UNIT II

IMAGE TRANSFORMS

- Need for transform
- 2D Orthogonal and Unitary transform and its properties
- 1D & 2D DFT – Properties – separability,translation, periodicity, conjugate symmetry, rotation, scaling,convolution and correlation
- Separable transforms
    - Walsh
    - Hadamard
    - Haar
    - Discrete Sine
    - DCT
    - Slant
    - SVD
    - KL transforms .

Need for transform

The need for transform is most of the signals or images are time domain signal (ie) signals can be measured with a function of time. This representation is not always best. For most image processing applications anyone of the mathematical transformation are applied to the signal or images to obtain further information from that signal.

2D Orthogonal and Unitary transform and its properties

As a one dimensional signal can be represented by an orthonormal set of basis vectors, an image can also be expanded in terms of a discrete set of basis arrays called basis images through a two dimensional (image) transform. For an N X N image f(x,y) the forward and inverse transforms are given below

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

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

where, again, $T(u,v,x,y)$ and $I(x,y,u,v)$ are called the **forward and inverse transformation kernels**, respectively.

The forward kernel is said to be **separable** if

$$T(u,v,x,y)=T_1(u,x)T_2(v,y)$$

It is said to be **symmetric** if T_1 is functionally equal to T_2 such that

$$T(u,v,x,y)=T_1(u,x)T_1(v,y)$$

The same comments are valid for the inverse kernel.

If the kernel $T(u,v,x,y)$ of an image transform is separable and symmetric, then the transform $g(u,v)=\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}T(u,v,x,y)f(x,y)=\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}T_1(u,x)T_1(v,y)f(x,y)$ can be written in matrix form as follows

$$\underline{g}=\underline{T}_1\cdot\underline{f}\cdot\underline{T}_1^{\,T}$$

where $\underline{f}$ is the original image of size $N\times N$, and $\underline{T}_1$ is an $N\times N$ transformation matrix with elements $t_{ij}=T_1(i,j)$. If, in addition, $\underline{T}_1$ is a unitary matrix then the transform is called **separable unitary** and the original image is recovered through the relationship

$$\underline{f}=\underline{T}_1^{\,*T}\cdot\underline{g}\cdot\underline{T}_1^{\,*}$$

Fundamental properties of unitary transforms

The property of energy preservation

In the unitary transformation

$$\underline{g} = \underline{T} \cdot \underline{f}$$

it is easily proven that

$$\|\underline{g}\|^2 = \|\underline{f}\|^2$$

Thus, a unitary transformation preserves the signal energy. This property is called energy preservation property. This means that every unitary transformation is simply a rotation of the vector f in the N - dimensional vector space.

The property of energy compaction

Most unitary transforms pack a large fraction of the energy of the image into relatively few of the transform coefficients. This means that relatively few of the transform coefficients have significant values and these are the coefficients that are close to the origin (small index coefficients).

1D & 2D DFT

Any function that periodically reports itself can be expressed as a sum of sines and cosines of different frequencies each multiplied by a different coefficient, this sum is called Fourier series. Even the functions which are non periodic but whose area under the curve if finite can also be represented in such form; this is now called Fourier transform. A function represented in either of these forms and can be completely reconstructed via an inverse process with no loss of information.

The discrete Fourier transform (DFT) of a sequence $\{u(n), n = 0, \ldots, N-1\}$ is defined as

$$v(k) = \sum_{n=0}^{N-1} u(n) W_N^{kn}, \qquad k = 0, 1, \ldots, N-1$$

where

$$W_N \triangleq \exp\left\{\frac{-j2\pi}{N}\right\}$$

The inverse transform is given by

$$u(n) = \frac{1}{N} \sum_{k=0}^{N-1} v(k) W_N^{-kn}, \qquad n = 0, 1, \ldots, N-1$$

The two-dimensional DFT of an $N \times N$ image $\{u(m, n)\}$ is a separable transform defined as

$$v(k,l) = \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} u(m,n) W_N^{km} W_N^{ln}, \qquad 0 \le k, l \le N-1$$

and the inverse transform is

$$u(m,n) = \frac{1}{N^2}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} v(k,l) W_N^{-km} W_N^{-ln}, \qquad 0 \le m, n \le N-1$$

The two-dimensional unitary DFT pair is defined as

$$v(k,l) = \frac{1}{N}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} u(m,n) W_N^{km} W_N^{ln}, \qquad 0 \le k, l \le N-1$$

$$u(m,n) = \frac{1}{N}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} v(k,l) W_N^{-km} W_N^{-ln}, \qquad 0 \le m, n \le N-1$$

In matrix notation this becomes

$$\mathbf{V} = \mathbf{F}\mathbf{U}\mathbf{F}$$

The properties of the two-dimensional unitarv DFT are quite similar to the one-dimensional case and are summarized next.

Symmetric, unitary.

$$\mathcal{F}^T = \mathcal{F}, \qquad \mathcal{F}^{-1} = \mathcal{F}^* = \mathbf{F}^* \otimes \mathbf{F}^*$$

Periodic extensions.

$$v(k+N, l+N) = v(k,l), \qquad \forall k, l$$

$$u(m+N, n+N) = u(m,n), \qquad \forall m, n$$

Sampled Fourier spectrum. If $\tilde{u}(m,n) = u(m,n), 0 \le m, n \le N-1$, and $\tilde{u}(m,n) = 0$ otherwise, then

$$\tilde{U}\left(\frac{2\pi k}{N}, \frac{2\pi l}{N}\right) = \text{DFT}\{u(m,n)\} = v(k,l)$$

where $\tilde{U}(\omega_1, \omega_2)$ is the Fourier transform of $\tilde{u}(m,n)$.

Fast transform. Since the two-dimensional DFT is separable, the transformation of ... is equivalent to $2N$ one-dimensional unitary DFTs, each of which can be performed in $O(N \log_2 N)$ operations via the FFT. Hence the total number of operations is $O(N^2 \log_2 N)$.

Conjugate symmetry. The DFT and unitary DFT of *real images* exhibit conjugate symmetry, that is,

$$v\left(\frac{N}{2} \pm k, \frac{N}{2} \pm l\right) = v^*\left(\frac{N}{2} \mp k, \frac{N}{2} \mp l\right), \qquad 0 \le k, l \le \frac{N}{2} - 1$$

THE DISCRETE COSINE TRANSFORM (DCT)

This is a transform that is similar to the Fourier transform in the sense that the new independent variable represents again frequency. The DCT is defined below.

$$C(u) = a(u) \sum_{x=0}^{N-1} f(x) \cos\left[\frac{(2x+1)u\pi}{2N}\right], \; u = 0,1,\ldots,N-1$$

with a(u) a parameter that is defined below.

$$a(u) = \begin{cases} \sqrt{1/N} & u = 0 \\ \sqrt{2/N} & u = 1,\ldots,N-1 \end{cases}$$

The inverse DCT (IDCT) is defined below.

$$f(x) = \sum_{u=0}^{N-1} a(u) C(u) \cos\left[\frac{(2x+1)u\pi}{2N}\right]$$

Two dimensional signals (images)

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

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

Properties of the DCT transform

- The cosine transform is real and orthogonal.
- The cosine transform is not a real part of the unitary DFT.
- The cosine transform of a sequence is related to the DFT of its antisymmetric extension
- The cosine transform is a fast transform
- The basis vectors of the cosine transform are the eigen vectors of the symmetric tridiagonal of Toeplitz matrix

- The cosine transform is close to the KL transform of first order Markov sequences. The cosine transform has very good to excellent energy compaction property of images,

The DCT is a real transform. This property makes it attractive in comparison to the Fourier transform.

The DCT has excellent energy compaction properties. For that reason it is widely used in image compression standards (as for example JPEG standards).

There are fast algorithms to compute the DCT, similar to the FFT for computing the DFT.

Sine Transform

The $N \times N$ sine transform matrix $\Psi = \{\psi(k, n)\}$, also called the *discrete sine transform* (DST), is defined as

$$\psi(k, n) = \sqrt{\frac{2}{N+1}} \sin \frac{\pi(k+1)(n+1)}{N+1}, \qquad 0 \le k, n \le N-1$$

The sine transform pair of one-dimensional sequences is defined as

$$v(k) = \sqrt{\frac{2}{N+1}} \sum_{n=0}^{N-1} u(n) \sin \frac{\pi(k+1)(n+1)}{N+1}, \qquad 0 \le k \le N-1$$

$$u(n) = \sqrt{\frac{2}{N+1}} \sum_{k=0}^{N-1} v(k) \sin \frac{\pi(k+1)(n+1)}{N+1}, \qquad 0 \le n \le N-1$$

Properties of the Discrete Sine Transform

- The sine transform is real, symmetric and orthogonal.
- The sine transform is not the imaginary part of the unitary DFT.
- The sine transform of a sequence is related to the DFT of its antisymmetric extension
- The sine transform is a fast transform
- The basis vectors of the sine transform are the eigen vectors of the symmetric tridiagonal of Toeplitz matrix
- The sine transform is close to the KL transform of first order Markov sequences. The sine transform has very good to excellent energy compaction property of images,

Hadamard Transform

In a similar form as the Walsh transform, the 2-D Hadamard transform is defined as follows.

Forward

$$H(u,v)=\frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}f(x,y)\left[\prod_{i=0}^{n-1}(-1)^{(b_i(x)b_i(u)+b_i(y)b_i(v))}\right],\ N=2^n \text{ or}$$

$$H(u,v)=\frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}f(x,y)(-1)^{\sum_{i=0}^{n-1}(b_i(x)b_i(u)+b_i(y)b_i(v))}$$

Inverse

$$f(x,y)=\frac{1}{N}\sum_{u=0}^{N-1}\sum_{v=0}^{N-1}H(u,v)\left[\prod_{i=0}^{n-1}(-1)^{(b_i(x)b_i(u)+b_i(y)b_i(v))}\right] \text{ etc.}$$

Properties of the Hadamard Transform

- The Hadamard transform is real, symmetric and orthogonal.
- The Hadamard transform is a fast transform
- The Hadamard transform has very good to excellent energy compaction property of images,

Properties are almost similar to that of *Walsh transform***.**

The Hadamard transform differs from the Walsh transform only in the order of basis functions. The order of basis functions of the Hadamard transform **does not** allow the fast computation of it by using a straightforward modification of the FFT. An extended version of the Hadamard transform is the **Ordered Hadamard Transform** for which a fast algorithm called **Fast Hadamard Transform (FHT)** can be applied.

Walsh Transform

The Walsh transform is defined as follows for two dimensional signals.

$$W(u,v)=\frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}f(x,y)\left[\prod_{i=0}^{n-1}(-1)^{(b_i(x)b_{n-1-i}(u)+b_i(y)b_{n-1-i}(v))}\right] \text{ or}$$

$$W(u,v)=\frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}f(x,y)(-1)^{\sum_{i=0}^{n-1}(b_i(x)b_{n-1-i}(u)+b_i(y)b_{n-1-i}(v))}$$

The inverse Walsh transform is defined as follows for two dimensional signals.

$$f(x,y)=\frac{1}{N}\sum_{u=0}^{N-1}\sum_{v=0}^{N-1}W(u,v)\left[\prod_{i=0}^{n-1}(-1)^{(b_i(x)b_{n-1-i}(u)+b_i(y)b_{n-1-i}(v))}\right] \text{ or}$$

$$f(x,y)=\frac{1}{N}\sum_{u=0}^{N-1}\sum_{v=0}^{N-1}W(u,v)(-1)^{\sum_{i=0}^{n-1}(b_i(x)b_{n-1-i}(u)+b_i(y)b_{n-1-i}(v))}$$

Properties of the Walsh Transform

- The Walsh transform is real, symmetric and orthogonal.
- The Walsh transform is a fast transform
- The Walsh transform has very good to excellent energy compaction property of images,
- Unlike the Fourier transform, which is based on trigonometric terms, the Walsh transform consists of a series expansion of basis functions whose values are only or and they have the form of square waves. These functions can be implemented more efficiently in a digital environment than the exponential basis functions of the Fourier transform.
- The forward and inverse Walsh kernels are identical except for a constant multiplicative factor of for 1-D signals.
- The forward and inverse Walsh kernels are identical for 2-D signals. This is because the array formed by the kernels is a symmetric matrix having orthogonal rows and columns, so its inverse array is the same as the array itself.

- The concept of frequency exists also in Walsh transform basis functions. We can think of frequency as the number of zero crossings or the number of transitions in a basis vector and we call this number **sequency**. The Walsh transform exhibits the property of energy compaction as all the transforms that we are currently studying
- For the fast computation of the Walsh transform there exists an algorithm called **Fast Walsh Transform (FWT).** This is a straightforward modification of the FFT.

Karhunen-Loeve Transform or KLT

The Karhunen-Loeve Transform or KLT was originally introduced as a series expansion for continuous random processes by Karhunen and Loeve. For discrete signals Hotelling first studied what was called a method of principal components, which is the discrete equivalent of the KL series expansion. Consequently, the KL transform is also called the Hotelling transform or the method of principal components.

Let e_i and □i, i □1,2,□,n , be this set of eigenvectors and corresponding eigenvalues of C_x , arranged in descending order so that $\lambda_i \geq \lambda_{i+1}$ for $i = 1,2,\ldots,n-1$. Let $\underline{A}$ be a matrix whose rows are formed from the eigenvectors of $\underline{C}_x$, ordered so that the first row of $\underline{A}$ is the eigenvector corresponding to the largest eigenvalue, and the last row the eigenvector corresponding to the smallest eigenvalue.

Suppose that $\underline{A}$ is a transformation matrix that maps the vectors $\underline{x}'s$ into vectors $\underline{y}'s$ by using the following transformation

$$\underline{y} = \underline{A}(\underline{x} - \underline{m}_x)$$

The above transform is called the **Karhunen-Loeve** or **Hotelling** transform. The mean of the $\underline{y}$ vectors resulting from the above transformation is zero **(try to prove that)**

$$\underline{m}_y = \underline{0}$$

the covariance matrix is **(try to prove that)**

$$\underline{C}_y = \underline{A}\underline{C}_x\underline{A}^T$$

and $\underline{C}_y$ is a diagonal matrix whose elements along the main diagonal are the eigenvalues of $\underline{C}_x$ **(try to prove that)**

$$\underline{C}_y = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

The off-diagonal elements of the covariance matrix are 0, so the elements of the $\underline{y}$ vectors are uncorrelated.

Lets try to reconstruct any of the original vectors $\underline{x}$ from its corresponding $\underline{y}$. Because the rows of $\underline{A}$ are orthonormal vectors (why?), then $\underline{A}^{-1} = \underline{A}^T$, and any vector $\underline{x}$ can by recovered from its corresponding vector $\underline{y}$ by using the relation

Properties of the Karhunen-Loeve transform

Despite its favourable theoretical properties, the KLT is not used in practice for the following reasons.

- Its basis functions depend on the covariance matrix of the image, and hence they have to recomputed and transmitted for every image.
- Perfect decorrelation is not possible, since images can rarely be modelled as realisations of ergodic fields.
- There are no fast computational algorithms for its implementation.

Harr Transform

The harr function $h_k(x)$ are defined on a continuous interval, x ϵ [0,1], and for k= 0 to N-1, where $N=2^n$, The integer k can be uniquely decomposed as

$$K = 2^p + q - 1$$

Where $0 \le p \le n-1$

$q=0,1$ for $p=0$ and $1 \le q \le 2^p$.

For example, N=4 we have

k	0	1	2	3
p	0	0	1	1
q	0	1	1	2

The harr function can be defined as

$$h_0(x) \triangleq h_{0,0}(x) = \frac{1}{\sqrt{N}}, \qquad x \in [0,1].$$

$$h_k(x) \triangleq h_{p,q}(x) = \frac{1}{\sqrt{N}} \begin{cases} 2^{p/2}, & \frac{q-1}{2^p} \le x < \frac{q-\frac{1}{2}}{2^p} \\ -2^{p/2}, & \frac{q-\frac{1}{2}}{2^p} \le x < \frac{q}{2^p} \\ 0, & \text{otherwise for } x \in [0,1] \end{cases}$$

For N=8.The harr transform is given by

$$\mathbf{Hr} = \frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} \\ 2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 \end{bmatrix} \quad \begin{matrix} \text{Sequency} \\ 0 \\ 1 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \end{matrix}$$

Properties of Harr transform

- The Harr transform is real and orthogonal.
- The Harr transform is a fast transform
- The basis vectors of the Harr transform are sequency ordered
- The Harr transform is close to the KL transform of first order Markov sequences. The Harr transform has poor energy compaction property for images

Slant transform

The NxN slant transform matrices are defined by the recursion

$$
\mathbf{S}_n = \frac{1}{\sqrt{2}}
\left[\begin{array}{cc|c|cc|c}
1 & 0 & & 1 & 0 & \\
 & & \mathbf{0} & & & \mathbf{0} \\
a_n & b_n & & -a_n & b_n & \\
\hline
\multicolumn{2}{c|}{\mathbf{0}} & \mathbf{I}_{(N/2)-2} & \multicolumn{2}{c|}{\mathbf{0}} & I_{(N/2)-2} \\
\hline
0 & 1 & & 0 & -1 & \\
 & & \mathbf{0} & & & \mathbf{0} \\
-b_n & a_n & & b_n & a_n & \\
\hline
\multicolumn{2}{c|}{\mathbf{0}} & I_{(N/2)-2} & \multicolumn{2}{c|}{\mathbf{0}} & -\mathbf{I}_{(N/2)-2}
\end{array}\right]
\left[\begin{array}{c|c}
\mathbf{S}_{n-1} & \mathbf{0} \\
\hline
\mathbf{0} & \mathbf{S}_{n-1}
\end{array}\right]
$$

The 4x4 slant transformation matrix is obtained as

$$
\mathbf{S}_2 = \frac{1}{2}
\begin{bmatrix}
1 & 1 & 1 & 1 \\
\frac{3}{\sqrt{5}} & \frac{1}{\sqrt{5}} & \frac{-1}{\sqrt{5}} & \frac{-3}{\sqrt{5}} \\
1 & -1 & -1 & 1 \\
\frac{1}{\sqrt{5}} & \frac{-3}{\sqrt{5}} & \frac{3}{\sqrt{5}} & \frac{-1}{\sqrt{5}}
\end{bmatrix}
\qquad
\begin{array}{c}
\text{Sequency} \\
0 \\
1 \\
2 \\
3
\end{array}
$$

Properties of slant transform

- The slant transform is real and orthogonal.
- The slant transform is a fast transform
- The basis vectors of the slant transform are not sequency ordered for n≥3.
- The slant transform is close to the KL transform of first order Markov sequences. The slant transform is very good to excellent energy compaction property for images

Questions for Practice:-

TWO MARKS:

1. What is meant by image transform?
2. What is the Need for transform?
3. Write note on unitary transform
4. List out the properties of Unitary transform
5. List out the applications of DCT
6. Mention the properties of Slant Transform
7. Write the transform pair of DST
8. What is hadamard transform? Mention it's applications
9. Write down 2D DFT
10. Write notes on SVD.

12 MARKS:

1. Explain the various properties of 2D DFT.
2. Explain about DST and DCT.
3. Briefly explain Hadamard and Slant Transform
4. Explain (a) Walsh (b) Harr transform