Image Transforms

4.1 INTRODUCTION

Definition: Image transform = operation to change the default representation space of a digital image (spatial domain -> another domain) so that:

(1) all the information present in the image is preserved in the transformed domain, but represented differently;

(2) the transform is reversible, i.e., we can revert to the spatial domain

Generally: in the transformed domain -> image information is represented in a more compact form => main goal of the transforms: **image compression.**

Other usage: image analysis - a new type of representation of different types of information present in the image.

Note: Most image transforms = "generalizations" of frequency transforms => the representation of the image by a DC component and several AC components.

Definition: "original representation space" of the image **U**[M×N] = a MN-dimensional space:

- each coordinate of the space = a spatial location (m,n) in the digital image;
- the value of the coordinate of **U** on an axis = the grey level in **U** in this spatial location (m,n).

x_1=(0,0); x_2=(0,1); x_3=(0,2); ... x_{MN}=(M-1,N-1).

=> A **unitary transform** of the image **U** = a **rotation** of the MN-dimensional space, defined by a rotation matrix **A** in MN-dimensions.

$\{u(n), \quad 0 \le n \le N-1\}$; **A** – unitary matrix, $\mathbf{A}^{-1} = \mathbf{A}^{*T}$

$$\mathbf{v} = \mathbf{A}\mathbf{u}, \quad or \quad v(k) = \sum_{n=0}^{N-1} a(k,n)u(n), \quad 0 \le k \le N-1 \qquad (4.1)$$

$$\mathbf{u} = \mathbf{A}^{*T}\mathbf{v} \quad or \quad u(n) = \sum_{k=0}^{N-1} a^{*}(k,n)v(k), \quad 0 \le n \le N-1 \qquad (4.2)$$

$$a_k^{*} = \{a^{*}(k,n), \quad 0 \le n \le N-1\},$$

4.2 UNITARY ORTHOGONAL TWO-DIMENSIONAL TRANSFORMS

$$v(k,l) = \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} a_{k,l}(m,n) \cdot u(m,n), \quad 0 \le k,l \le N-1 \qquad (4.3)$$

$$u(m,n) = \sum_{k=0}^{N-1}\sum_{l=0}^{N-1} a_{k,l}^{*}(m,n) \cdot v(k,l), \quad 0 \le k,l \le N-1 \qquad (4.4)$$

Mathematical Background: Complex Numbers (cont'd)

- Euler's formula

$$e^{\pm j\theta} = cos(\theta) \pm jsin(\theta)$$

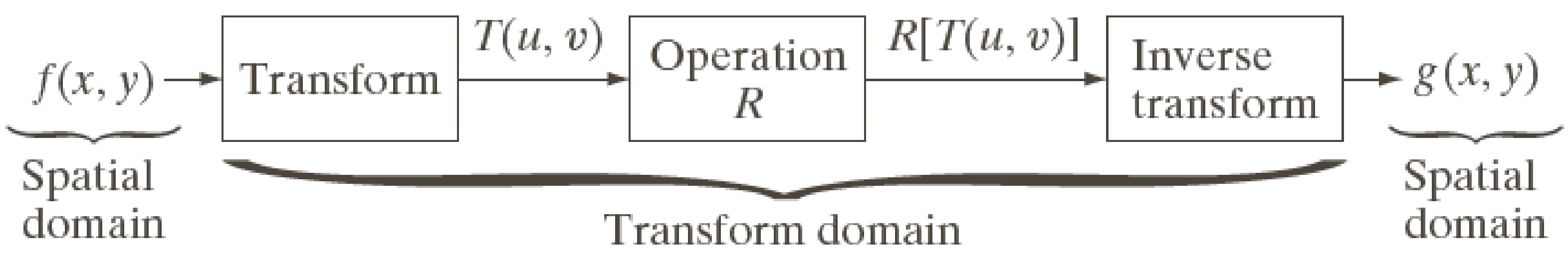
Image Transforms

- Many times, image processing tasks are best performed in a domain other than the *spatial domain.*
- Key steps:

  (1) Transform the image

  (2) Carry the task(s) in the *transformed domain*.

  (3) Apply *inverse transform* to return to the spatial domain.



Notation

- Continuous Fourier Transform (FT)
- Discrete Fourier Transform (DFT)
- Discrete Cosine Transform (DCT)

Fourier Series Theorem

- Any periodic function f(t) can be expressed as a weighted sum (infinite) of sine and cosine functions of varying frequency:

$$f(t) = a_0 + \sum_{n=1}^{\infty} a_n cos(nf_0 t) + \sum_{n=1}^{\infty} b_n sin(nf_0 t)$$

f_0 is called the "fundamental frequency"

Continuous Fourier Transform (FT)

- Transforms a signal (i.e., function) from the **spatial** (x) domain to the **frequency** (u) domain.

Forward FT: $$F(f(x)) = F(u) = \int_{-\infty}^{\infty} f(x) e^{-j2\pi ux} dx$$

Inverse FT: $$F^{-1}(F(u)) = f(x) = \int_{-\infty}^{\infty} F(u) e^{j2\pi ux} du$$

(IFT)

where $$e^{\pm j\theta} = cos(\theta) \pm jsin(\theta)$$

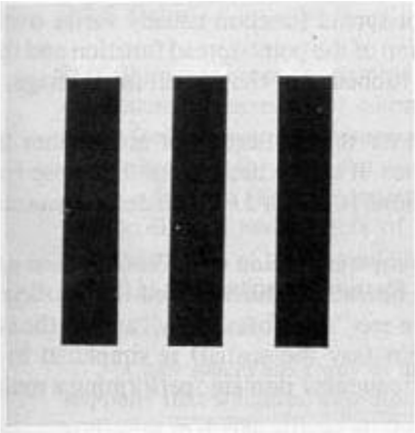
Why is FT Useful?

- Easier to remove undesirable frequencies.

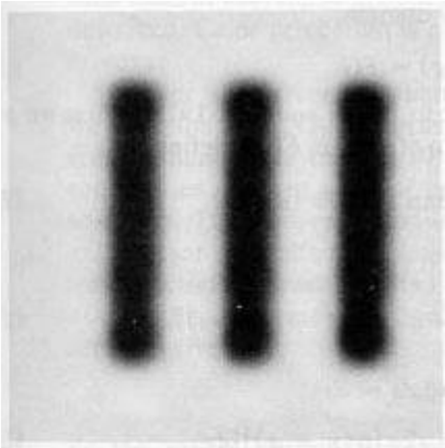
- Faster perform certain operations in the **frequency** domain than in the **spatial** domain.

How do frequencies show up in an image?

- Low frequencies correspond to slowly varying information (e.g., continuous surface).
- High frequencies correspond to quickly varying information (e.g., edges)

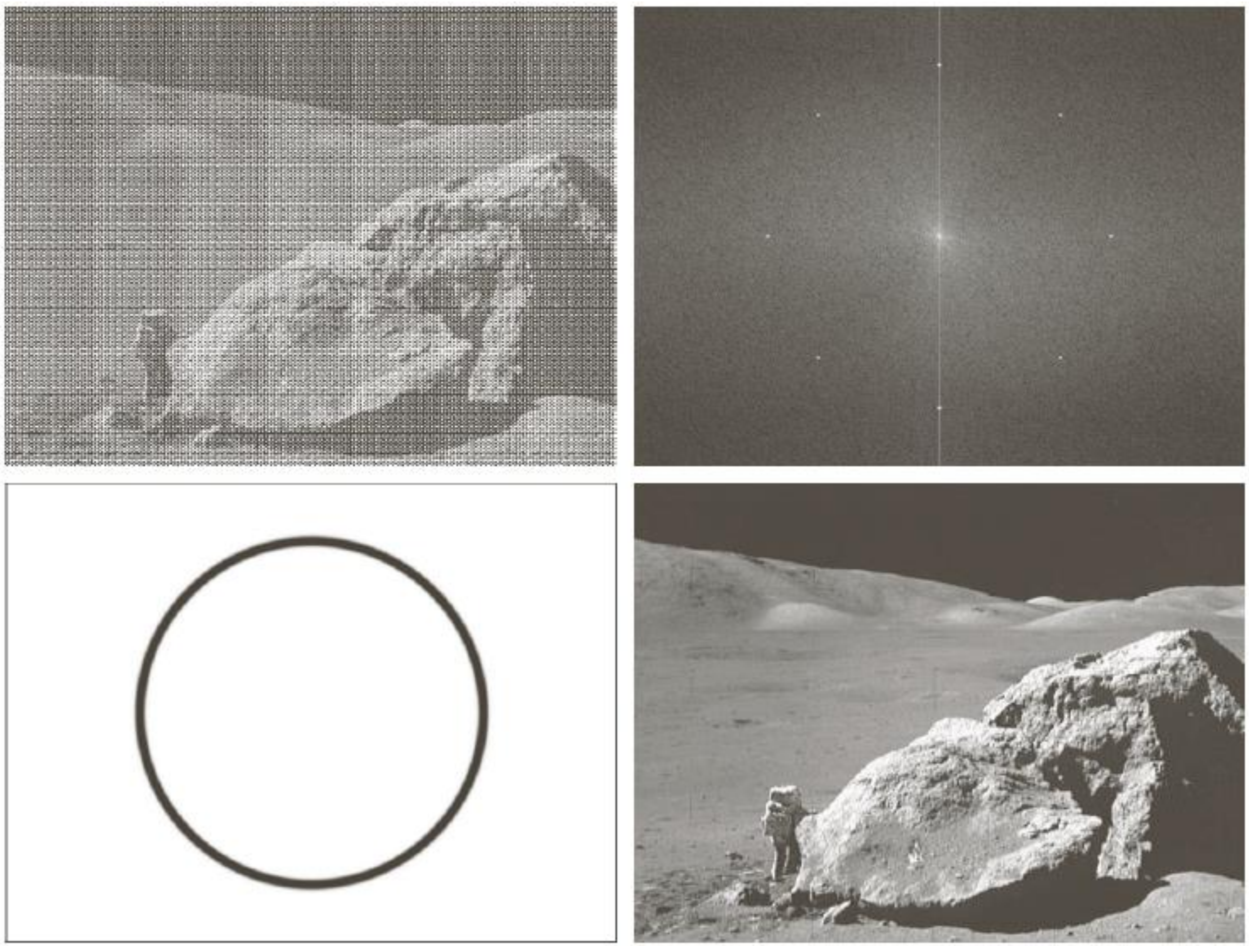


Original Image



Low-passed

Example of noise reduction using FT



Extending FT in 2D

- Forward FT

$$F(f(x,y)) = F(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)e^{-j2\pi(ux+vy)}dxdy$$

- Inverse F

$$F^{-1}(F(u,v)) = f(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u,v)e^{j2\pi(ux+uy)}dudv$$

Discrete Fourier Transform (DFT) (cont'd)

- Forward DFT

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

- Inverse DFT

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

Extending DFT to 2D

- Assume that f(x,y) is M x N.
- Forward DFT

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

$$(u = 0, 1, \ldots, M-1,\ v = 0, 1, \ldots, N-1)$$

- Inverse DFT:

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

$$(x = 0, 1, \ldots, M-1,\ y = 0, 1, \ldots, N-1)$$

Extending DFT to 2D (cont'd)

- Special case: f(x,y) is N x N.

- Forward DFT

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

u,v = 0,1,2, …, N-1

- Inverse DFT

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

x,y = 0,1,2, …, N-1

Visualizing DFT

- Typically, we visualize |F(u,v)|
- The dynamic range of |F(u,v)| is typically very large

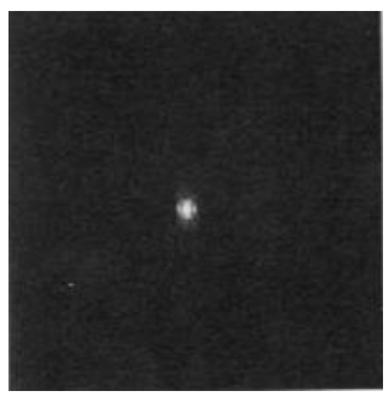
- Apply streching: t)

$$D(u,v) = c\ log(1 + |F(u,v)|)$$

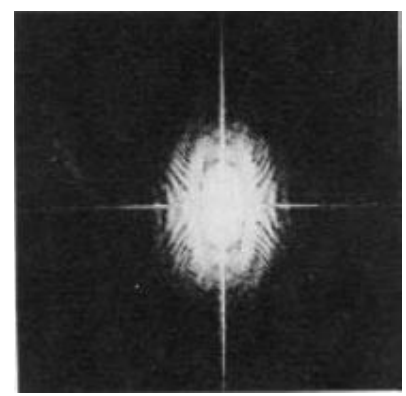
|F(u,v)|

|D(u,v)|

original image



before stretching



after stretching

2-D DCT using a 1-D DCT Pair

- **1-D DCT:**

$$X(k) = \sqrt{\frac{2}{N}}C(k)\sum_{i=0}^{N-1} x(i)\cos[\frac{(2i+1)k\pi}{2N}]$$

- **1-D IDCT:**

$$x(i) = \sqrt{\frac{2}{N}}\sum_{k=0}^{N-1} C(k)X(k)\cos[\frac{(2i+1)k\pi}{2N}]$$

k = 0, 1, 2, …, N-1.
and i = 0, 1, 2, …, N-1.

$$C(k) = \begin{cases} \frac{1}{\sqrt{2}} & \text{for } k = 0 \\ 1 & \text{otherwise.} \end{cases}$$

Implementation of the DCT

- DCT-based codecs use a two-dimensional version of the transform.
- The 2-D DCT and its inverse (IDCT) of an N x N block are shown below:

- **2-D DCT:**

$$F(u,v) = \frac{2}{N} C(u)C(v) \sum_{y=0}^{N-1} \sum_{x=0}^{N-1} f(x,y) \cos[\frac{(2x+1)u\pi}{2N}] \cos[\frac{(2y+1)v\pi}{2N}]$$

- **2-D IDCT:**

$$f(x,y) = \frac{2}{N} \sum_{v=0}^{N-1} \sum_{u=0}^{N-1} C(u)C(v)F(u,v) \cos[\frac{(2x+1)u\pi}{2N}] \cos[\frac{(2y+1)v\pi}{2N}]$$

$$C(k) = \begin{cases} \frac{1}{\sqrt{2}} & \text{for } k = 0 \\ 1 & \text{otherwise.} \end{cases}$$

- **Note: The DCT is similar to the DFT since it decomposes a signal into a series of harmonic cosine functions.**