v)}\right]$
- The power spectrum is $P(u,v) = |F(u,v)|^2 = R^2(u,v) + I^2(u,v)$

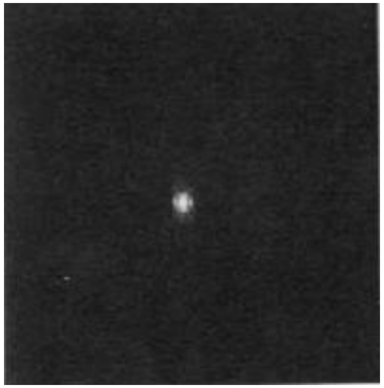
The visualisation of the range of values of 2D DFT

- The range of values of $F(u,v)$ is typically very large.
- Due to quantization of $F(u,v)$ small values are not distinguishable when we attempt to display the amplitude of $F(u,v)$.
- We, therefore, apply a logarithmic transformation to enhance small values.
  $D(u,v) = c\log(1 + |F(u,v)|)$
- The parameter c is chosen so that the range of $D(u,v)$ is $[0,255]$, i.e.,
  $c = \frac{255}{\log(1+\max\{|F(u,v)|\})}$.

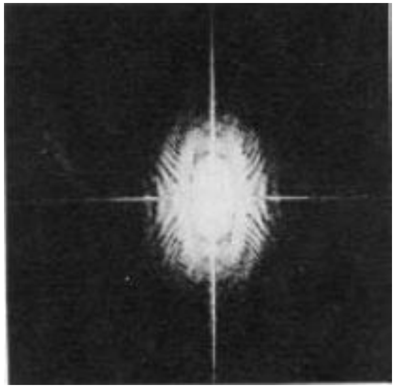




Original image

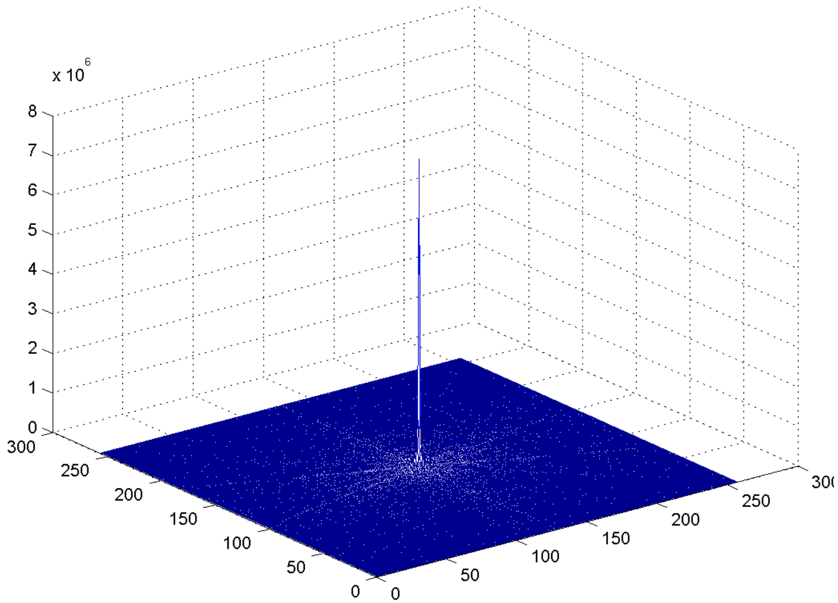


Display of amplitude of DFT

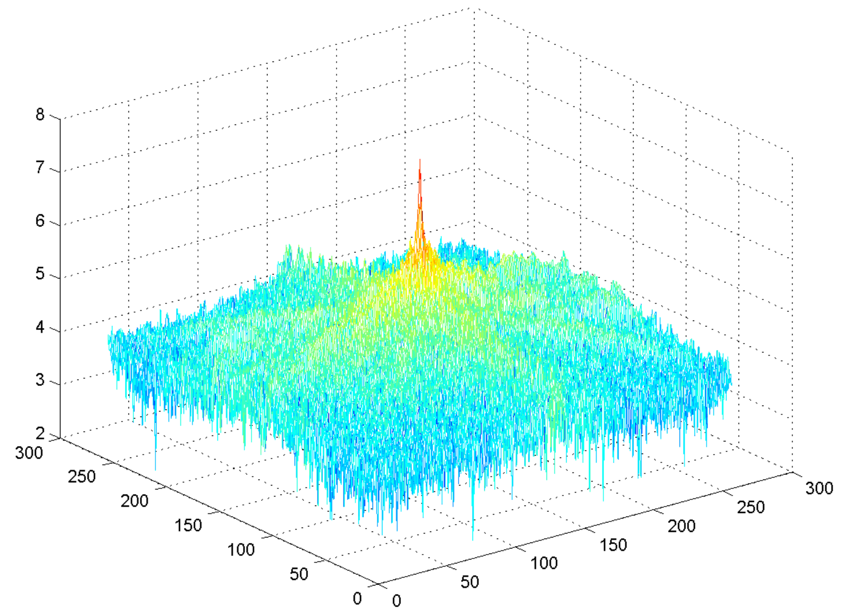


Display of the logarithmic amplitude of DFT

The visualisation of the range of values of 2D DFT cont.

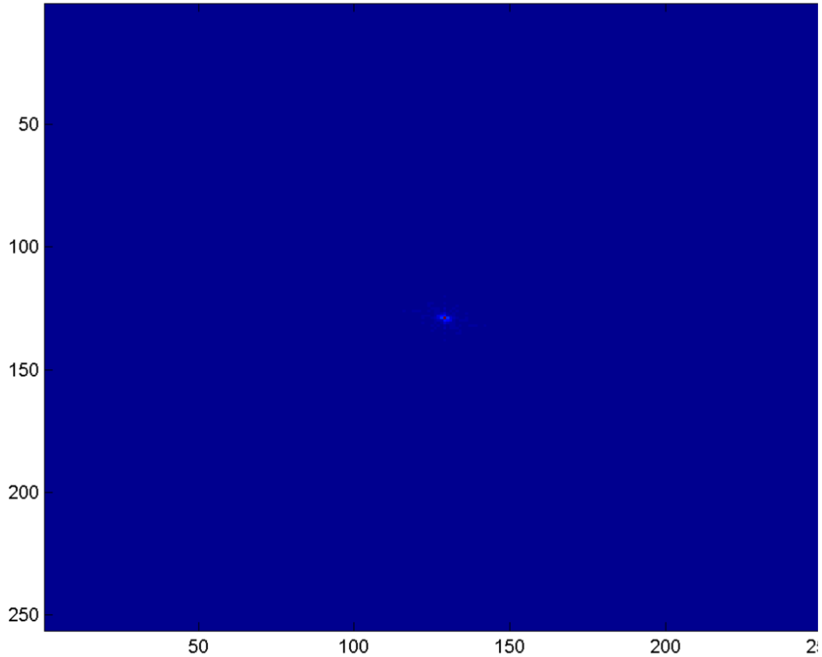


Display of amplitude of DFT

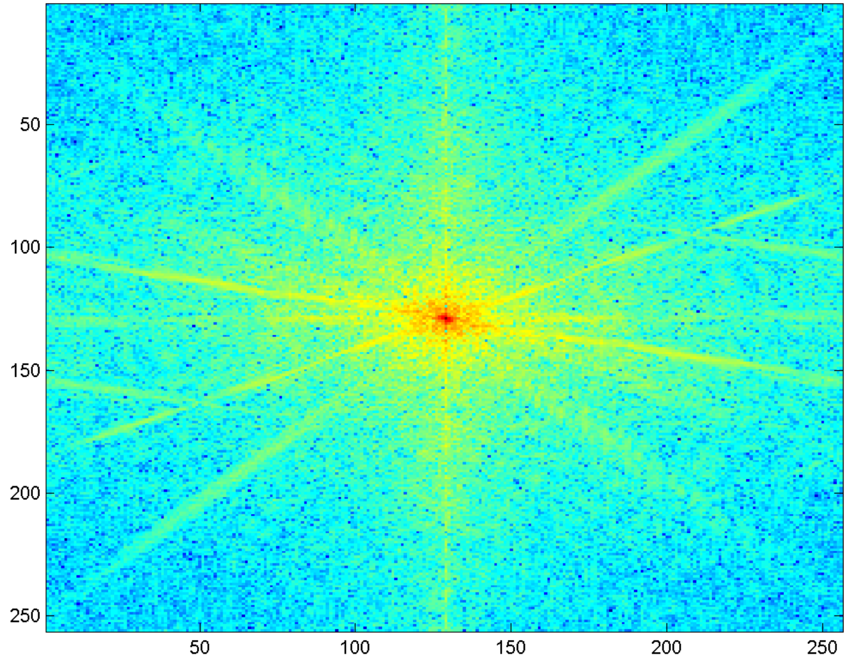


Display of the logarithmic amplitude of DFT

The visualisation of the range of values of the 2D DFT cont.



Display of amplitude of DFT



Display of the logarithmic amplitude of DFT

Properties of 2D DFT

- **Periodicity**: The 2D DFT and its inverse are periodic
  $F(u,v) = F(u+M,v) = F(u,v+N) = F(u+M,v+N)$
  Therefore, use of DFT implies virtual periodicity in space.
- **Conjugate symmetry**: $F(u,v) = F^*(-u+pM,-v+qN)$ with p,q any integers. This property also implies that $|F(u,v)| = |F(-u,-v)|$.
- If $f(x,y)$ is real and even then $F(u,v)$ is real and even.
- If $f(x,y)$ is real and odd then $F(u,v)$ is imaginary and odd.
- $\mathcal{F}\{f(x,y)+g(x,y)\} = \mathcal{F}\{f(x,y)\} + \mathcal{F}\{g(x,y)\}$ where $\mathcal{F}\{\cdot\}$ indicates the Discrete Fourier Transform operator.
- $\mathcal{F}\{f(x,y)\cdot g(x,y)\} \neq \mathcal{F}\{f(x,y)\}\cdot\mathcal{F}\{g(x,y)\}$
- **Translation** in spatial and frequency domain:

  $$f(x-x_0, y-y_0) \leftrightarrow F(u,v)e^{-j2\pi\left(\frac{ux_0}{M}+\frac{vy_0}{N}\right)}$$

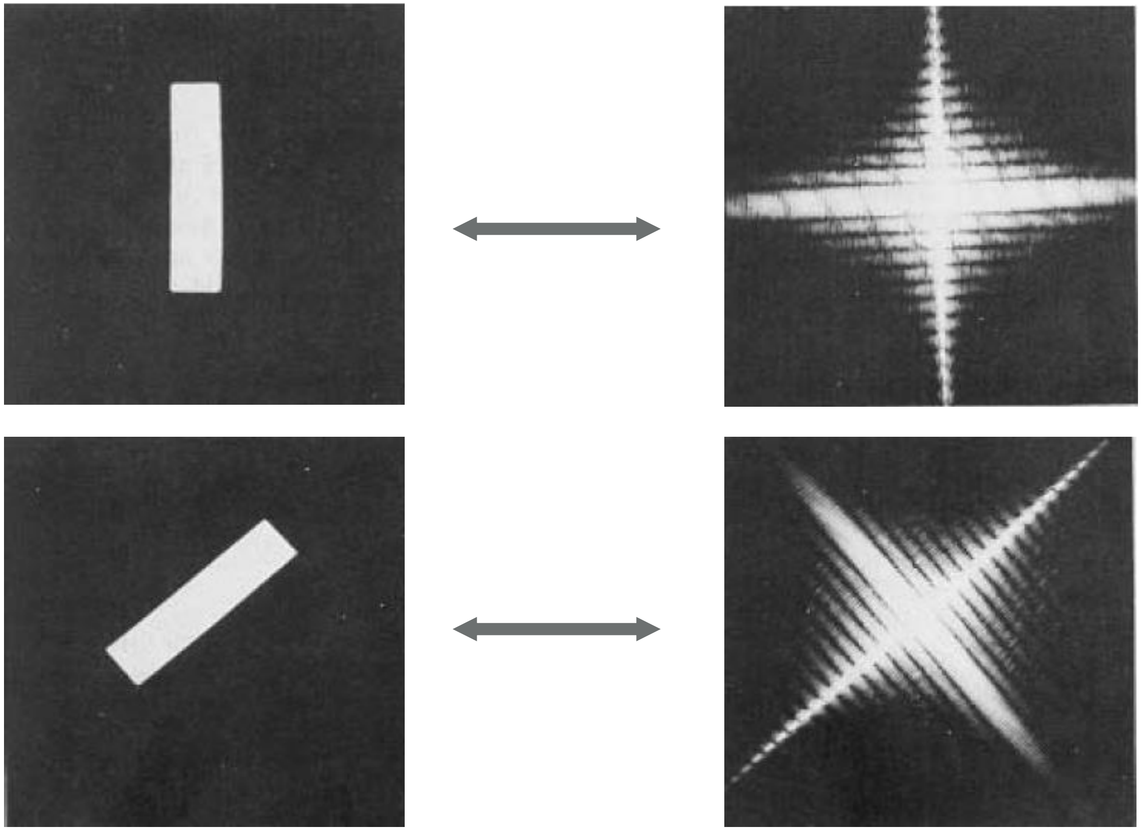
  $$f(x,y)e^{j2\pi\left(\frac{u_0x}{M}+\frac{v_0y}{N}\right)} \leftrightarrow F(u-u_0, v-v_0)$$

- Average value of the signal $\bar{f}(x,y) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)$

  $$F(0,0) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y) \Rightarrow \bar{f}(x,y) = F(0,0)$$

Properties of 2D DFT cont.

- **Rotation**: Rotating $f(x,y)$ by θ rotates $F(u,v)$ by θ.

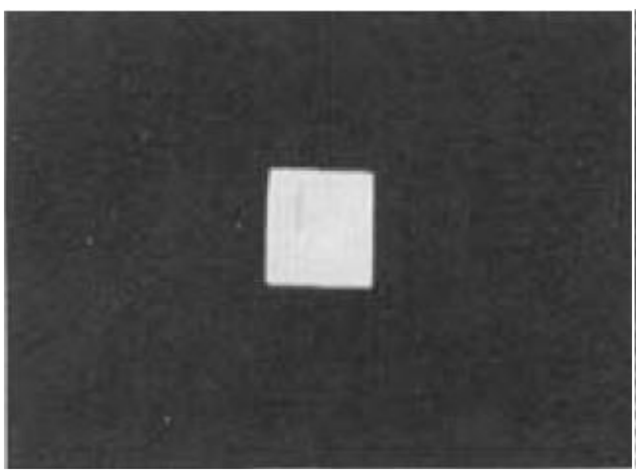


Visualisation of a full period of 2D DFT

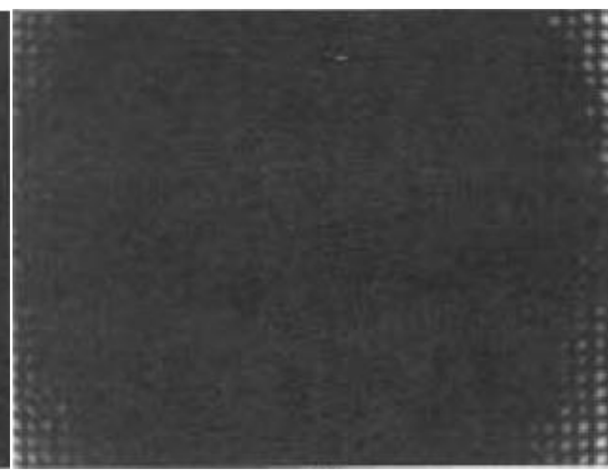
- In order to display a full period of the 2D DFT in the center of the image, we need to translate the origin of the transform at $(u,v)=(M/2,N/2)$.
- To move $F(u,v)$ at $(M/2,N/2)$ we use the second translation property in the previous slide with $u_0=M/2$ and $v_0=N/2$.

In that case $\mathcal{F}\left\{f(x,y)e^{j2\pi\left(\frac{u_0 x}{M}+\frac{v_0 y}{N}\right)}\right\}=F(u-u_0,v-v_0)$ becomes

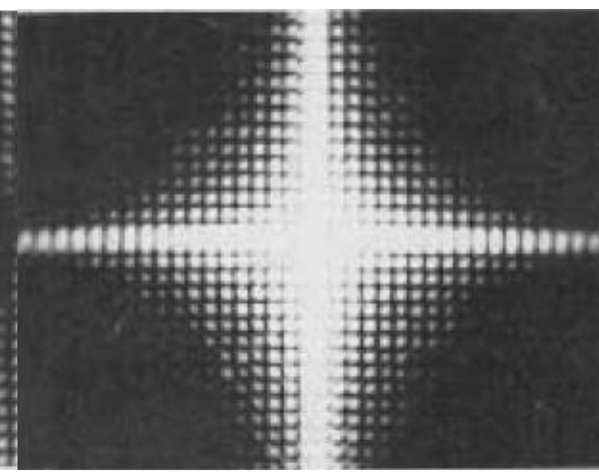
$$\mathcal{F}\{f(x,y)e^{j\pi(x+y)}\}=\mathcal{F}\{f(x,y)(-1)^{(x+y)}\}=F(u-M/2,v-N/2)$$



Original image

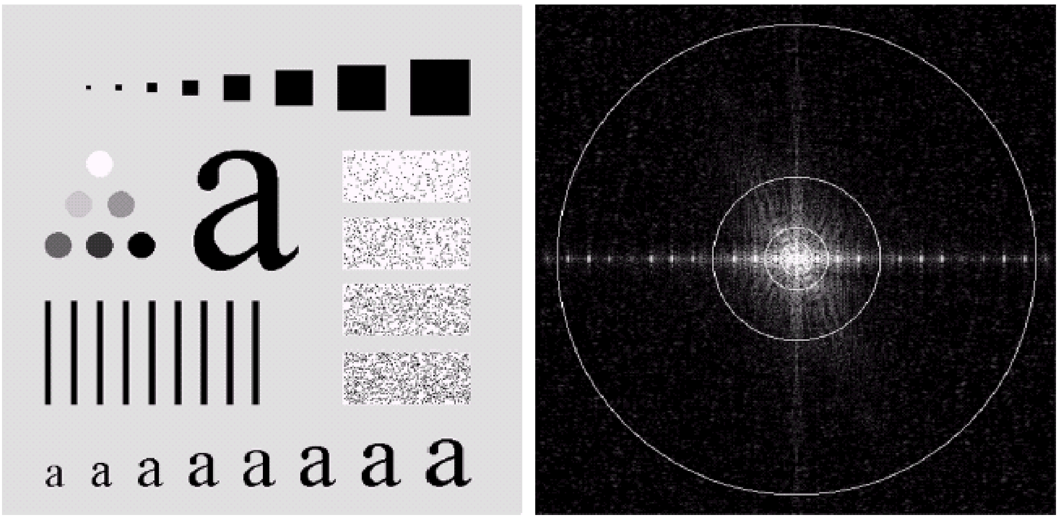


Original DFT amplitude



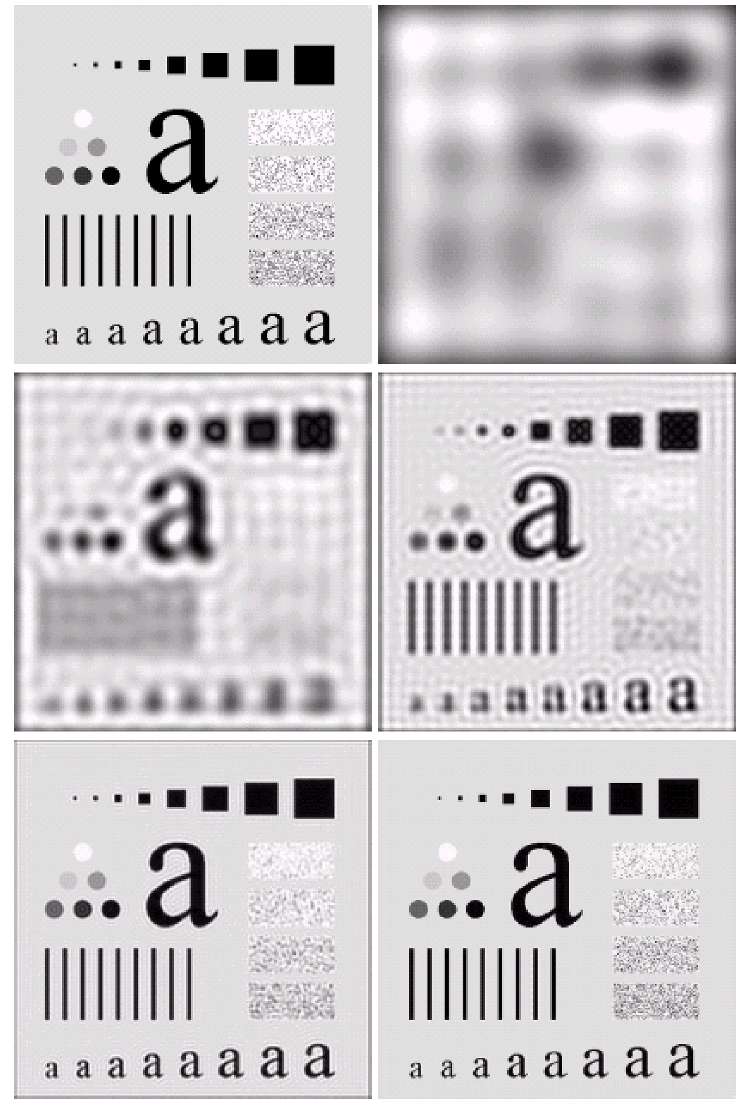
Translated DFT amplitude

Applications:
Exploiting the property of energy compaction of 2D DFT



a b

FIGURE 4.11 (a) An image of size 500 × 500 pixels and (b) its Fourier spectrum. The superimposed circles have radii values of 5, 15, 30, 80, and 230, which enclose 92.0, 94.6, 96.4, 98.0, and 99.5% of the image power, respectively.



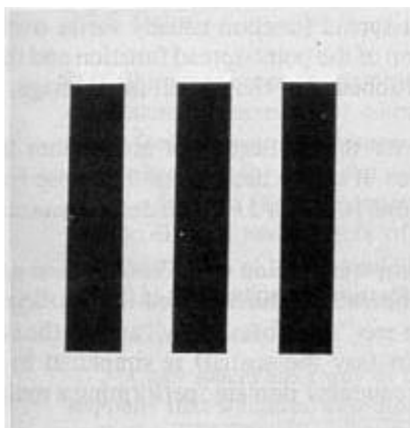
a b
c d
e f

FIGURE 4.12 (a) Original image. (b)–(f) Results of ideal lowpass filtering with cutoff frequencies set at radii values of 5, 15, 30, 80, and 230, as shown in Fig. 4.11(b). The power removed by these filters was 8, 5.4, 3.6, 2, and 0.5% of the total, respectively.

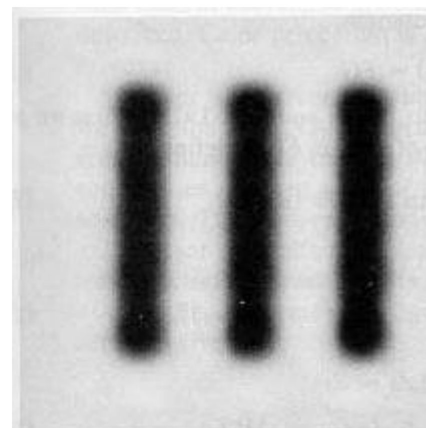
Imperial College London

Applications: Low- and high- pass filtering: How do frequencies show up in an image?

- Low frequencies correspond to slowly varying pixel intensities (e.g., smooth surfaces).
- High frequencies correspond to quickly varying pixel intensities (e.g., edges)



original image



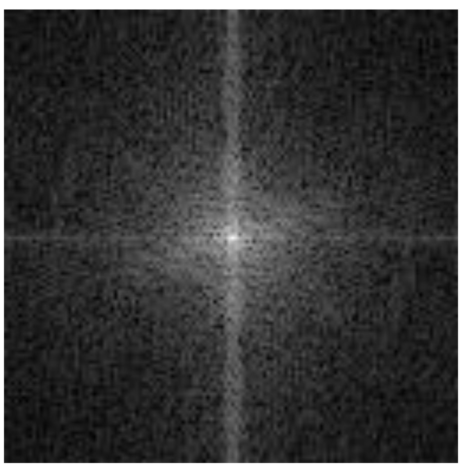
low-pass filtered image

Low- and high- pass filtering using amplitude of 2D DFT

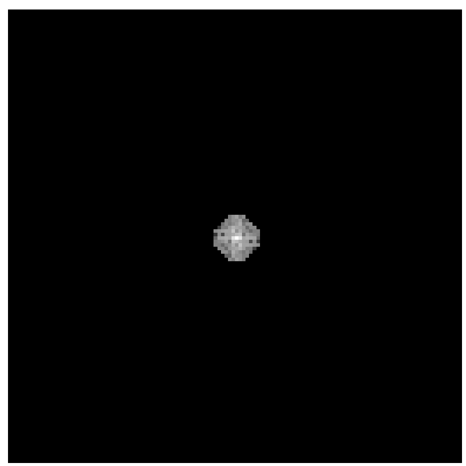
original image

low pass filtered

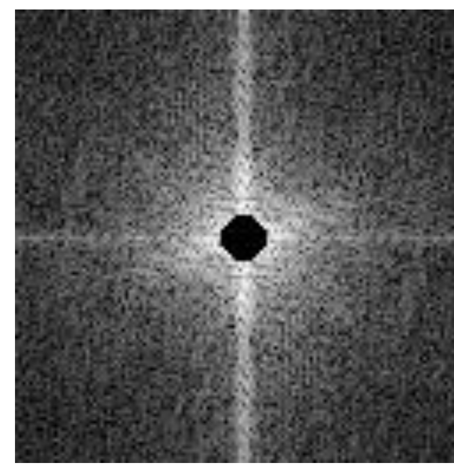
high pass filtered



amplitude of DFT



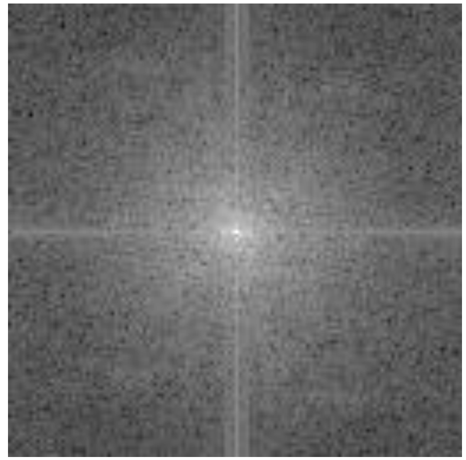
amplitude of DFT with eliminated high frequencies



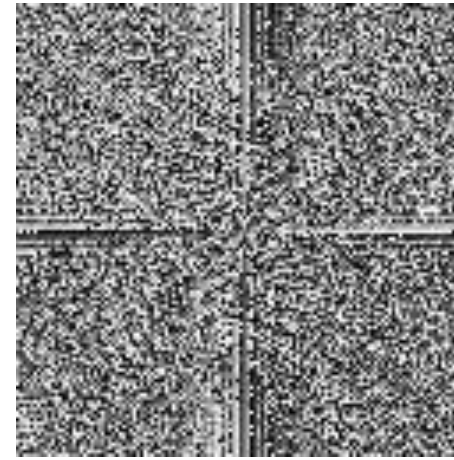
amplitude of DFT with eliminated low frequencies

Amplitude and phase of 2D DFT

original image

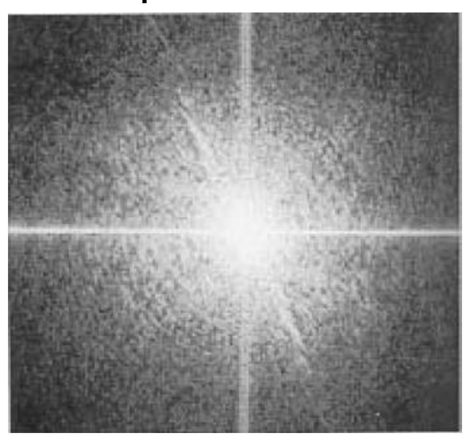


amplitude of DFT

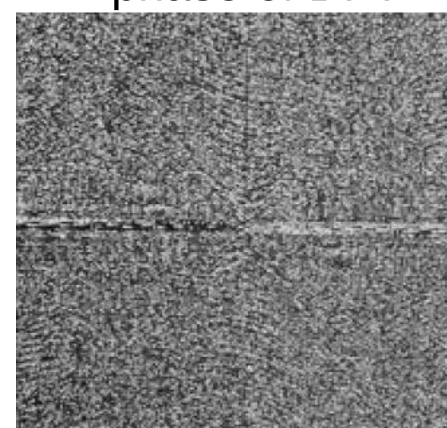


phase of DFT

amplitude of DFT



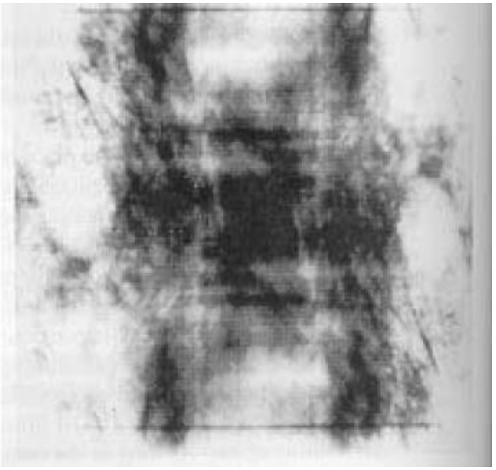
amplitude of DFT



phase of DFT

Exercise: use IDFT to reconstruct the image using magnitude or phase only information

Amplitude and phase of 2D DFT cont.



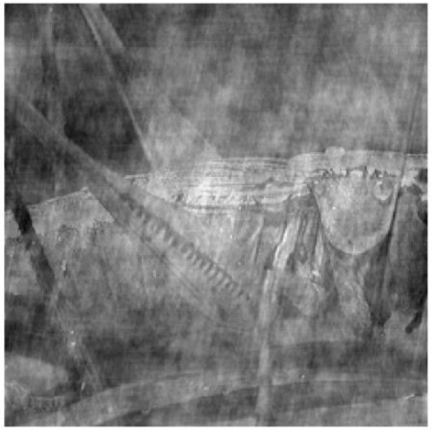
Reconstructed image using the amplitude of DFT only and zero phase. The magnitude determines the strength of each frequency component but not its location.

Reconstructed image using the phase of DFT only and amplitude equal to 1.

- The phase determines the locations of individual frequencies within the image.
- High frequencies → abrupt changes → edges → object silhouettes → image content.

Reconstruction from the amplitude of one image and the phase of another

phase of cameraman
amplitude of grasshopper



phase of grasshopper
amplitude of cameraman

Reconstruction from the amplitude of one image and the phase of another cont.

phase of buffalo
amplitude of rocks

phase of rocks
amplitude of buffalo