5

Image Transforms

5.1 INTRODUCTION

The term *image transforms* usually refers to a class of unitary matrices used for representing images. Just as a one-dimensional signal can be represented by an orthogonal series of *basis functions*, an image can also be expanded in terms of a discrete set of basis arrays called *basis images*. These basis images can be generated by unitary matrices. Alternatively, a given $N \times N$ image can be viewed as an $N^2 \times 1$ vector. An image transform provides a set of coordinates or basis vectors for the vector space.

For continuous functions, orthogonal series expansions provide series coefficients which can be used for any further processing or analysis of the functions. For a one-dimensional sequence $\{u(n), 0 \le n \le N-1\}$, represented as a vector $\mathbf{u}$ of size N, a unitary transformation is written as

$$\mathbf{v} = \mathbf{A}\mathbf{u} \quad \Rightarrow v(k) = \sum_{n=0}^{N-1} a(k,n)u(n), \qquad 0 \le k \le N-1 \tag{5.1}$$

where $\mathbf{A}^{-1} = \mathbf{A}^{*T}$ (*unitary*). This gives

$$\mathbf{u} = \mathbf{A}^{*T}\mathbf{v} \quad \Rightarrow u(n) = \sum_{k=0}^{N-1} v(k)a^*(k,n) \qquad 0 \le n \le N-1 \tag{5.2}$$

Equation (5.2) can be viewed as a series representation of the sequence $u(n)$. The columns of $\mathbf{A}^{*T}$, that is, the vectors $\mathbf{a}_k^* \triangleq \{a^*(k,n), 0 \le n \le N-1\}^T$ are called the *basis vectors* of $\mathbf{A}$. Figure 5.1 shows examples of basis vectors of several orthogonal transforms encountered in image processing. The series coefficients $v(k)$ give a representation of the original sequence $u(n)$ and are useful in filtering, data compression, feature extraction, and other analyses.

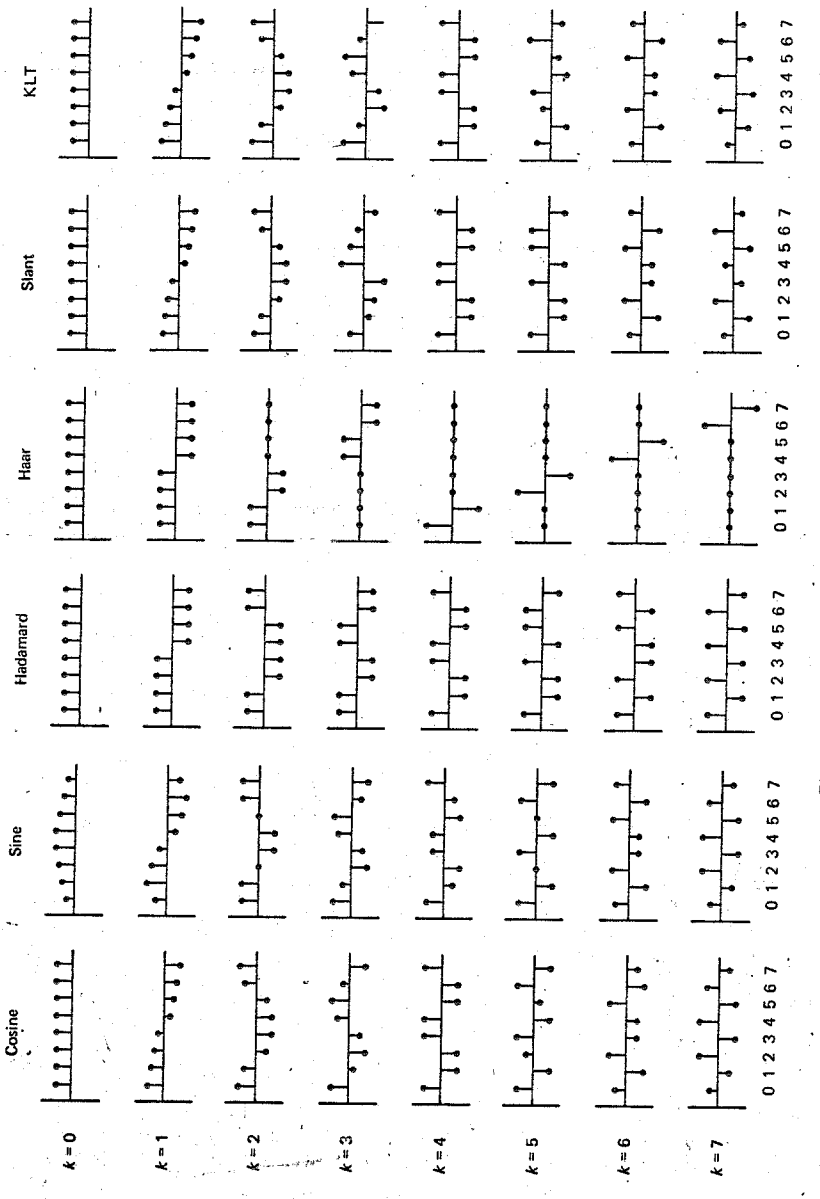


Figure 5.1 Basic vectors of the 8×8 transforms.

5.2 TWO-DIMENSIONAL ORTHOGONAL AND UNITARY TRANSFORMS

In the context of image processing a general orthogonal series expansion for an $N \times N$ image $u(m, n)$ is a pair of transformations of the form

$$v(k, l) = \sum_{m,n=0}^{N-1} \sum u(m, n)\alpha_{k,l}(m, n), \qquad 0 \le k, l \le N-1 \tag{5.3}$$

$$u(m, n) = \sum_{k,l=0}^{N-1} \sum v(k, l)\alpha^*_{k,l}(m, n), \qquad 0 \le m, n \le N-1 \tag{5.4}$$

where $\{\alpha_{k,l}(m, n)\}$, called an *image transform,* is a set of complete orthonormal discrete basis functions satisfying the properties

$$\textit{Orthonormality:} \quad \sum_{m,n=0}^{N-1} \sum \alpha_{k,l}(m, n)\alpha^*_{k',l'}(m, n) = \delta(k-k', l-l') \tag{5.5}$$

$$\textit{Completeness:} \quad \sum_{k,l=0}^{N-1} \sum \alpha_{k,l}(m, n)\alpha^*_{k,l}(m', n') = \delta(m-m', n-n') \tag{5.6}$$

The elements $v(k, l)$ are called the *transform coefficients* and $\mathbf{V} \triangleq \{v(k, l)\}$ is called the *transformed image.* The orthonormality property assures that any truncated series expansion of the form

$$u_{P,Q}(m, n) \triangleq \sum_{k=0}^{P-1} \sum_{l=0}^{Q-1} v(k, l)\alpha^*_{k,l}(m, n), \qquad P \le N, \quad Q \le N \tag{5.7}$$

will minimize the sum of squares error

$$\sigma_e^2 = \sum_{m,n=0}^{N-1} \sum [u(m, n) - u_{P,Q}(m, n)]^2 \tag{5.8}$$

when the coefficients $v(k, l)$ are given by (5.3). The completeness property assures that this error will be zero for $P = Q = N$ (Problem 5.1).

Separable Unitary Transforms

The number of multiplications and additions required to compute the transform coefficients $v(k, l)$ using (5.3) is $O(N^4)$, which is quite excessive for practical-size images. The dimensionality of the problem is reduced to $O(N^3)$ when the transform is restricted to be separable, that is,

$$\alpha_{k,l}(m, n) = a_k(m)b_l(n) \triangleq a(k, m)b(l, n) \tag{5.9}$$

where $\{a_k(m), k = 0, \ldots, N-1\}$, $\{b_l(n), l = 0, \ldots, N-1\}$ are one-dimensional complete orthonormal sets of basis vectors. Imposition of (5.5) and (5.6) shows that $\mathbf{A} \triangleq \{a(k, m)\}$ and $\mathbf{B} \triangleq \{b(l, n)\}$ should be unitary matrices themselves, for example,

$$\mathbf{A}\mathbf{A}^{*T} = \mathbf{A}^T\mathbf{A}^* = \mathbf{I} \tag{5.10}$$

Often one chooses $\mathbf{B}$ to be the same as $\mathbf{A}$ so that (5.3) and (5.4) reduce to

$$v(k, l) = \sum_{m,n=0}^{N-1} \sum a(k, m)u(m, n)a(l, n) \leftrightarrow \mathbf{V} = \mathbf{A}\mathbf{U}\mathbf{A}^T \tag{5.11}$$

$$u(m, n) = \sum_{k,l=0}^{N-1} \sum a^*(k, m)v(k, l)a^*(l, n) \leftrightarrow \mathbf{U} = \mathbf{A}^{*T}\mathbf{V}\mathbf{A}^* \tag{5.12}$$

For an $M \times N$ rectangular image, the transform pair is

$$\mathbf{V} = \mathbf{A}_M\mathbf{U}\mathbf{A}_N \tag{5.13}$$

$$\mathbf{U} = \mathbf{A}_M^{*T}\mathbf{V}\mathbf{A}_N^{*T} \tag{5.14}$$

where $\mathbf{A}_M$ and $\mathbf{A}_N$ are $M \times M$ and $N \times N$ unitary matrices, respectively. These are called two-dimensional separable transformations. Unless otherwise stated, we will always imply the preceding separability when we mention two-dimensional unitary transformations. Note that (5.11) can be written as

$$\mathbf{V}^T = \mathbf{A}[\mathbf{A}\mathbf{U}]^T \tag{5.15}$$

which means (5.11) can be performed by first transforming each column of $\mathbf{U}$ and then transforming each row of the result to obtain the rows of $\mathbf{V}$.

Basis Images

Let $\mathbf{a}_k^*$ denote the kth column of $\mathbf{A}^{*T}$. Define the matrices

$$\mathbf{A}^*_{k,l} = \mathbf{a}_k^*\,\mathbf{a}_l^{*T} \tag{5.16}$$

and the matrix *inner product* of two $N \times N$ matrices $\mathbf{F}$ and $\mathbf{G}$ as

$$\langle \mathbf{F}, \mathbf{G} \rangle = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} f(m, n)g^*(m, n) \tag{5.17}$$

Then (5.4) and (5.3) give a *series representation for the image* as

$$\mathbf{U} = \sum_{k,l=0}^{N-1} \sum v(k, l)\mathbf{A}^*_{k,l} \tag{5.18}$$

$$v(k, l) = \langle \mathbf{U}, \mathbf{A}^*_{k,l} \rangle \tag{5.19}$$

Equation (5.18) expresses any image $\mathbf{U}$ as a linear combination of the N^2 matrices $\mathbf{A}^*_{k,l}$, $k, l = 0, \ldots, N-1$, which are called the *basis images.* Figure 5.2 shows 8×8 basis images for the same set of transforms in Fig. 5.1. The transform coefficient $v(k, l)$ is simply the inner product of the (k, l)th basis image with the given image. It is also called the projection of the image on the (k, l)th basis image. Therefore, any $N \times N$ image can be expanded in a series using a complete set of N^2 basis images. If $\mathbf{U}$ and $\mathbf{V}$ are mapped into vectors by row ordering, then (5.11), (5.12), and (5.16) yield (see Section 2.8, on Kronecker products)

$$\begin{cases} \boldsymbol{v} = (\mathbf{A} \otimes \mathbf{A})\boldsymbol{u} \triangleq \mathcal{A}\boldsymbol{u} & (5.20) \\ \boldsymbol{u} = (\mathbf{A} \otimes \mathbf{A})^{*T}\boldsymbol{v} = \mathcal{A}^{*T}\boldsymbol{v} & (5.21) \end{cases}$$

where

$$\mathcal{A} \triangleq \mathbf{A} \otimes \mathbf{A} \tag{5.22}$$

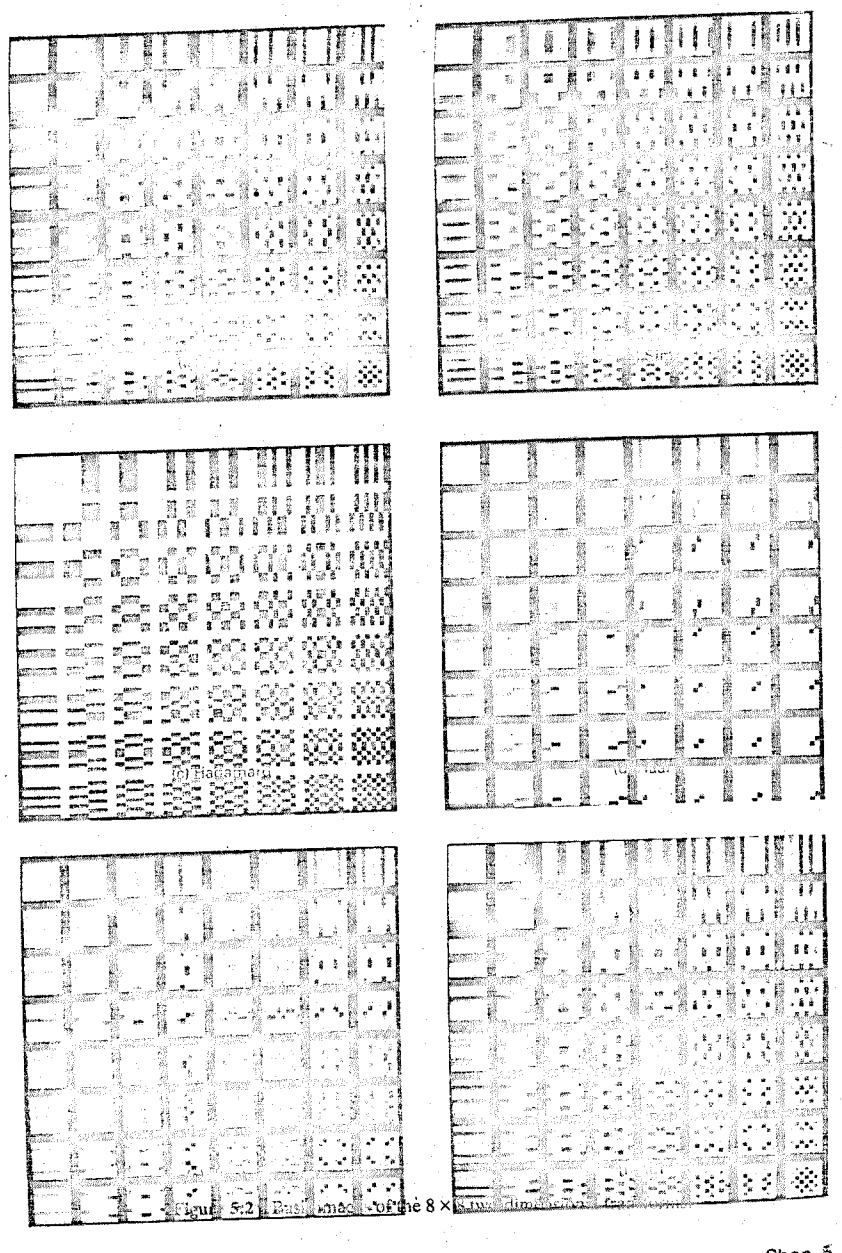


Figure 5.2 Basis images of the 8 × 8 two-dimensional fast

is a unitary matrix. Thus, given any unitary transform $\mathbf{A}$, a two-dimensional separable unitary transformation can be defined via (5.20) or (5.13).

Example 5.1

For the given orthogonal matrix $\mathbf{A}$ and image $\mathbf{U}$

$$\mathbf{A} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \qquad \mathbf{U} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

the transformed image, obtained according to (5.11), is

$$\mathbf{V} = \tfrac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \tfrac{1}{2}\begin{pmatrix} 4 & 6 \\ -2 & -2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 5 & -1 \\ -2 & 0 \end{pmatrix}$$

To obtain the basis images, we find the *outer product of the columns* of $\mathbf{A}^{*T}$, which gives

$$\mathbf{A}^*_{0,0} = \tfrac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}(1 \quad 1) = \tfrac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

and similarly

$$\mathbf{A}^*_{0,1} = \tfrac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} = \mathbf{A}^{*T}_{1,0}, \qquad \mathbf{A}^*_{1,1} = \tfrac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

The inverse transformation gives

$$\mathbf{A}^{*T}\mathbf{V}\mathbf{A}^* = \tfrac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 5 & -1 \\ -2 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \tfrac{1}{2}\begin{pmatrix} 3 & -1 \\ 7 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

which is $\mathbf{U}$, the original image.

Kronecker Products and Dimensionality

Dimensionality of image transforms can also be studied in terms of their Kronecker product separability. An arbitrary one-dimensional transformation

$$\mathcal{y} = \mathcal{A}\mathcal{x} \tag{5.23}$$

is called separable if

$$\mathcal{A} = \mathbf{A}_1 \otimes \mathbf{A}_2 \tag{5.24}$$

This is because (5.23) can be reduced to the separable two-dimensional transformation

$$\mathbf{Y} = \mathbf{A}_1 \mathbf{X} \mathbf{A}_2^T \tag{5.25}$$

where $\mathbf{X}$ and $\mathbf{Y}$ are matrices that map into vectors $\mathcal{x}$ and $\mathcal{y}$, respectively, by row ordering. If $\mathcal{A}$ is $N^2 \times N^2$ and $\mathbf{A}_1, \mathbf{A}_2$ are $N \times N$, then the number of operations required for implementing (5.25) reduces from N^4 to about $2N^3$. The number of operations can be reduced further if $\mathbf{A}_1$ and $\mathbf{A}_2$ are also separable. Image transforms such as discrete Fourier, sine, cosine, Hadamard, Haar, and Slant can be factored as Kronecker products of several smaller-sized matrices, which leads to fast algorithms for their implementation (see Problem 5.2). In the context of image processing such matrices are also called *fast image transforms.*

Dimensionality of Image Transforms

The $2N^3$ computations for $\mathbf{V}$ can also be reduced by restricting the choice of $\mathbf{A}$ to the *fast transforms*, whose matrix structure allows a factorization of the type

$$\mathbf{A} = \mathbf{A}_{(1)}\mathbf{A}_{(2)}, \ldots, \mathbf{A}_{(p)} \tag{5.26}$$

where $\mathbf{A}_{(i)}, i = 1, \ldots, p\ (p \ll N)$ are matrices with just a few nonzero entries (say r, with $r \ll N$). Thus, a multiplication of the type $\mathbf{y} = \mathbf{Ax}$ is accomplished in rpN operations. For Fourier, sine, cosine, Hadamard, Slant, and several other transforms, $p \simeq \log_2 N$, and the operations reduce to the order of $N \log_2 N$ (or $N^2 \log_2 N$ for $N \times N$ images). Depending on the actual transform, one operation can be defined as one multiplication and one addition or subtraction, as in the Fourier transform, or one addition or subtraction, as in the Hadamard transform.

Transform Frequency

For a one-dimensional signal $f(x)$, frequency is defined by the Fourier domain variable ξ. It is related to the number of zero crossings of the real or imaginary part of the basis function $\exp\{j2\pi\xi x\}$. This concept can be generalized to arbitrary unitary transforms. Let the rows of a unitary matrix $\mathbf{A}$ be arranged so that the number of zero crossings increases with the row number. Then in the transformation

$$\mathbf{y} = \mathbf{Ax}$$

the elements $y(k)$ are ordered according to increasing *wave number* or *transform frequency*. In the sequel any reference to frequency will imply the transform frequency, that is, discrete Fourier frequency, cosine frequency, and so on. The term *spatial frequency* generally refers to the continuous Fourier transform frequency and is not the same as the discrete Fourier frequency. In the case of Hadamard transform, a term called *sequency* is also used. It should be noted that this concept of frequency is useful only on a relative basis for a particular transform. A low-frequency term of one transform could contain the high-frequency harmonics of another transform.

The Optimum Transform

Another important consideration in selecting a transform is its performance in filtering and data compression of images based on the mean square criterion. The Karhunen-Loeve transform (KLT) is known to be optimum with respect to this criterion and is discussed in Section 5.11.

5.3 PROPERTIES OF UNITARY TRANSFORMS

Energy Conservation and Rotation

In the unitary transformation,

$$\mathbf{v} = \mathbf{Au}$$
$$\|\mathbf{v}\|^2 = \|\mathbf{u}\|^2 \tag{5.27}$$

This is easily proven by noting that

$$\|\mathbf{v}\|^2 \triangleq \sum_{k=0}^{N-1} |v(k)|^2 = \mathbf{v}^{*T}\mathbf{v} = \mathbf{u}^{*T}\mathbf{A}^{*T}\mathbf{Au} = \mathbf{u}^{*T}\mathbf{u} = \sum_{n=0}^{N-1} |u(n)|^2 \triangleq \|\mathbf{u}\|^2$$

Thus a unitary transformation preserves the signal energy or, equivalently, the length of the vector $\mathbf{u}$ in the N-dimensional vector space. This means every unitary transformation is simply a rotation of the vector $\mathbf{u}$ in the N-dimensional vector space. Alternatively, a unitary transformation is a rotation of the basis coordinates and the components of $\mathbf{v}$ are the projections of $\mathbf{u}$ on the new basis (see Problem 5.4). Similarly, for the two-dimensional unitary transformations such as (5.3), (5.4), and (5.11) to (5.14), it can be proven that

$$\sum_{m,n=0}^{N-1} \sum |u(m,n)|^2 = \sum_{k,l=0}^{N-1} \sum |v(k,l)|^2 \tag{5.28}$$

Energy Compaction and Variances of Transform Coefficients

Most unitary transforms have a tendency to pack a large fraction of the average energy of the image into a relatively few components of the transform coefficients. Since the total energy is preserved, this means many of the transform coefficients will contain very little energy. If $\boldsymbol{\mu}_u$ and $\mathbf{R}_u$ denote the mean and covariance of a vector $\mathbf{u}$, then the corresponding quantities for the transformed vector $\mathbf{v}$ are given by

$$\boldsymbol{\mu}_v \triangleq E[\mathbf{v}] = E[\mathbf{Au}] = \mathbf{A}E[\mathbf{u}] = \mathbf{A}\boldsymbol{\mu}_u \tag{5.29}$$

$$\begin{aligned}\mathbf{R}_v &= E[(\mathbf{v} - \boldsymbol{\mu}_v)(\mathbf{v} - \boldsymbol{\mu}_v)^{*T}] \\ &= \mathbf{A}(E[(\mathbf{u} - \boldsymbol{\mu}_u)(\mathbf{u} - \boldsymbol{\mu}_u)^{*T}])\mathbf{A}^{*T} = \mathbf{AR}_u\mathbf{A}^{*T}\end{aligned} \tag{5.30}$$

The transform coefficient variances are given by the diagonal elements of $\mathbf{R}_v$, that is

$$\sigma_v^2(k) = [\mathbf{R}_v]_{k,k} = [\mathbf{AR}_u\mathbf{A}^{*T}]_{k,k} \tag{5.31}$$

Since A is unitary, it follows that

$$\sum_{k=0}^{N-1} |\mu_v(k)|^2 = \boldsymbol{\mu}_v^{*T}\boldsymbol{\mu}_v = \boldsymbol{\mu}_u^{*T}\mathbf{A}^{*T}\mathbf{A}\boldsymbol{\mu}_u = \sum_{n=0}^{N-1} |\mu_u(n)|^2 \tag{5.32}$$

$$\sum_{k=0}^{N-1} \sigma_v^2(k) = \mathrm{Tr}[\mathbf{AR}_u\mathbf{A}^{*T}] = \mathrm{Tr}[\mathbf{R}_u] = \sum_{n=0}^{N-1} \sigma_u^2(n) \tag{5.33}$$

$$\Rightarrow \sum_{k=0}^{N-1} E[|v(k)|^2] = \sum_{n=0}^{N-1} E[|u(n)|^2] \tag{5.34}$$

The average energy $E[|v(k)|^2]$ of the transform coefficients $v(k)$ tends to be unevenly distributed, although it may be evenly distributed for the input sequence $u(n)$. For a two-dimensional random field $u(m, n)$ whose mean is $\mu_u(m, n)$ and covariance is $r(m, n; m', n')$, its transform coefficients $v(k, l)$ satisfy the properties

$$\mu_v(k,l) = \sum_m \sum_n a(k,m)a(l,n)\mu_u(m,n) \tag{5.35}$$

$$\sigma_v^2(k,l) = E[|v(k,l) - \mu_v(k,l)|^2]$$
$$= \sum_m \sum_n \sum_{m'} \sum_{n'} a(k,m)a(l,n)r(m,n;m',n')a^*(k,m')a^*(l,n') \tag{5.36}$$

If the covariance of $u(m,n)$ is separable, that is

$$r(m,n;m',n') = r_1(m,m')r_2(n,n') \tag{5.37}$$

then the variances of the transform coefficients can be written as a separable product

$$\sigma_v^2(k,l) = \sigma_1^2(k)\sigma_2^2(l)$$
$$\triangleq [\mathbf{A}\mathbf{R}_1\mathbf{A}^{*T}]_{k,k}[\mathbf{A}\mathbf{R}_2\mathbf{A}^{*T}]_{l,l} \tag{5.38}$$

where

$$\mathbf{R}_1 = \{r_1(m,m')\} \quad \text{and} \quad \mathbf{R}_2 = \{r_2(n,n')\}$$

Decorrelation

When the input vector elements are highly correlated, the transform coefficients tend to be uncorrelated. This means the off-diagonal terms of the covariance matrix $\mathbf{R}_v$ tend to become small compared to the diagonal elements.

With respect to the preceding two properties, the KL transform is optimum, that is, it packs the maximum average energy in a given number of transform coefficients while completely decorrelating them. These properties are presented in greater detail in Section 5.11.

Other Properties

Unitary transforms have other interesting properties. For example, the determinant and the eigenvalues of a unitary matrix have unity magnitude. Also, the entropy of a random vector is preserved under a unitary transformation. Since entropy is a measure of average information, this means information is preserved under a unitary transformation.

Example 5.2 (Energy compaction and decorrelation)

A 2×1 zero mean vector $\mathbf{u}$ is unitarily transformed as

$$\mathbf{v} = \tfrac{1}{2}\begin{pmatrix} \sqrt{3} & 1 \\ -1 & \sqrt{3} \end{pmatrix}\mathbf{u}, \quad \text{where} \quad \mathbf{R}_u \triangleq \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad 0 < \rho < 1$$

The parameter ρ measures the correlation between $u(0)$ and $u(1)$. The covariance of $\mathbf{v}$ is obtained as

$$\mathbf{R}_v = \begin{pmatrix} 1 + \sqrt{3}\rho/2 & \rho/2 \\ \rho/2 & 1 - \sqrt{3}\rho/2 \end{pmatrix}$$

From the expression for $\mathbf{R}_u$, $\sigma_u^2(0) = \sigma_u^2(1) = 1$, that is, the total average energy of 2 is distributed equally between $u(0)$ and $u(1)$. However, $\sigma_v^2(0) = 1 + \sqrt{3}\rho/2$ and $\sigma_v^2(1) = 1 - \sqrt{3}\rho/2$. The total average energy is still 2, but the average energy in $v(0)$ is greater than in $v(1)$. If $\rho = 0.95$, then 91.1% of the total average energy has been packed in the first sample. The correlation between $v(0)$ and $v(1)$ is given by

$$\rho_v(0,1) \triangleq \frac{E[v(0)v(1)]}{\sigma_v(0)\sigma_v(1)} = \frac{\rho}{2(1 - \frac{3}{4}\rho^2)^{1/2}}$$

which is less in absolute value than $|\rho|$ for $|\rho| < 1$. For $\rho = 0.95$, we find $\rho_v(0,1) = 0.83$. Hence the correlation between the transform coefficients has been reduced. If the foregoing procedure is repeated for the 2×2 transform $\mathbf{A}$ of Example 5.1, then we find $\sigma_v^2(0) = 1 + \rho$, $\sigma_v^2(1) = 1 - \rho$, and $\rho_v(0,1) = 0$. For $\rho = 0.95$, now 97.5% of the energy is packed in $v(0)$. Moreover, $v(0)$ and $v(1)$ become uncorrelated.

5.4 THE ONE-DIMENSIONAL DISCRETE FOURIER TRANSFORM (DFT)

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

$$v(k) = \sum_{n=0}^{N-1} u(n)W_N^{kn}, \qquad k = 0, 1, \ldots, N-1 \tag{5.39}$$

where

$$W_N \triangleq \exp\left\{\frac{-j2\pi}{N}\right\} \tag{5.40}$$

The inverse transform is given by

$$u(n) = \frac{1}{N}\sum_{k=0}^{N-1} v(k)W_N^{-kn}, \qquad n = 0, 1, \ldots, N-1 \tag{5.41}$$

The pair of equations (5.39) and (5.41) are not scaled properly to be unitary transformations. In image processing it is more convenient to consider the *unitary DFT*, which is defined as

$$v(k) = \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1} u(n)W_N^{kn}, \qquad k = 0, \ldots, N-1 \tag{5.42}$$

$$u(n) = \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} v(k)W_N^{-kn}, \qquad n = 0, \ldots, N-1 \tag{5.43}$$

The $N \times N$ unitary DFT matrix $\mathbf{F}$ is given by

$$\mathbf{F} = \left\{\frac{1}{\sqrt{N}}W_N^{kn}\right\}, \qquad 0 \le k, n \le N-1 \tag{5.44}$$

Future references to DFT and unitary DFT will imply the definitions of (5.39) and (5.42), respectively. The DFT is one of the most important transforms in digital signal and image processing. It has several properties that make it attractive for image processing applications.

Properties of the DFT/Unitary DFT

Let $u(n)$ be an arbitrary sequence defined for $n = 0, 1, \ldots, N-1$. A circular shift of $u(n)$ by l, denoted by $u(n-l)_c$, is defined as $u[(n-l) \text{ modulo } N]$. See Fig. 5.3 for $l = 2, N = 5$.

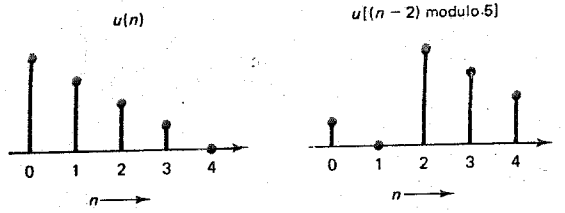


Figure 5.3 Circular shift of $u(n)$ by 2.

The DFT and Unitary DFT matrices are symmetric. By definition the matrix $\mathbf{F}$ is symmetric. Therefore,

$$\mathbf{F}^{-1} = \mathbf{F}^* \tag{5.45}$$

The extensions are periodic. The extensions of the DFT and unitary DFT of a sequence and their inverse transforms are periodic with period N. If for example, in the definition of (5.42) we let k take all integer values, then the sequence $v(k)$ turns out to be periodic, that is, $v(k) = v(k+N)$ for every k.

The DFT is the sampled spectrum of the finite sequence $u(n)$ extended by zeros outside the interval $[0, N-1]$. If we define a zero-extended sequence

$$\tilde{u}(n) \triangleq \begin{cases} u(n), & 0 \le n \le N-1 \\ 0, & \text{otherwise} \end{cases} \tag{5.46}$$

then its Fourier transform is

$$\tilde{U}(\omega) = \sum_{n=-\infty}^{\infty} \tilde{u}(n) \exp(-j\omega n) = \sum_{n=0}^{N-1} u(n) \exp(-j\omega n) \tag{5.47}$$

Comparing this with (5.39) we see that

$$v(k) = \tilde{U}\left(\frac{2\pi k}{N}\right) \tag{5.48}$$

Note that the unitary DFT of (5.42) would be $\tilde{U}(2\pi k/N)/\sqrt{N}$.

The DFT and unitary DFT of dimension N can be implemented by a fast algorithm in $O(N \log_2 N)$ operations. There exists a class of algorithms, called the *fast Fourier transform* (FFT), which requires $O(N \log_2 N)$ operations for implementing the DFT or unitary DFT, where one operation is a real multiplication and a real addition. The exact operation count depends on N as well as the particular choice of the algorithm in that class. Most common FFT algorithms require $N = 2^p$, where p is a positive integer.

The DFT or unitary DFT of a real sequence $\{x(n), n = 0, \ldots, N-1\}$ is conjugate symmetric about $N/2$. From (5.42) we obtain

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

$$\Rightarrow \quad v\left(\frac{N}{2}-k\right) = v^*\left(\frac{N}{2}+k\right), \qquad k = 0, \ldots, \frac{N}{2}-1 \tag{5.49}$$

$$\Rightarrow \quad \left|v\left(\frac{N}{2}-k\right)\right| = \left|v\left(\frac{N}{2}+k\right)\right| \tag{5.50}$$

Figure 5.4 shows a 256-sample scan line of an image. The magnitude of its DFT is shown in Fig. 5.5, which exhibits symmetry about the point 128. If we consider the periodic extension of $v(k)$, we see that

$$v(-k) = v(N-k)$$

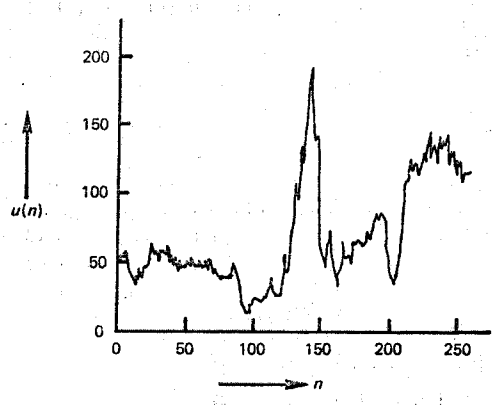


Figure 5.4 A 256-sample scan line of an image.

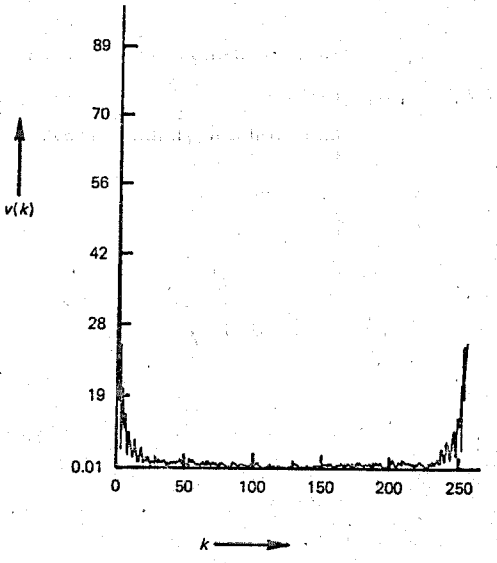


Figure 5.5 Unitary discrete Fourier transform of Fig. 5.4.

Hence the (unitary) DFT frequencies $N/2+k$, $k=0,\ldots,N/2-1$, are simply the negative frequencies at $\omega=(2\pi/N)(-N/2+k)$ in the Fourier spectrum of the finite sequence $\{u(n), 0\le n\le N-1\}$. Also, from (5.39) and (5.49), we see that $v(0)$ and $v(N/2)$ are real, so that the $N\times 1$ real sequence

$$v(0),\left\{\mathrm{Re}\{v(k)\},k=1,\ldots,\frac{N}{2}-1\right\},\left\{\mathrm{Im}\{v(k)\},k=1,\ldots,\frac{N}{2}-1\right\},v\left(\frac{N}{2}\right) \tag{5.51}$$

completely defines the DFT of the real sequence $u(n)$. Therefore, it can be said that the DFT or unitary DFT of an $N\times 1$ real sequence has N degrees of freedom and requires the same storage capacity as the sequence itself.

The basis vectors of the unitary DFT are the orthonormal eigenvectors of any circulant matrix. Moreover, the eigenvalues of a circulant matrix are given by the DFT of its first column. Let $\mathbf{H}$ be an $N\times N$ circulant matrix. Therefore, its elements satisfy

$$[\mathbf{H}]_{m,n}=h(m-n)=h[(m-n)\ \text{modulo}\ N],\qquad 0\le m,n\le N-1 \tag{5.52}$$

The basis vectors of the unitary DFT are columns of $\mathbf{F}^{*T}=\mathbf{F}^*$, that is,

$$\boldsymbol{\phi}_k=\left\{\frac{1}{\sqrt{N}}W_N^{-kn},0\le n\le N-1\right\}^T,\qquad k=0,\ldots,N-1 \tag{5.53}$$

Consider the expression

$$[\mathbf{H}\boldsymbol{\phi}_k]_m=\frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}h(m-n)W_N^{-kn} \tag{5.54}$$

Writing $m-n=l$ and rearranging terms, we can write

$$[\mathbf{H}\boldsymbol{\phi}_k]_m=\frac{1}{\sqrt{N}}W_N^{-km}\left[\sum_{l=0}^{N-1}h(l)W_N^{kl}+\sum_{l=-N+m+1}^{-1}h(l)W_N^{kl}-\sum_{l=m+1}^{N-1}h(l)W_N^{kl}\right] \tag{5.55}$$

Using (5.52) and the fact that $W_N^{-l}=W_N^{N-l}$ (since $W_N^N=1$), the second and third terms in the brackets cancel, giving the desired eigenvalue equation

$$[\mathbf{H}\boldsymbol{\phi}_k]_m=\lambda_k\boldsymbol{\phi}_k(m)$$

or

$$\mathbf{H}\boldsymbol{\phi}_k=\lambda_k\boldsymbol{\phi}_k \tag{5.56}$$

where λ_k, the eigenvalues of $\mathbf{H}$, are defined as

$$\lambda_k\triangleq\sum_{l=0}^{N-1}h(l)W_N^{kl},\qquad 0\le k\le N-1 \tag{5.57}$$

This is simply the DFT of the first column of $\mathbf{H}$.

Based on the preceding properties of the DFT, the following additional properties can be proven (Problem 5.9).

Circular convolution theorem. *The DFT of the circular convolution of two sequences is equal to the product of their DFTs,* that is, if

$$x_2(n)=\sum_{k=0}^{N-1}h(n-k)_c x_1(k),\qquad 0\le n\le N-1 \tag{5.58}$$

then

$$\mathrm{DFT}\{x_2(n)\}_N=\mathrm{DFT}\{h(n)\}_N\mathrm{DFT}\{x_1(n)\}_N \tag{5.59}$$

where $\mathrm{DFT}\{x(n)\}_N$ denotes the DFT of the sequence $x(n)$ of size N. This means we can calculate the circular convolution by first calculating the DFT of $x_2(n)$ via (5.59) and then taking its inverse DFT. Using the FFT this will take $O(N\log_2 N)$ operations, compared to N^2 operations required for direct evaluation of (5.58).

A linear convolution of two sequences can also be obtained via the FFT by imbedding it into a circular convolution. In general, the linear convolution of two sequences $\{h(n),n=0,\ldots,N'-1\}$ and $\{x_1(n),n=0,\ldots,N-1\}$ is a sequence $\{x_2(n),0\le n\le N'+N-2\}$ and can be obtained by the following algorithm:

Step 1: Let $M\ge N'+N-1$ be an integer for which an FFT algorithm is available.

Step 2: Define $\tilde{h}(n)$ and $\tilde{x}_1(n),0\le n\le M-1$, as zero extended sequences corresponding to $h(n)$ and $x_1(n)$, respectively.

Step 3: Let $\tilde{y}_1(k)=\mathrm{DFT}\{\tilde{x}_1(n)\}_M,\lambda_k=\mathrm{DFT}\{\tilde{h}(n)\}_M$. Define $\tilde{y}_2(k)=\lambda_k\tilde{y}_1(k)$, $k=0,\ldots,M-1$.

Step 4: Take the inverse DFT of $\tilde{y}_2(k)$ to obtain $\tilde{x}_2(n)$. Then $x_2(n)=\tilde{x}_2(n)$ for $0\le n\le N+N'-2$.

Any circulant matrix can be diagonalized by the DFT/unitary DFT. That is,

$$\mathbf{FHF}^*=\boldsymbol{\Lambda} \tag{5.60}$$

where $\boldsymbol{\Lambda}=\mathrm{Diag}\{\lambda_k,0\le k\le N-1\}$ and λ_k are given by (5.57). It follows that if $\mathbf{C}$, $\mathbf{C}_1$ and $\mathbf{C}_2$ are circulant matrices, then the following hold.

1. $\mathbf{C}_1\mathbf{C}_2=\mathbf{C}_2\mathbf{C}_1$, that is, circulant matrices commute.
2. $\mathbf{C}^{-1}$ is a circulant matrix and can be computed in $O(N\log N)$ operations.
3. $\mathbf{C}^T$, $\mathbf{C}_1+\mathbf{C}_2$, and $f(\mathbf{C})$ are all circulant matrices, where $f(x)$ is an arbitrary function of x.

5.5 THE TWO-DIMENSIONAL DFT

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

$$v(k,l)=\sum_{m=0}^{N-1}\sum_{n=0}^{N-1}u(m,n)W_N^{km}W_N^{ln},\qquad 0\le k,l\le N-1 \tag{5.61}$$

and the inverse transform is

$$u(m,n)=\frac{1}{N^2}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1}v(k,l)W_N^{-km}W_N^{-ln},\qquad 0\le m,n\le N-1 \tag{5.62}$$

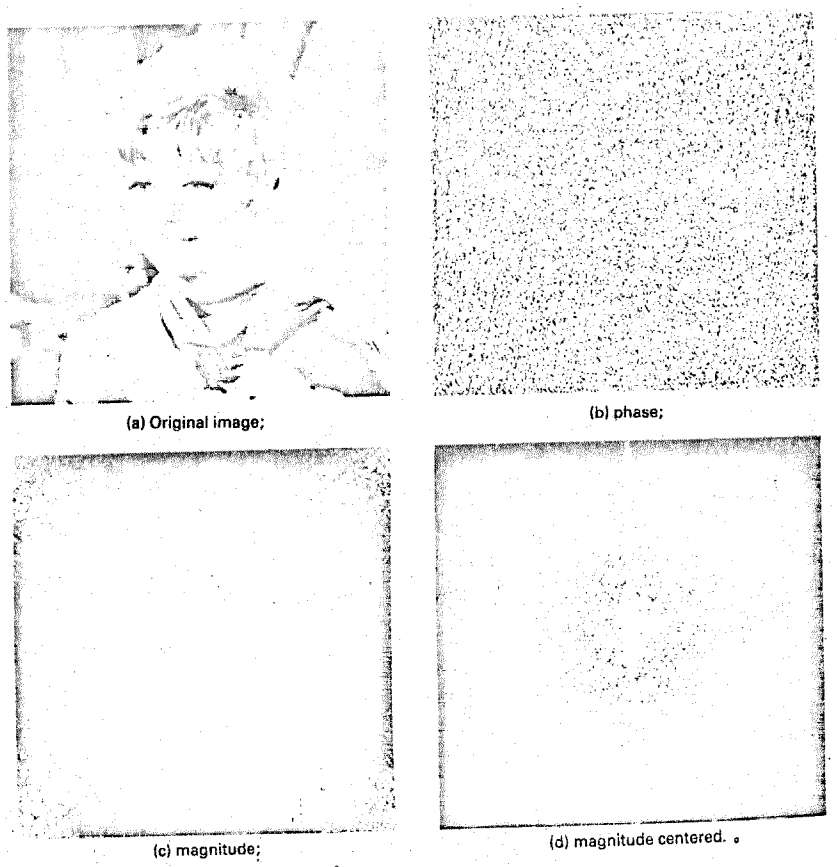
The two-dimensional unitary DFT pair is defined as

$$v(k,l)=\frac{1}{N}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1}u(m,n)W_N^{km}W_N^{ln},\qquad 0\le k,l\le N-1 \tag{5.63}$$

$$u(m,n)=\frac{1}{N}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1}v(k,l)W_N^{-km}W_N^{-ln},\qquad 0\le m,n\le N-1 \tag{5.64}$$

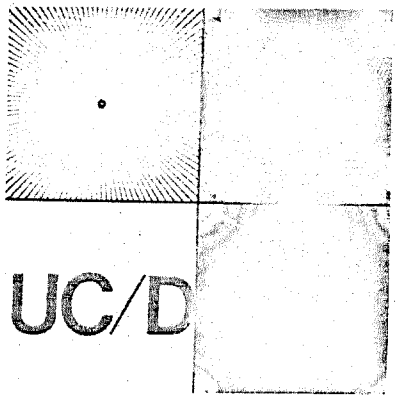
In matrix notation this becomes

$$\mathbf{V}=\mathbf{FUF} \tag{5.65}$$



(a) Original image; (b) phase; (c) magnitude; (d) magnitude centered.

Figure 5.6 Two-dimensional unitary DFT of a 256 × 256 image.



a	b
c	d

Figure 5.7 Unitary DFT of images

(a) Resolution chart;
(b) its DFT;
(c) binary image;
(d) its DFT. The two parallel lines are due to the '/' sign in the binary image.

$$\mathbf{U}=\mathbf{F}^*\mathbf{V}\mathbf{F}^* \tag{5.66}$$

If **U** and **V** are mapped into row-ordered vectors u and v, respectively, then

$$v=\mathcal{F}u,\qquad u=\mathcal{F}^*v \tag{5.67}$$

$$\mathcal{F}=\mathbf{F}\otimes\mathbf{F} \tag{5.68}$$

The $N^2\times N^2$ matrix $\mathcal{F}$ represents *the $N\times N$ two-dimensional unitary* DFT. Figure 5.6 shows an original image and the magnitude and phase components of its unitary DFT. Figure 5.7 shows magnitudes of the unitary DFTs of two other images.

Properties of the Two-Dimensional DFT

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

Symmetric, unitary.

$$\mathcal{F}^T=\mathcal{F},\qquad \mathcal{F}^{-1}=\mathcal{F}^*=\mathbf{F}^*\otimes\mathbf{I}^* \tag{5.69}$$

Periodic extensions.

$$\begin{aligned}v(k+N,l+N)&=v(k,l),\qquad \forall k,l\\ u(m+N,n+N)&=u(m,n),\qquad \forall m,n\end{aligned} \tag{5.70}$$

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

$$\tilde{U}\left(\frac{2\pi k}{N},\frac{2\pi l}{N}\right)=\mathrm{DFT}\{u(m,n)\}=v(k,l) \tag{5.71}$$

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 (5.65) 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,

$$v\left(\frac{N}{2} \pm k, \frac{N}{2} \pm l\right) = v^*\left(\frac{N}{2} \mp k, \frac{N}{2} \mp l\right), \qquad 0 \le k, l \le \frac{N}{2} - 1 \tag{5.72}$$

or

$$v(k, l) = v^*(N - k, N - l), \qquad 0 \le k, l \le N - 1 \tag{5.73}$$

From this, it can be shown that $v(k, l)$ has only N^2 independent real elements. For example, the samples in the shaded region of Fig. 5.8 determine the complete DFT or unitary DFT (see problem 5.10).

Basis images. The basis images are given by definition [see (5.16) and (5.53)]:

$$\mathbf{A}^*_{k,l} = \boldsymbol{\phi}_k \boldsymbol{\phi}_l^T = \frac{1}{N}\{W_N^{-(km+ln)}, \qquad 0 \le m, n \le N-1\}, \qquad 0 \le k, l \le N-1 \tag{5.74}$$

Two-dimensional circular convolution theorem. *The DFT of the two-dimensional circular convolution of two arrays is the product of their DFTs.*

Two-dimensional circular convolution of two $N \times N$ arrays $h(m, n)$ and $u_1(m, n)$ is defined as

$$u_2(m, n) = \sum_{m'=0}^{N-1} \sum_{n'=0}^{N-1} h(m - m', n - n')_c u_1(m', n'), \qquad 0 \le m, n \le N-1 \tag{5.75}$$

where

$$h(m, n)_c = h(m \text{ modulo } N, n \text{ modulo } N) \tag{5.76}$$

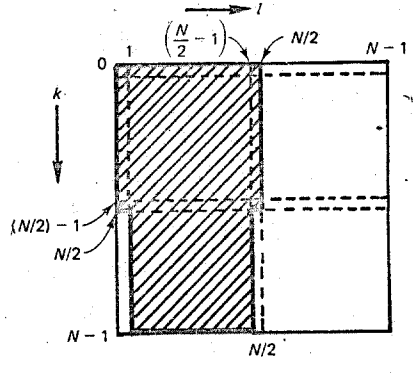
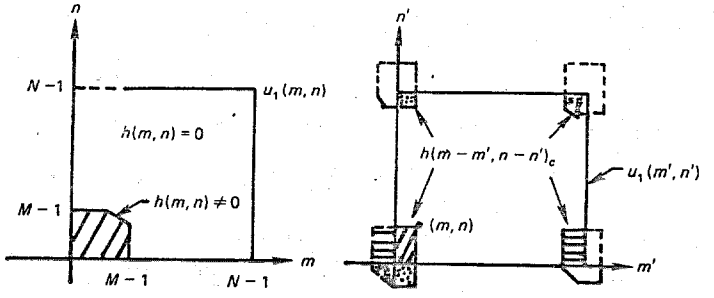


Figure 5.8 Discrete Fourier transform coefficients $v(k, l)$ in the shaded area determine the remaining coefficients.



(a) Array $h(m, n)$.

(b) Circular convolution of $h(m, n)$ with $u_1(m, n)$ over $N \times N$ region.

Figure 5.9 Two-dimensional circular convolution.

Figure 5.9 shows the meaning of circular convolution. It is the same when a periodic extension of $h(m, n)$ is convolved over an $N \times N$ region with $u_1(m, n)$. The two-dimensional DFT of $h(m - m', n - n')_c$ for fixed m', n' is given by

$$\begin{aligned}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} h(m - m', n - n')_c W_N^{(mk+nl)} &= W_N^{(m'k+n'l)} \sum_{i=-m'}^{N-1-m'} \sum_{j=-n'}^{N-1-n'} h(i, j)_c W_N^{(ik+jl)} \\ &= W_N^{(m'k+n'l)} \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} h(m, n) W_N^{(mk+nl)} \\ &= W_N^{(m'k+n'l)} \text{DFT}\{h(m, n)\}_N\end{aligned} \tag{5.77}$$

where we have used (5.76). Taking the DFT of both sides of (5.75) and using the preceding result, we obtain†

$$\text{DFT}\{u_2(m, n)\}_N = \text{DFT}\{h(m, n)\}_N \text{DFT}\{u_1(m, n)\}_N \tag{5.78}$$

From this and the fast transform property (page 142), it follows that an $N \times N$ circular convolution can be performed in $O(N^2 \log_2 N)$ operations. This property is also useful in calculating two-dimensional convolutions such as

$$x_3(m, n) = \sum_{m'=0}^{M-1} \sum_{n'=0}^{M-1} x_2(m - m', n - n') x_1(m', n') \tag{5.79}$$

where $x_1(m, n)$ and $x_2(m, n)$ are assumed to be zero for $m, n \notin [0, M-1]$. The region of support for the result $x_3(m, n)$ is $\{0 \le m, n \le 2M - 2\}$. Let $N \ge 2M - 1$ and define $N \times N$ arrays

$$\tilde{h}(m, n) \triangleq \begin{cases} x_2(m, n), & 0 \le m, n \le M-1 \\ 0, & \text{otherwise} \end{cases} \tag{5.80}$$

$$\tilde{u}_1(m, n) \triangleq \begin{cases} x_1(m, n), & 0 \le m, n \le M-1 \\ 0, & \text{otherwise} \end{cases} \tag{5.81}$$

†We denote $\text{DFT}\{x(m, n)\}_N$ as the two-dimensional DFT of an $N \times N$ array $x(m, n)$, $0 \le m, n \le N-1$.

Evaluating the circular convolution of $\tilde{h}(m, n)$ and $\tilde{u}_1(m, n)$ according to (5.75), it can be seen with the aid of Fig. 5.9 that

$$x_3(m, n) = u_2(m, n), \qquad 0 \le m, n \le 2M - 2 \tag{5.82}$$

This means the two-dimensional linear convolution of (5.79) can be performed in $O(N^2 \log_2 N)$ operations.

Block circulant operations. Dividing both sides of (5.77) by N and using the definition of Kronecker product, we obtain

$$(\mathbf{F} \otimes \mathbf{F})\mathscr{H} = \mathscr{D}(\mathbf{F} \otimes \mathbf{F}) \tag{5.83}$$

where $\mathscr{H}$ is doubly circulant and $\mathscr{D}$ is diagonal whose elements are given by

$$[\mathscr{D}]_{kN+l, kN+l} \triangleq d_{k,l} = \text{DFT}\{h(m, n)\}_N, \qquad 0 \le k, l \le N - 1 \tag{5.84}$$

Eqn. (5.83) can be written as

$$\mathscr{F}\mathscr{H} = \mathscr{D}\mathscr{F} \qquad \text{or} \qquad \mathscr{F}\mathscr{H}\mathscr{F}^* = \mathscr{D} \tag{5.85}$$

that is, a doubly block circulant matrix is diagonalized by the two-dimensional unitary DFT. From (5.84) and the fast transform property (page 142), we conclude that a doubly block circulant matrix can be diagonalized in $O(N^2 \log_2 N)$ operations. The eigenvalues of $\mathscr{H}$, given by the two-dimensional DFT of $h(m, n)$, are the same as operating $N\mathscr{F}$ on the first column of $\mathscr{H}$. This is because the elements of the first column of $\mathscr{H}$ are the elements $h(m, n)$ mapped by lexicographic ordering.

Block Toeplitz operations. Our discussion on linear convolution implies that any doubly block Toeplitz matrix operation can be imbedded into a double block circulant operation, which, in turn, can be implemented using the two-dimensional unitary DFT.

5.6 THE COSINE TRANSFORM

The $N \times N$ cosine transform matrix $\mathbf{C} = \{c(k, n)\}$, also called the *discrete cosine transform* (DCT), is defined as

$$c(k, n) = \begin{cases} \dfrac{1}{\sqrt{N}}, & k = 0,\ 0 \le n \le N - 1 \\ \sqrt{\dfrac{2}{N}} \cos \dfrac{\pi(2n + 1)k}{2N}, & 1 \le k \le N - 1,\ 0 \le n \le N - 1 \end{cases} \tag{5.86}$$

The one-dimensional DCT of a sequence $\{u(n), 0 \le n \le N - 1\}$ is defined as

$$v(k) = \alpha(k) \sum_{n=0}^{N-1} u(n) \cos\left[\frac{\pi(2n + 1)k}{2N}\right], \qquad 0 \le k \le N - 1 \tag{5.87}$$

where

$$\alpha(0) \triangleq \sqrt{\frac{1}{N}}, \qquad \alpha(k) \triangleq \sqrt{\frac{2}{N}} \quad \text{for} \quad 1 \le k \le N - 1 \tag{5.88}$$

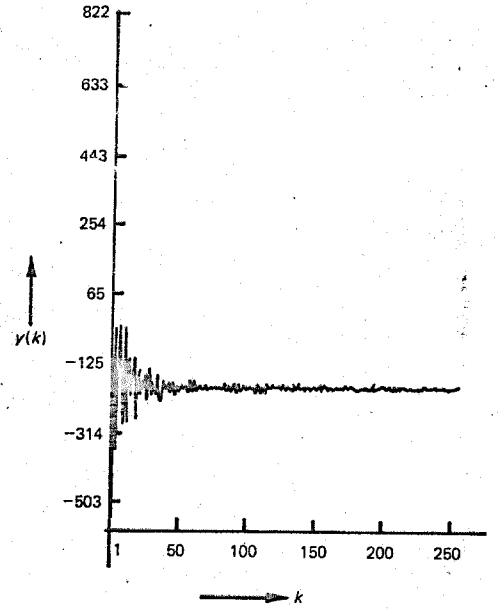


Figure 5.10 Cosine transform of the image scan line shown in Fig. 5.4.

The inverse transformation is given by

$$u(n) = \sum_{k=0}^{N-1} \alpha(k) v(k) \cos\left[\frac{\pi(2n + 1)k}{2N}\right], \qquad 0 \le n \le N - 1 \tag{5.89}$$

The basis vectors of the 8×8 DCT are shown in Fig. 5.1. Figure 5.10 shows the cosine transform of the image scan line shown in Fig. 5.4. Note that many transform coefficients are small, that is, most of the energy of the data is packed in a few transform coefficients.

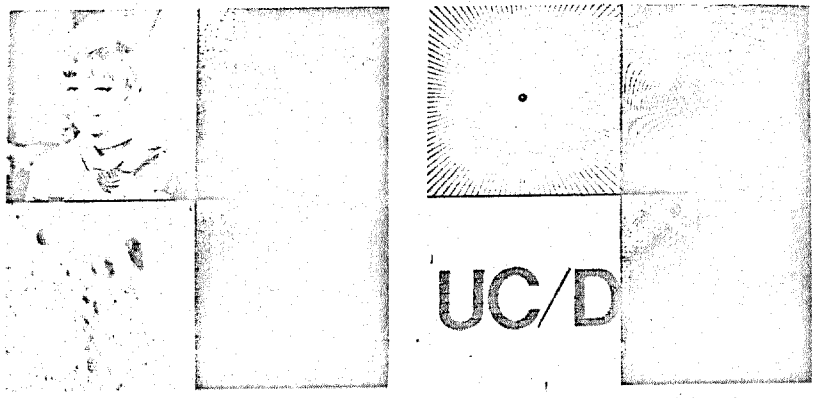
The two-dimensional cosine transform pair is obtained by substituting $\mathbf{A} = \mathbf{A}^* = \mathbf{C}$ in (5.11) and (5.12). The basis images of the 8×8 two-dimensional cosine transform are shown in Fig. 5.2. Figure 5.11 shows examples of the cosine transform of different images.

Properties of the Cosine Transform

1. The cosine transform is real and orthogonal, that is,

$$\mathbf{C} = \mathbf{C}^* \Rightarrow \mathbf{C}^{-1} = \mathbf{C}^T \tag{5.90}$$

2. The cosine transform is not the real part of the unitary DFT. This can be seen by inspection of $\mathbf{C}$ and the DFT matrix $\mathbf{F}$. (Also see Problem 5.13.) However, the cosine transform of a sequence is related to the DFT of its symmetric extension (see Problem 5.16).



(a) Cosine transform examples of monochrome images; (b) Cosine transform examples of binary images.

Figure 5.11

3. The cosine transform is a fast transform. The cosine transform of a vector of N elements can be calculated in $O(N \log_2 N)$ operations via an N-point FFT [19]. To show this we define a new sequence $\tilde{u}(n)$ by reordering the even and odd elements of $u(n)$ as

$$\left.\begin{aligned} \tilde{u}(n) &= u(2n) \\ \tilde{u}(N-n-1) &= u(2n+1) \end{aligned}\right\}, \qquad 0 \le n \le \left(\frac{N}{2}\right) - 1 \tag{5.91}$$

Now, we split the summation term in (5.87) into even and odd terms and use (5.91) to obtain

$$\begin{aligned} v(k) &= \alpha(k)\left\{\sum_{n=0}^{(N/2)-1} u(2n)\cos\left[\frac{\pi(4n+1)k}{2N}\right] + \sum_{n=0}^{(N/2)-1} u(2n+1)\cos\left[\frac{\pi(4n+3)k}{2N}\right]\right\} \\ &= \alpha(k)\left\{\left[\sum_{n=0}^{(N/2)-1} \tilde{u}(n)\cos\left[\frac{\pi(4n+1)k}{2N}\right] + \sum_{n=0}^{(N/2)-1} \tilde{u}(N-n-1)\cos\left[\frac{\pi(4n+3)k}{2N}\right]\right\}\right. \end{aligned}$$

Changing the index of summation in the second term to $n' = N - n - 1$ and combining terms, we obtain

$$v(k) = \alpha(k)\sum_{n=0}^{N-1} \tilde{u}(n)\cos\left[\frac{\pi(4n+1)k}{2N}\right] \tag{5.92}$$

$$= \mathrm{Re}\left[\alpha(k)e^{-j\pi k/2N}\sum_{n=0}^{N-1}\tilde{u}(n)e^{-j2\pi kn/N}\right] = \mathrm{Re}[\alpha(k)W_{2N}^{k/2}\,\mathrm{DFT}\{\tilde{u}(n)\}_N]$$

which proves the previously stated result. For inverse cosine transform we write (5.89) for even data points as

$$u(2n) = \hat{u}(2n) \triangleq \mathrm{Re}\left[\sum_{k=0}^{N-1}[\alpha(k)v(k)e^{j\pi k/2N}]e^{j2\pi nk/N}\right], \qquad 0 \le n \le \left(\frac{N}{2}\right) - 1 \tag{5.93}$$

The odd data points are obtained by noting that

$$u(2n+1) = \hat{u}[2(N-1-n)], \qquad 0 \le n \le \left(\frac{N}{2}\right) - 1 \tag{5.94}$$

Therefore, if we calculate the N-point inverse FFT of the sequence $\alpha(k)v(k)\exp(j\pi k/2N)$, we can also obtain the inverse DCT in $O(N \log N)$ operations. Direct algorithms that do not require FFT as an intermediate step, so that complex arithmetic is avoided, are also possible [18]. The computational complexity of the direct as well as the FFT based methods is about the same.

4. The cosine transform has excellent energy compaction for highly correlated data. This is due to the following properties.
5. The basis vectors of the cosine transform (that is, rows of $\mathbf{C}$) are the eigenvectors of the symmetric tridiagonal matrix $\mathbf{Q}_c$, defined as

$$\mathbf{Q}_c = \begin{bmatrix} 1-\alpha & -\alpha & 0 & & \\ -\alpha & 1 & & & \\ & & \ddots & & \\ & & & 1 & -\alpha \\ 0 & & & -\alpha & 1-\alpha \end{bmatrix} \tag{5.95}$$

The proof is left as an exercise.

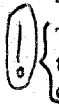
6. The $N \times N$ cosine transform is very close to the KL transform of a first-order stationary Markov sequence of length N whose covariance matrix is given by (2.68) when the correlation parameter ρ is close to 1. The reason is that $\mathbf{R}^{-1}$ is a symmetric tridiagonal matrix, which for a scalar $\beta^2 \triangleq (1-\rho^2)/(1+\rho^2)$, and $\alpha \triangleq \rho/(1+\rho^2)$ satisfies the relation

$$\beta^2\mathbf{R}^{-1} = \begin{bmatrix} 1-\rho\alpha & -\alpha & 0 & & \\ -\alpha & 1 & & & \\ & & \ddots & & \\ & & & 1 & -\alpha \\ 0 & & & -\alpha & 1-\rho\alpha \end{bmatrix} \tag{5.96}$$

This gives the approximation

$$\beta^2 R^{-1} \cong \mathbf{Q}_c \quad \text{for} \quad \rho \cong 1 \tag{5.97}$$

Hence the eigenvectors of $\mathbf{R}$ and the eigenvectors of $\mathbf{Q}_c$, that is, the cosine transform, will be quite close. These aspects are considered in greater depth in Section 5.12 on sinusoidal transforms.



This property of the cosine transform together with the fact that it is a fast transform has made it a useful substitute for the KL transform of highly correlated first-order Markov sequences.

5.7 THE SINE TRANSFORM

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

$$\psi(k, n) = \sqrt{\frac{2}{N+1}} \sin \frac{\pi(k+1)(n+1)}{N+1}, \qquad 0 \le k, n \le N-1 \tag{5.98}$$

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

$$v(k) = \sqrt{\frac{2}{N+1}} \sum_{n=0}^{N-1} u(n) \sin \frac{\pi(k+1)(n+1)}{N+1}, \qquad 0 \le k \le N-1 \tag{5.99}$$

$$u(n) = \sqrt{\frac{2}{N+1}} \sum_{k=0}^{N-1} v(k) \sin \frac{\pi(k+1)(n+1)}{N+1}, \qquad 0 \le n \le N-1 \tag{5.100}$$

The two-dimensional sine transform pair for $N \times N$ images is obtained by substituting $\mathbf{A} = \mathbf{A}^* = \mathbf{A}^T = \mathbf{\Psi}$ in (5.11) and (5.12). The basis vectors and the basis images of the sine transform are shown in Figs. 5.1 and 5.2. Figure 5.12 shows the sine transform of a 255×255 image. Once again it is seen that a large fraction of the total energy is concentrated in a few transform coefficients.

Properties of the Sine Transform

1. The sine transform is real, symmetric, and orthogonal, that is,

$$\mathbf{\Psi}^* = \mathbf{\Psi} = \mathbf{\Psi}^T = \mathbf{\Psi}^{-1} \tag{5.101}$$

Thus, the forward and inverse sine transforms are identical.

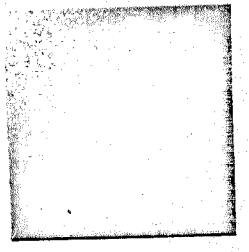


Figure 5.12 Sine transform of a 255×255 portion of the 256×256 image shown in Fig. 5.6a.

2. 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 (see Problem 5.16).
3. The sine transform is a fast transform. The sine transform (or its inverse) of a vector of N elements can be calculated in $O(N \log_2 N)$ operations via a $2(N+1)$-point FFT.
   Typically this requires $N + 1 = 2^p$, that is, the fast sine transform is usually defined for $N = 3, 7, 15, 31, 63, 255, \ldots$. Fast sine transform algorithms that do not require complex arithmetic (or the FFT) are also possible. In fact, these algorithms are somewhat faster than the FFT and the fast cosine transform algorithms [20].
4. The basis vectors of the sine transform are the eigenvectors of the symmetric tridiagonal Toeplitz matrix

$$\mathbf{Q} = \begin{bmatrix} 1 & -\alpha & 0 \\ -\alpha & \ddots & -\alpha \\ 0 & -\alpha & 1 \end{bmatrix} \tag{5.102}$$

5. The sine transform is close to the KL transform of first order stationary Markov sequences, whose covariance matrix is given in (2.68), when the correlation parameter ρ lies in the interval $(-0.5, 0.5)$. In general it has very good to excellent energy compaction property for images.
6. The sine transform leads to a fast KL transform algorithm for Markov sequences, whose boundary values are given. This makes it useful in many image processing problems. Details are considered in greater depth in Chapter 6 (Sections 6.5 and 6.9) and Chapter 11 (Section 11.5).

5.8 THE HADAMARD TRANSFORM

Unlike the previously discussed transforms, the elements of the basis vectors of the Hadamard transform take only the binary values ± 1 and are, therefore, well suited for digital signal processing. The Hadamard transform matrices, $\mathbf{H}_n$, are $N \times N$ matrices, where $N \triangleq 2^n$, $n = 1, 2, 3$. These can be easily generated by the core matrix

$$\mathbf{H}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{5.103}$$

and the Kronecker product recursion

$$\mathbf{H}_n = \mathbf{H}_{n-1} \otimes \mathbf{H}_1 = \mathbf{H}_1 \otimes \mathbf{H}_{n-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{H}_{n-1} & \mathbf{H}_{n-1} \\ \mathbf{H}_{n-1} & -\mathbf{H}_{n-1} \end{pmatrix} \tag{5.104}$$

As an example, for $n = 3$, the Hadamard matrix becomes

$$\mathbf{H}_3 = \mathbf{H}_1 \otimes \mathbf{H}_2 \tag{5.105}$$

$$\mathbf{H}_2 = \mathbf{H}_1 \otimes \mathbf{H}_1 \tag{5.106}$$

which gives

$$\mathbf{H}_3 = \frac{1}{\sqrt{8}} \left[\begin{array}{cccc|cccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ \hline 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{array}\right] \begin{array}{c} \text{Sequency} \\ 0 \\ 7 \\ 3 \\ 4 \\ \\ 1 \\ 6 \\ 2 \\ 5 \end{array} \tag{5.107}$$

The basis vectors of the Hadamard transform can also be generated by sampling a class of functions called the *Walsh functions*. These functions also take only the binary values ± 1 and form a complete orthonormal basis for square integrable functions. For this reason the Hadamard transform just defined is also called the Walsh-Hadamard transform.

The number of zero crossings of a Walsh function or the number of transitions in a basis vector of the Hadamard transform is called its *sequency*. Recall that for sinusoidal signals, frequency can be defined in terms of the zero crossings. In the Hadamard matrix generated via (5.104), the row vectors are not sequency ordered. The existing sequency order of these vectors is called the *Hadamard order*. The Hadamard transform of an $N \times 1$ vector $\mathbf{u}$ is written as

$$\mathbf{v} = \mathbf{H}\mathbf{u} \tag{5.108}$$

and the inverse transform is given by

$$\mathbf{u} = \mathbf{H}\mathbf{v} \tag{5.109}$$

where $\mathbf{H} \triangleq \mathbf{H}_n, n = \log_2 N$. In series form the transform pair becomes

$$v(k) = \frac{1}{\sqrt{N}} \sum_{m=0}^{N-1} u(m)(-1)^{b(k,m)}, \qquad 0 \le k \le N-1 \tag{5.110}$$

$$u(m) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} v(k)(-1)^{b(k,m)}, \qquad 0 \le m \le N-1 \tag{5.111}$$

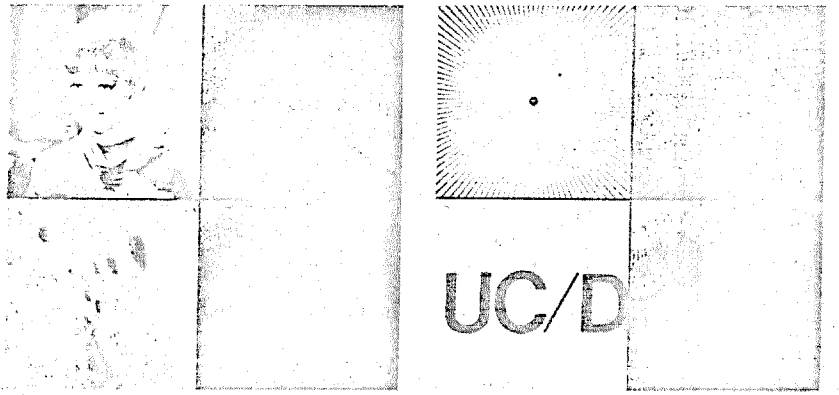
where

$$b(k,m) = \sum_{i=0}^{n-1} k_i m_i; \qquad k_i, m_i = 0, 1 \tag{5.112}$$

and $\{k_i\}, \{m_i\}$ are the binary representations of k and m, respectively, that is,

$$\left.\begin{aligned} k &= k_0 + 2k_1 + \cdots + 2^{n-1}k_{n-1} \\ m &= m_0 + 2m_1 + \cdots + 2^{n-1}m_{n-1} \end{aligned}\right\} \tag{5.113}$$

The two-dimensional Hadamard transform pair for $N \times N$ images is obtained by substituting $\mathbf{A} = \mathbf{A}^* = \mathbf{A}^T = \mathbf{H}$ in (5.11) and (5.12). The basis vectors and the basis



(a) Hadamard transforms of monochrome images. (b) Hadamard transforms of binary images.

Figure 5.13 Examples of Hadamard transforms.

images of the Hadamard transform are shown in Figs. 5.1 and 5.2. Examples of two-dimensional Hadamard transforms of images are shown in Fig. 5.13.

Properties of the Hadamard Transform

1. The Hadamard transform $\mathbf{H}$ is real, symmetric, and orthogonal, that is,

$$\mathbf{H} = \mathbf{H}^* = \mathbf{H}^T = \mathbf{H}^{-1} \tag{5.114}$$

2. The Hadamard transform is a fast transform. The one-dimensional transformation of (5.108) can be implemented in $O(N \log_2 N)$ additions and subtractions.

   Since the Hadamard transform contains only ± 1 values, no multiplications are required in the transform calculations. Moreover, the number of additions or subtractions required can be reduced from N^2 to about $N \log_2 N$. This is due to the fact that $\mathbf{H}_n$ can be written as a product of n sparse matrices, that is,

$$\mathbf{H} = \mathbf{H}_n = \tilde{\mathbf{H}}^n, \qquad n = \log_2 N \tag{5.115}$$

where

$$\tilde{\mathbf{H}} \triangleq \frac{1}{\sqrt{2}} \left[\begin{array}{cccccccc} 1 & 1 & 0 & 0 & & & & \\ 0 & 0 & 1 & 1 & 0 & 0 & & \\ \vdots & \vdots & \vdots & \vdots & & & & \\ 0 & 0 & & & \cdots & 1 & 1 & \\ \hline 1 & -1 & 0 & 0 & & & & \\ 0 & 0 & 1 & -1 & & & & \\ \vdots & \vdots & \vdots & \vdots & & & & \\ 0 & 0 & & & \cdots & 1 & -1 & \end{array}\right] \begin{array}{l} \uparrow \\ \frac{N}{2} \text{ rows} \\ \downarrow \\ \\ \uparrow \\ \frac{N}{2} \text{ rows} \\ \downarrow \end{array} \tag{5.116}$$

Since $\tilde{\mathbf{H}}$ contains only two nonzero terms per row, the transformation

$$\mathbf{v} = \tilde{\mathbf{H}}_n^n \mathbf{u} = \underbrace{\tilde{\mathbf{H}}\tilde{\mathbf{H}} \ldots \tilde{\mathbf{H}}}_{n \text{ terms}} \mathbf{u}, \qquad n = \log_2 N \tag{5.117}$$

can be accomplished by operating $\tilde{\mathbf{H}}$ n times on $\mathbf{u}$. Due to the structure of $\tilde{\mathbf{H}}$ only N additions or subtractions are required each time $\tilde{\mathbf{H}}$ operates on a vector, giving a total of $Nn = N \log_2 N$ additions or subtractions.

3. The natural order of the Hadamard transform coefficients turns out to be equal to the bit reversed gray code representation of its sequency s. If the sequency s has the binary representation $b_n b_{n-1} \ldots b_1$ and if the corresponding gray code is $g_n g_{n-1} \ldots g_1$, then the bit-reversed representation $g_1 g_2 \ldots g_n$ gives the natural order. Table 5.1 shows the conversion of sequency s to natural order h, and vice versa, for $N = 8$. In general,

$$\begin{aligned} g_k &= b_k \oplus b_{k+1}, \qquad k = 1, \ldots, n-1 \\ g_n &= b_n \\ h_k &= g_{n-k+1} \end{aligned} \tag{5.118}$$

and

$$\begin{aligned} g_k &= h_{n-k+1} \\ b_k &= g_k \oplus b_{k+1}, \qquad k = n-1, \ldots, 1 \\ b_n &= g_n \end{aligned} \tag{5.119}$$

give the forward and reverse conversion formulas for the sequency and natural ordering.

4. The Hadamard transform has good to very good energy compaction for highly correlated images. Let $\{u(n), 0 \le n \le N-1\}$ be a stationary random sequence with autocorrelation $r(n), 0 \le n \le N-1$. The fraction of the expected energy packed in the first $N/2^j$ sequency ordered Hadamard transform coefficients is given by [23]

$$\mathcal{E}\left(\frac{N}{2^j}\right) \triangleq \frac{\sum_{k=0}^{(N/2^j)-1} D_k'}{\sum_{k=0}^{N-1} D_k} = \frac{\left[1 + 2\sum_{k=1}^{2^j-1}\left(1 - \frac{k}{2^j}\right)\frac{r(k)}{r(0)}\right]}{2^j} \tag{5.120}$$

$$D_k \triangleq [\mathbf{HRH}]_{k,k}, \qquad \mathbf{R} \triangleq \{r(m-n)\} \tag{5.121}$$

where D_k' are the first $N/2^j$ sequency ordered elements D_k. Note that the D_k are simply the mean square values of the transform coefficients $[\mathbf{Hu}]_k$. The significance of this result is that $\mathcal{E}(N/2^j)$ depends on the first 2^j autocorrelations only. For $j = 1$, the fractional energy packed in the first $N/2$ sequency ordered coefficients will be $(1 + r(1)/r(0)/2$ and depends only upon the one-step correlation $\rho \triangleq r(1)/r(0)$. Thus for $\rho = 0.95$, 97.5% of the total energy is concentrated in half of the transform coefficients. The result of (5.120) is useful in calculating the energy compaction efficiency of the Hadamard transform.

TABLE 5.1 Natural Ordering versus Sequency Ordering of Hadamard Transform Coefficients for $N = 8$

Natural order h	Binary representation $h_3 h_2 h_1$	Gray code of s or reverse binary representation $g_3 g_2 g_1 = h_1 h_2 h_3$	Sequency binary representation $b_3 b_2 b_1$	Sequency s
0	000	000	000	0
1	001	100	111	7
2	010	010	011	3
3	011	110	100	4
4	100	001	001	1
5	101	101	110	6
6	110	011	010	2
7	111	111	101	5

Example 5.3

Consider the covariance matrix $\mathbf{R}$ of (2.68) for $N = 4$. Using the definition of $\mathbf{H}_2$ we obtain

$$\mathbf{D} = \text{diagonal}\,[\mathbf{H}_2 \mathbf{R} \mathbf{H}_2] = \tfrac{1}{4}\begin{bmatrix} 4+6\rho+4\rho^2+2\rho^3 & 0 & & \\ & 4-6\rho+4\rho^2-2\rho^3 & & \\ & & 4+2\rho-4\rho^2-2\rho^3 & \\ 0 & & & 4-2\rho-4\rho^2+2\rho^3 \end{bmatrix} \begin{matrix} 0 \\ 3 \\ 1 \\ 2 \end{matrix} \quad \text{Sequency}$$

This gives $D_0' = D_0$, $D_1' = D_2$, $D_2' = D_3$, $D_3' = D_1$ and

$$\mathcal{E}\left(\frac{1}{2}\right) = \frac{1}{4}\sum_{k=0}^{1} D_k' = \frac{1}{16}(4 + 6\rho + 4\rho^2 + 2\rho^3 + 4 + 2\rho - 4\rho^2 - 2\rho^3) = \frac{(1+\rho)}{2}$$

as expected according to (5.120).

5.9 THE HAAR TRANSFORM

The Haar functions $h_k(x)$ are defined on a continuous interval, $x \in [0, 1]$, and for $k = 0, \ldots, N-1$, where $N = 2^n$. The integer k can be uniquely decomposed as

$$k = 2^p + q - 1 \tag{5.122}$$

where $0 \le p \le n-1$; $q = 0, 1$ for $p = 0$ and $1 \le q \le 2^p$ for $p \ne 0$. For example, when $N = 4$, we have

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

Representing k by (p, q), the Haar functions are defined as

$$h_0(x) \triangleq h_{0,0}(x) = \frac{1}{\sqrt{N}}, \qquad x \in [0,1]. \tag{5.123a}$$

$$h_k(x) \triangleq h_{p,q}(x) = \frac{1}{\sqrt{N}} \begin{cases} 2^{p/2}, & \dfrac{q-1}{2^p} \le x < \dfrac{q-\frac{1}{2}}{2^p} \\ -2^{p/2}, & \dfrac{q-\frac{1}{2}}{2^p} \le x < \dfrac{q}{2^p} \\ 0, & \text{otherwise for } x \in [0,1] \end{cases} \tag{5.123b}$$

The Haar transform is obtained by letting x take discrete values at m/N, $m = 0, 1, \ldots, N-1$. For $N = 8$, the Haar transform is given by

$$\mathbf{Hr} = \frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} \\ 2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 \end{bmatrix} \begin{matrix} \text{Sequency} \\ 0 \\ 1 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \end{matrix} \tag{5.124}$$

The basis vectors and the basis images of the Haar transform are shown in Figs. 5.1 and 5.2. An example of the Haar transform of an image is shown in Fig. 5.14. From the structure of **Hr** [see (5.124)] we see that the Haar transform takes differences of the samples or differences of local averages of the samples of the input vector. Hence the two-dimensional Haar transform coefficients $y(k, l)$, except for $k = l = 0$, are the differences along rows and columns of the local averages of pixels in the image. These are manifested as several "edge extractions" of the original image, as is evident from Fig. 5.14.

Although some work has been done for using the Haar transform in image data compression problems, its full potential in feature extraction and image analysis problems has not been determined.

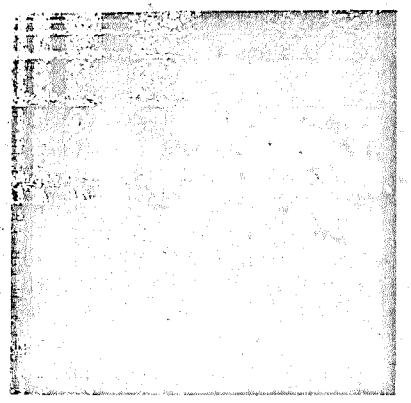


Figure 5.14 Haar transform of the 256 × 256 image shown in Fig. 5.6a.

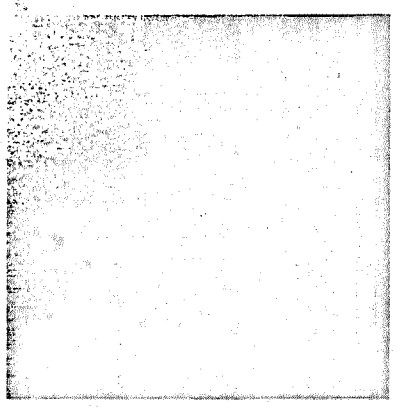


Figure 5.15 Slant transform of the 256 × 256 image shown in Fig. 5.6a.

Properties of the Haar Transform

1. The Haar transform is real and orthogonal. Therefore,
$$\mathbf{Hr} = \mathbf{Hr}^*$$
$$\mathbf{Hr}^{-1} = \mathbf{Hr}^T \tag{5.125}$$
2. The Haar transform is a very fast transform. On an $N \times 1$ vector it can be implemented in $O(N)$ operations.
3. The basis vectors of the Haar matrix *are* sequency ordered.
4. The Haar transform has poor energy compaction for images.

5.10 THE SLANT TRANSFORM

The $N \times N$ Slant transform matrices are defined by the recursion

$$\mathbf{S}_n = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & & 1 & 0 & & \\ a_n & b_n & \mathbf{0} & -a_n & b_n & \mathbf{0} & \\ \hline \mathbf{0} & & \mathbf{I}_{(N/2)-2} & \mathbf{0} & & I_{(N/2)-2} & \\ \hline 0 & 1 & & 0 & -1 & & \\ -b_n & a_n & \mathbf{0} & b_n & a_n & \mathbf{0} & \\ \hline \mathbf{0} & & I_{(N/2)-2} & \mathbf{0} & & -\mathbf{I}_{(N/2)-2} & \end{bmatrix} \begin{bmatrix} \mathbf{S}_{n-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_{n-1} \end{bmatrix} \tag{5.126}$$

where $N = 2^n$, $\mathbf{I}_M$ denotes an $M \times M$ identity matrix, and

$$\mathbf{S}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{5.127}$$

The parameters a_n and b_n are defined by the recursions

$$\left.\begin{aligned} b_n &= (1 + 4a_{n-1}^2)^{-1/2}, \qquad a_1 = 1 \\ a_n &= 2b_n a_{n-1} \end{aligned}\right\} \tag{5.128}$$

which solve to give

$$a_{n+1} = \left(\frac{3N^2}{4N^2 - 1}\right)^{1/2}, \qquad b_{n+1} = \left(\frac{N^2 - 1}{4N^2 - 1}\right)^{1/2}, \qquad N = 2^n \tag{5.129}$$

Using these formulas, the 4×4 Slant transformation matrix is obtained as

$$\mathbf{S}_2 = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ \frac{3}{\sqrt{5}} & \frac{1}{\sqrt{5}} & \frac{-1}{\sqrt{5}} & \frac{-3}{\sqrt{5}} \\ 1 & -1 & -1 & 1 \\ \frac{1}{\sqrt{5}} & \frac{-3}{\sqrt{5}} & \frac{3}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \end{bmatrix} \begin{matrix} \text{Sequency} \\ 0 \\ 1 \\ 2 \\ 3 \end{matrix} \tag{5.130}$$

Figure 5.1 shows the basis vectors of the 8×8 Slant transform. Figure 5.2 shows the basis images of the 8×8 two dimensional Slant transform. Figure 5.15 shows the Slant transform of a 256×256 image.

Properties of the Slant Transform

1. The Slant transform is real and orthogonal. Therefore,
$$\mathbf{S} = \mathbf{S}^*, \qquad \mathbf{S}^{-1} = \mathbf{S}^T \tag{5.131}$$
2. The Slant transform is a fast transform, which can be implemented in $O(N \log_2 N)$ operations on an $N \times 1$ vector.
3. It has very good to excellent energy compaction for images.
4. The basis vectors of the Slant transform matrix $\mathbf{S}$ are not sequency ordered for $n \geq 3$. If $\mathbf{S}_{n-1}$ is sequency ordered, the ith row sequency of $\mathbf{S}_n$ is given as follows.

$$\begin{aligned} i &= 0, && \text{sequency} = 0 \\ i &= 1, && \text{sequency} = 1 \\ 2 \leq i \leq \frac{N}{2} - 1, && \text{sequency} = \begin{cases} 2i, & i = \text{even} \\ 2i + 1, & i = \text{odd} \end{cases} \\ i &= \frac{N}{2}, && \text{sequency} = 2 \end{aligned}$$

$$\begin{aligned} i &= \frac{N}{2} + 1, && \text{sequency} = 3 \\ \frac{N}{2} + 2 \leq i \leq N - 1, && \text{sequency} = \begin{cases} 2\left(i - \frac{N}{2}\right) + 1, & i = \text{even} \\ 2\left(i - \frac{N}{2}\right), & i = \text{odd} \end{cases} \end{aligned}$$

5.11 THE KL TRANSFORM

The KL transform was originally introduced as a series expansion for continuous random processes by Karhunen [27] and Loeve [28]. For random sequences Hotelling [26] 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.

For a real $N \times 1$ random vector $\mathbf{u}$, the basis vectors of the KL transform (see Section 2.9) are given by the orthonormalized eigenvectors of its autocorrelation matrix $\mathbf{R}$, that is,

$$\mathbf{R}\boldsymbol{\phi}_k = \lambda_k \boldsymbol{\phi}_k, \qquad 0 \leq k \leq N - 1 \tag{5.132}$$

The KL transform of $\mathbf{u}$ is defined as

$$\mathbf{v} = \boldsymbol{\Phi}^{*T}\mathbf{u} \tag{5.133}$$

and the inverse transform is

$$\mathbf{u} = \boldsymbol{\Phi}\mathbf{v} = \sum_{k=0}^{N-1} v(k)\boldsymbol{\phi}_k \tag{5.134}$$

where $\boldsymbol{\phi}_k$ is the kth column of $\boldsymbol{\Phi}$. From (2.44) we know $\boldsymbol{\Phi}$ reduces $\mathbf{R}$ to its diagonal form, that is,

$$\boldsymbol{\Phi}^{*T}\mathbf{R}\boldsymbol{\Phi} = \boldsymbol{\Lambda} = \text{Diag}\{\lambda_k\} \tag{5.135}$$

We often work with the covariance matrix rather than the autocorrelation matrix. With $\boldsymbol{\mu} \triangleq E[\mathbf{u}]$, then

$$\mathbf{R}_0 \triangleq \text{cov}[\mathbf{u}] \triangleq E[(\mathbf{u} - \boldsymbol{\mu})(\mathbf{u} - \boldsymbol{\mu})^T] = E[\mathbf{u}\mathbf{u}^T] - \boldsymbol{\mu}\boldsymbol{\mu}^T = \mathbf{R} - \boldsymbol{\mu}\boldsymbol{\mu}^T \tag{5.136}$$

If the vector $\boldsymbol{\mu}$ is known, then the eigenmatrix of $\mathbf{R}_0$ determines the KL transform of the zero mean random process $\mathbf{u} - \boldsymbol{\mu}$. In general, the KL transform of $\mathbf{u}$ and $\mathbf{u} - \boldsymbol{\mu}$ need not be identical.

Note that whereas the image transforms considered earlier were functionally independent of the data, the KL transform depends on the (second-order) statistics of the data.

Example 5.4 (KL Transform of Markov-1 Sequences)

The covariance matrix of a zero mean Markov sequence of N elements is given by (2.68). Its eigenvalues λ_k and the eigenvectors $\boldsymbol{\phi}_k$ are given by

$$\lambda_k = \frac{1 - \rho^2}{1 - 2\rho\cos\omega_k + \rho^2}$$

$$\phi_k(m) = \phi(m,k) \tag{5.137}$$
$$= \left(\frac{2}{N+\lambda_k}\right)^{1/2} \sin\left(\omega_k\left(m+1-\frac{N+1}{2}\right)+\frac{(k+1)\pi}{2}\right), \qquad 0 \le m,k \le N-1$$

where the $\{\omega_k\}$ are the positive roots of the equation

$$\tan(N\omega) = -\frac{(1-\rho^2)\sin\omega}{\cos\omega - 2\rho + \rho^2\cos\omega}, \qquad N \text{ even} \tag{5.138}$$

A similar result holds when N is odd. This is a transcendental equation that gives rise to nonharmonic sinusoids $\phi_k(m)$. Figure 5.1 shows the basis vectors of this 8×8 KL transform for $\rho = 0.95$. Note the basis vectors of the KLT and the DCT are quite similar. Because $\phi_k(m)$ are nonharmonic, a fast algorithm for this transform does not exist. Also note, the KL transform matrix is $\mathbf{\Phi}^T \triangleq \{\phi(k,m)\}$.

Example 5.5

Since the unitary DFT reduces any circulant matrix to a diagonal form, it is the KL transform of all random sequences with circulant autocorrelation matrices, that is, for all periodic random sequences.

The DCT is the KL transform of a random sequence whose autocorrelation matrix $\mathbf{R}$ commutes with $\mathbf{Q}_c$ of (5.95) (that is, if $\mathbf{RQ}_c = \mathbf{Q}_c\mathbf{R}$). Similarly, the DST is the KL transform of all random sequences whose autocorrelation matrices commute with $\mathbf{Q}$ of (5.102).

KL Transform of Images

If an $N\times N$ image $u(m,n)$ is represented by a random field whose autocorrelation function is given by

$$E[u(m,n)u(m',n')] = r(m,n;m',n'), \qquad 0 \le m,m',n,n' \le N-1 \tag{5.139}$$

then the basis images of the KL transform are the orthonormalized eigenfunctions $\psi_{k,l}(m,n)$ obtained by solving

$$\sum_{m'=0}^{N-1}\sum_{n'=0}^{N-1} r(m,n;m',n')\psi_{k,l}(m',n') = \lambda_{k,l}\psi_{k,l}(m,n), \qquad 0 \le k,l \le N-1,\ 0 \le m,n \le N-1 \tag{5.140}$$

In matrix notation this can be written as

$$\mathcal{R}\psi_i = \lambda_i\psi_i, \qquad i = 0,\ldots,N^2-1 \tag{5.141}$$

where ψ_i is an $N^2\times 1$ vector representation of $\psi_{k,l}(m,n)$ and $\mathcal{R}$ is an $N^2\times N^2$ autocorrelation matrix of the image mapped into an $N^2\times 1$ vector $\boldsymbol{u}$. Thus

$$\mathcal{R} = E[\boldsymbol{u}\boldsymbol{u}^T] \tag{5.142}$$

If $\mathcal{R}$ is separable, then the $N^2\times N^2$ matrix $\mathbf{\Psi}$ whose columns are $\{\psi_i\}$ becomes separable (see Table 2.7). For example, let

$$r(m,n;m',n') = r_1(m,m')r_2(n,n') \tag{5.143}$$

$$\psi_{k,l}(m,n) = \phi_1(m,k)\phi_2(n,l) \tag{5.144}$$

In matrix notation this means

$$\mathcal{R} = \mathbf{R}_1 \otimes \mathbf{R}_2, \qquad \mathbf{\Psi} = \mathbf{\Phi}_1 \otimes \mathbf{\Phi}_2 \tag{5.145}$$

where

$$\mathbf{\Phi}_j\mathbf{R}_j\mathbf{\Phi}_j^{*T} = \mathbf{\Lambda}_j, \qquad j = 1,2 \tag{5.146}$$

and the KL transform of $\boldsymbol{u}$ is

$$\boldsymbol{v} = \mathbf{\Psi}^{*T}\boldsymbol{u} = [\mathbf{\Phi}_1^{*T}\otimes\mathbf{\Phi}_2^{*T}]\boldsymbol{u} \tag{5.147}$$

For row-ordered vectors this is equivalent to

$$\mathbf{V} = \mathbf{\Phi}_1^{*T}\mathbf{U}\mathbf{\Phi}_2^{*} \tag{5.148}$$

and the inverse KL transform is

$$\mathbf{U} = \mathbf{\Phi}_1\mathbf{V}\mathbf{\Phi}_2^{T} \tag{5.149}$$

The advantage in modeling the image autocorrelation by a separable function is that instead of solving the $N^2\times N^2$ matrix eigenvalue problem of (5.141), only two $N\times N$ matrix eigenvalue problems of (5.146) need to be solved. Since an $N\times N$ matrix eigenvalue problem requires $O(N^3)$ computations, the reduction in dimensionality achieved by the separable model is $O(N^6)/O(N^3) = O(N^3)$, which is very significant. Also, the transformation calculations of (5.148) and (5.149) require $2N^3$ operations compared to N^4 operations required for $\mathbf{\Psi}^{*T}\boldsymbol{u}$.

Example 5.6

Consider the separable covariance function for a zero mean random field

$$r(m,n;m',n') = \rho^{|m-m'|}\rho^{|n-n'|} \tag{5.150}$$

This gives $\mathcal{R} = \mathbf{R}\otimes\mathbf{R}$, where $\mathbf{R}$ is given by (2.68). The eigenvectors of $\mathbf{R}$ are given by $\boldsymbol{\phi}_k$ in Example 5.4. Hence $\mathbf{\Psi} = \mathbf{\Phi}\otimes\mathbf{\Phi}$ and the KL transform matrix is $\mathbf{\Phi}^T\otimes\mathbf{\Phi}^T$. Figure 5.2 shows the basis images of this 8×8 two-dimensional KL transform for $\rho = 0.95$.

Properties of the KL Transform

The KL transform has many desirable properties, which make it optimal in many signal processing applications. Some of these properties are discussed here. For simplicity we assume $\mathbf{u}$ has zero mean and a positive definite covariance matrix $\mathbf{R}$.

Decorrelation. The KL transform coefficients $\{v(k),\ k = 0,\ldots,N-1\}$ are uncorrelated and have zero mean, that is,

$$E[v(k)] = 0$$
$$E[v(k)v^*(l)] = \lambda_k\delta(k-l) \tag{5.151}$$

The proof follows directly from (5.133) and (5.135), since

$$E[\mathbf{v}\mathbf{v}^{*T}] \triangleq \mathbf{\Phi}^{*T}E[\mathbf{u}\mathbf{u}^T]\mathbf{\Phi} = \mathbf{\Phi}^{*T}\mathbf{R}\mathbf{\Phi} = \mathbf{\Lambda} = \text{diagonal} \tag{5.152}$$

which implies the latter relation in (5.151). It should be noted that Φ is not a *unique matrix* with respect to this property. There could be many matrices (unitary and nonunitary) that would decorrelate the transformed sequence. For example, a lower triangular matrix Φ could be found that satisfies (5.152).

Example 5.7

The covariance matrix $\mathbf{R}$ of (2.68) is diagonalized by the lower triangular matrix

$$\mathbf{L} \triangleq \begin{bmatrix} 1 & & 0 \\ -\rho & & \\ 0 & -\rho & 1 \end{bmatrix} \Rightarrow \mathbf{L}^T\mathbf{R}\mathbf{L} = \begin{bmatrix} 1-\rho^2 & & 0 \\ & 1-\rho^2 & \\ 0 & & 1-\rho^2 \\ & & & 1 \end{bmatrix} \triangleq \mathbf{D} \tag{5.153}$$

Hence the transformation $\mathbf{v} = \mathbf{L}^T\mathbf{u}$, will cause the sequence $v(k)$ to be uncorrelated. Comparing with Example 5.4, we see that $\mathbf{L} \neq \mathbf{\Phi}$. Moreover, $\mathbf{L}$ is not unitary and the diagonal elements of $\mathbf{D}$ are not the eigenvalues of $\mathbf{R}$.

Basis restriction mean square error. Consider the operations in Fig. 5.16. The vector $\mathbf{u}$ is first transformed to $\mathbf{v}$. The elements of $\mathbf{w}$ are chosen to be the first m elements of $\mathbf{v}$ and zeros elsewhere. Finally, $\mathbf{w}$ is transformed to $\mathbf{z}$. $\mathbf{A}$ and $\mathbf{B}$ are $N \times N$ matrices and $\mathbf{I}_m$ is a matrix with 1s along the first m diagonal terms and zeros elsewhere. Hence

$$w(k) = \begin{cases} v(k), & 0 \le k \le m-1 \\ 0, & k \ge m \end{cases} \tag{5.154}$$

Therefore, whereas $\mathbf{u}$ and $\mathbf{v}$ are vectors in an N-dimensional vector space, $\mathbf{w}$ is a vector restricted to an $m \le N$-dimensional subspace. The average mean square error between the sequences $u(n)$ and $z(n)$ is defined as

$$J_m \triangleq \frac{1}{N} E\left(\sum_{n=0}^{N-1} |u(n) - z(n)|^2\right) = \frac{1}{N} \operatorname{Tr}[E\{(\mathbf{u}-\mathbf{z})(\mathbf{u}-\mathbf{z})^{*T}\}] \tag{5.155}$$

This quantity is called the *basis restriction error*. It is desired to find the matrices $\mathbf{A}$ and $\mathbf{B}$ such that J_m is minimized for each and every value of $m \in [1, N]$. This minimum is achieved by the KL transform of $\mathbf{u}$.

Theorem 5.1. The error J_m in (5.155) is minimum when

$$\mathbf{A} = \mathbf{\Phi}^{*T}, \qquad \mathbf{B} = \mathbf{\Phi}, \qquad \mathbf{AB} = \mathbf{I} \tag{5.156}$$

where the columns of $\mathbf{\Phi}$ are arranged according to the decreasing order of the eigenvalues of $\mathbf{R}$.

Proof. From Fig. 5.16, we have

$$\mathbf{v} = \mathbf{A}\mathbf{u}, \qquad \mathbf{w} = \mathbf{I}_m\mathbf{v}, \quad \text{and} \quad \mathbf{z} = \mathbf{B}\mathbf{w} \tag{5.157}$$

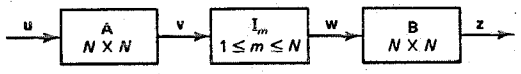


Figure 5.16 KL transform basis restriction.

Using these we can rewrite (5.155) as

$$J_m = \frac{1}{N} \operatorname{Tr}[(\mathbf{I} - \mathbf{B}\mathbf{I}_m\mathbf{A})\mathbf{R}(\mathbf{I} - \mathbf{B}\mathbf{I}_m\mathbf{A})^{*T}]$$

To minimize J_m we first differentiate it with respect to the elements of $\mathbf{A}$ and set the result to zero [see Problem 2.15 for the differentiation rules]. This gives

$$\mathbf{I}_m\mathbf{B}^T(\mathbf{I} - \mathbf{B}\mathbf{I}_m\mathbf{A})^*\mathbf{R} = \mathbf{0} \tag{5.158}$$

which yields

$$J_m = \frac{1}{N} \operatorname{Tr}[(\mathbf{I} - \mathbf{B}\mathbf{I}_m\mathbf{A})\mathbf{R}] \tag{5.159}$$

$$\mathbf{I}_m\mathbf{B}^{*T} = \mathbf{I}_m\mathbf{B}^{*T}\mathbf{B}\mathbf{I}_m\mathbf{A} \tag{5.160}$$

At $m = N$, the minimum value of J_N must be zero, which requires

$$\mathbf{I} - \mathbf{B}\mathbf{A} = \mathbf{0} \quad \text{or} \quad \mathbf{B} = \mathbf{A}^{-1} \tag{5.161}$$

Using this in (5.160) and rearranging terms, we obtain

$$\mathbf{I}_m\mathbf{B}^{*T}\mathbf{B} = \mathbf{I}_m\mathbf{B}^{*T}\mathbf{B}\mathbf{I}_m, \qquad 1 \le m \le N \tag{5.162}$$

For (5.162) to be true for every m, it is necessary that $\mathbf{B}^{*T}\mathbf{B}$ be diagonal. Since $\mathbf{B} = \mathbf{A}^{-1}$, it is easy to see that (5.160) remains invariant if $\mathbf{B}$ is replaced by $\mathbf{DB}$ or $\mathbf{BD}$, where $\mathbf{D}$ is a diagonal matrix. Hence, without loss of generality we can normalize $\mathbf{B}$ so that $\mathbf{B}^{*T}\mathbf{B} = \mathbf{I}$, that is, $\mathbf{B}$ is a unitary matrix. Therefore, $\mathbf{A}$ is also unitary and $\mathbf{B} = \mathbf{A}^{*T}$. This gives

$$J_m = \frac{1}{N} \operatorname{Tr}((\mathbf{I} - \mathbf{A}^{*T}\mathbf{I}_m\mathbf{A})\mathbf{R}) = \frac{1}{N} \operatorname{Tr}(\mathbf{R} - \mathbf{I}_m\mathbf{A}\mathbf{R}\mathbf{A}^{*T}) \tag{5.163}$$

Since $\mathbf{R}$ is fixed, J_m is minimized if the quantity

$$\tilde{J}_m \triangleq \operatorname{Tr}(\mathbf{I}_m\mathbf{A}\mathbf{R}\mathbf{A}^{*T}) = \sum_{k=0}^{m-1} \mathbf{a}_k^T\mathbf{R}\mathbf{a}_k^* \tag{5.164}$$

is *maximized* where $\mathbf{a}_k^T$ is the kth row of $\mathbf{A}$. Since $\mathbf{A}$ is unitary,

$$\mathbf{a}_k^T\mathbf{a}_k^* = 1 \tag{5.165}$$

To maximize $\tilde{J}_m$ subject to (5.165), we form the Lagrangian

$$\tilde{J}_m = \sum_{k=0}^{m-1} \mathbf{a}_k^T\mathbf{R}\mathbf{a}_k^* + \sum_{k=0}^{m-1} \lambda_k(1 - \mathbf{a}_k^T\mathbf{a}_k^*) \tag{5.166}$$

and differentiate it with respect to $\mathbf{a}_j$. The result gives a necessary condition

$$\mathbf{R}\mathbf{a}_j^* = \lambda_j\mathbf{a}_j^* \tag{5.167}$$

where $\mathbf{a}_j^*$ are orthonormalized eigenvectors of $\mathbf{R}$. This yields

$$\tilde{J}_m = \sum_{k=0}^{m-1} \lambda_k \tag{5.168}$$

which is maximized if $\{\mathbf{a}_j^*,\ 0 \le j \le m-1\}$ correspond to the largest m eigenvalues of $\mathbf{R}$. Because J_m must be maximized for every m, it is necessary to arrange $\lambda_0 \ge \lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_{N-1}$. Then $\mathbf{a}_j^T$, the rows of $\mathbf{A}$, are the conjugate transpose of the eigenvectors of $\mathbf{R}$, that is, $\mathbf{A}$ is the KL transform of $\mathbf{u}$.

Distribution of variances. Among all the unitary transformations $\mathbf{v} = \mathbf{A}\mathbf{u}$, the KL transform Φ^{*T} packs the maximum average energy in $m \le N$ samples of $\mathbf{v}$. Define

$$\sigma_k^2 \triangleq E[|v(k)|^2], \qquad \sigma_0^2 \ge \sigma_1^2 \cdots \ge \sigma_{N-1}^2$$
$$S_m(\mathbf{A}) \triangleq \sum_{k=0}^{m-1} \sigma_k^2 \tag{5.169}$$

Then for any fixed $m \in [1, N]$

$$S_m(\Phi^{*T}) \ge S_m(\mathbf{A}) \tag{5.170}$$

Proof. Note that

$$\begin{aligned} S_m(\mathbf{A}) &= \sum_{k=0}^{m-1} (\mathbf{A}\mathbf{R}\mathbf{A}^{*T})_{k,k} \\ &= \mathrm{Tr}(\mathbf{I}_m \mathbf{A}^{*T}\mathbf{R}\mathbf{A}) \\ &= J_m \end{aligned}$$

which, we know from the last property [see (5.164)], is maximized when $\mathbf{A}$ is the KL transform. Since $\sigma_k^2 = \lambda_k$ when $\mathbf{A} = \Phi^{*T}$, from (5.168)

$$\sum_{k=0}^{m-1} \lambda_k \ge \sum_{k=0}^{m-1} \sigma_k^2, \qquad 1 \le m \le N \tag{5.171}$$

Threshold representation. The KL transform also minimizes $E[m]$, the expected number of transform coefficients required, so that their energy just exceeds a prescribed threshold (see Problem 5.26 and [33]).

A fast KL transform. In application of the KL transform to images, there are dimensionality difficulties. The KL transform depends on the statistics as well as the size of the image and, in general, the basis vectors are not known analytically. After the transform matrix has been computed, the operations for performing the transformation are quite large for images.

It has been shown that certain statistical image models yield a fast KL transform algorithm as an alternative to the conventional KL transform for images. It is based on a stochastic decomposition of an image as a sum of two random sequences. *The first random sequence is such that its KL transform is a fast transform* and the second sequence, called the *boundary response,* depends only on information at the boundary points of the image. For details see Sections 6.5 and 6.9.

The rate-distortion function. Suppose a random vector $\mathbf{u}$ is unitary transformed to $\mathbf{v}$ and transmitted over a communication channel (Fig. 5.17). Let $\mathbf{v}^*$ and $\mathbf{u}^*$

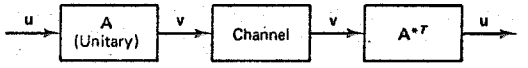


Figure 5.17 Unitary transform data transmission. Each element of $\mathbf{v}$ is coded independently.

be the reproduced values of $\mathbf{v}$ and $\mathbf{u}$, respectively. Further, assume that $\mathbf{u}$, $\mathbf{v}$, $\mathbf{v}^*$, and $\mathbf{u}^*$ are Gaussian. The average distortion in $\mathbf{u}$ is

$$D = \frac{1}{N} E[(\mathbf{u} - \mathbf{u}^*)^{*T}(\mathbf{u} - \mathbf{u}^*)] \tag{5.172}$$

Since $\mathbf{A}$ is unitary and $\mathbf{u} = \mathbf{A}^{*T}\mathbf{v}$ and $\mathbf{u}^* = \mathbf{A}^{*T}\mathbf{v}^*$, we have

$$\begin{aligned} D &= \frac{1}{N} E[(\mathbf{v} - \mathbf{v}^*)^{*T}\mathbf{A}\mathbf{A}^{*T}(\mathbf{v} - \mathbf{v}^*)] \\ &= \frac{1}{N} E[(\mathbf{v} - \mathbf{v}^*)^{*T}(\mathbf{v} - \mathbf{v}^*)] = \frac{1}{N} E(\delta\mathbf{v}^{*T}\delta\mathbf{v}) \end{aligned} \tag{5.173}$$

where $\delta\mathbf{v} = \mathbf{v} - \mathbf{v}^*$ represents the error in the reproduction of $\mathbf{v}$. From the preceding, D is invariant under all unitary transformations. The rate-distortion function is now obtained, following Section 2.13, as

$$R = \frac{1}{N} \sum_{k=0}^{N-1} \max\left[0, \tfrac{1}{2}\log_2 \frac{\sigma_k^2}{\theta}\right] \tag{5.174}$$

$$D = \frac{1}{N} \sum_{k=0}^{N-1} \min[\theta, \sigma_k^2] \tag{5.175}$$

where

$$\sigma_k^2 = E[|v(k)|^2] = [\mathbf{A}\mathbf{R}\mathbf{A}^{*T}]_{k,k} \tag{5.176}$$

depend on the transform $\mathbf{A}$. Therefore, the rate

$$R = R(\mathbf{A}) \tag{5.177}$$

also depends on $\mathbf{A}$. For each fixed D, *the KL transform achieves the minimum rate among all unitary transforms,* that is,

$$R(\Phi^{*T}) \le R(\mathbf{A}) \tag{5.178}$$

This property is discussed further in Chapter 11, on transform coding.

Example 5.8

Consider a 2×1 vector $\mathbf{u}$, whose covariance matrix is

$$\mathbf{R} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}, \qquad |\rho| < 1$$

The KL transform is

$$\Phi^{*T} = \Phi = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

The transformation $\mathbf{v} = \mathbf{\Phi u}$ gives

$$E\{[v(0)]^2\} = \lambda_0 = 1 + \rho, \qquad E\{[v(1)]^2\} = 1 - \rho$$

$$R(\Phi) = \tfrac{1}{2}\left[\max\left(0, \tfrac{1}{2}\log\frac{1+\rho}{\theta}\right) + \max\left(0, \tfrac{1}{2}\log\frac{1-\rho}{\theta}\right)\right]$$

Compare this with the case when $\mathbf{A} = \mathbf{I}$ (that is, $\mathbf{u}$ is transmitted), which gives $\sigma_0^2 = \sigma_1^2 = 1$, and

$$R(\mathbf{I}) = \tfrac{1}{4}[-2\log\theta], \qquad 0 < \theta < 1$$

Suppose we let θ be small, say $\theta < 1 - |\rho|$. Then it is easy to show that

$$R(\Phi) < R(\mathbf{I})$$

This means for a fixed level of distortion, the number of bits required to transmit the KLT sequence would be less than those required for transmission of the original sequence.

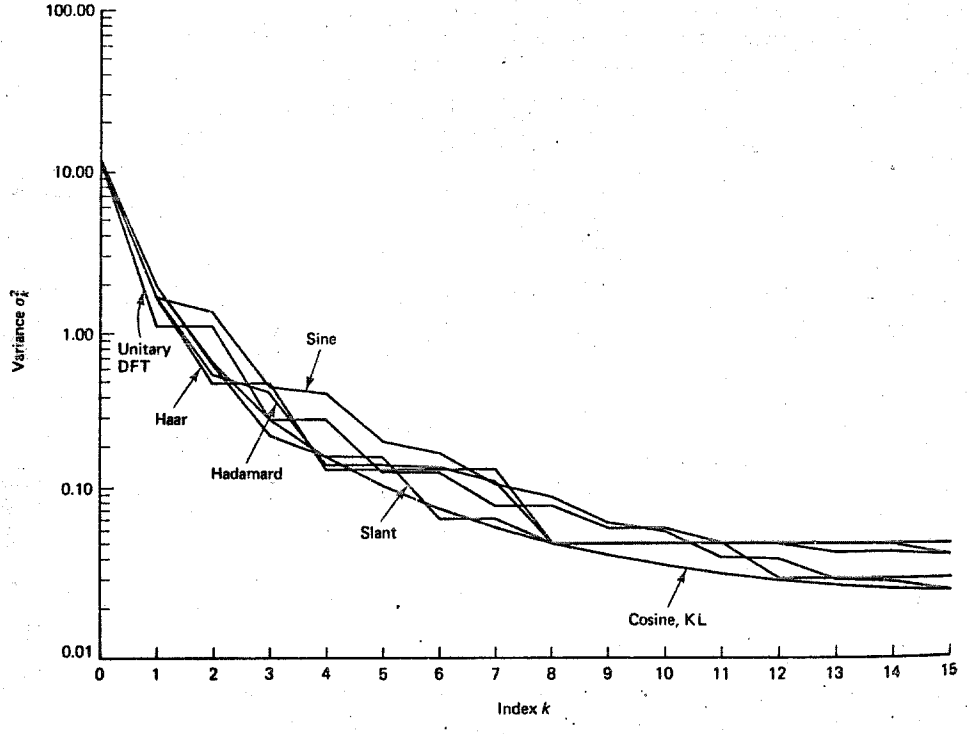


Figure 5.18 Distribution of variances of the transform coefficients (in decreasing order) of a stationary Markov sequence with $N = 16$, $\rho = 0.95$ (see Example 5.9).

TABLE 5.2 Variances σ_k^2 of Transform Coefficients of a Stationary Markov Sequence with $\rho = 0.95$ and $N = 16$. See Example 5.9.

$\downarrow k$ \ Transform	KL	Cosine	Sine	Unitary DFT	Hadamard	Haar	Slant
0	12.442	12.406	11.169	12.406	12.406	12.406	12.406
1	1.946	1.943	1.688	1.100	1.644	1.644	1.904
2	0.615	0.648	1.352	0.292	0.544	0.487	0.641
3	0.292	0.295	0.421	0.139	0.431	0.487	0.233
4	0.171	0.174	0.463	0.086	0.153	0.144	0.173
5	0.114	0.114	0.181	0.062	0.152	0.144	0.172
6	0.082	0.083	0.216	0.051	0.149	0.144	0.072
7	0.063	0.063	0.098	0.045	0.121	0.144	0.072
8	0.051	0.051	0.116	0.043	0.051	0.050	0.051
9	0.043	0.043	0.060	0.045	0.051	0.050	0.051
10	0.037	0.037	0.067	0.051	0.051	0.050	0.051
11	0.033	0.033	0.040	0.062	0.051	0.050	0.051
12	0.030	0.030	0.042	0.086	0.051	0.050	0.031
13	0.028	0.028	0.031	0.139	0.051	0.050	0.031
14	0.027	0.027	0.029	0.292	0.050	0.050	0.031
15	0.026	0.026	0.026	1.100	0.043	0.050	0.031

Example 5.9 (Comparison Among Unitary Transforms for a Markov Sequence)

Consider a first-order zero mean stationary Markov sequence of length N whose covariance matrix $\mathbf{R}$ is given by (2.68) with $\rho = 0.95$. Figure 5.18 shows the distribution of variances σ_k^2 of the transform coefficients (in decreasing order) for different transforms. Table 5.2 lists σ_k^2 for the various transforms.

Define the normalized basis restriction error as

$$J_m = \frac{\sum_{k=m}^{N-1} \sigma_k^2}{\sum_{k=0}^{N-1} \sigma_k^2}, \qquad m = 0, \ldots, N-1 \tag{5.179}$$

where σ_k^2 have been arranged in decreasing order.

Figure 5.19 shows J_m versus m for the various transforms. It is seen that the cosine transform performance is indistinguishable from that of the KL transform for $\rho = 0.95$. In general it seems possible to find a fast sinusoidal transform (that is, a transform consisting of sine or cosine functions) as a good substitute for the KL transform for different values of ρ as well as for higher-order stationary random sequences (see Section 5.12).

The mean square performance of the various transforms also depends on the dimension N of the transform. Such comparisons are made in Section 5.12.

Example 5.10 (Performance of Transforms on Images)

The mean square error test of the last example can be extended to actual images. Consider an $N \times N$ image $u(m, n)$ from which its mean is subtracted out to make it zero mean. The transform coefficient variances are estimated as

$$\sigma^2(k, l) = E[|v(k, l)|^2] \cong |v(k, l)|^2$$

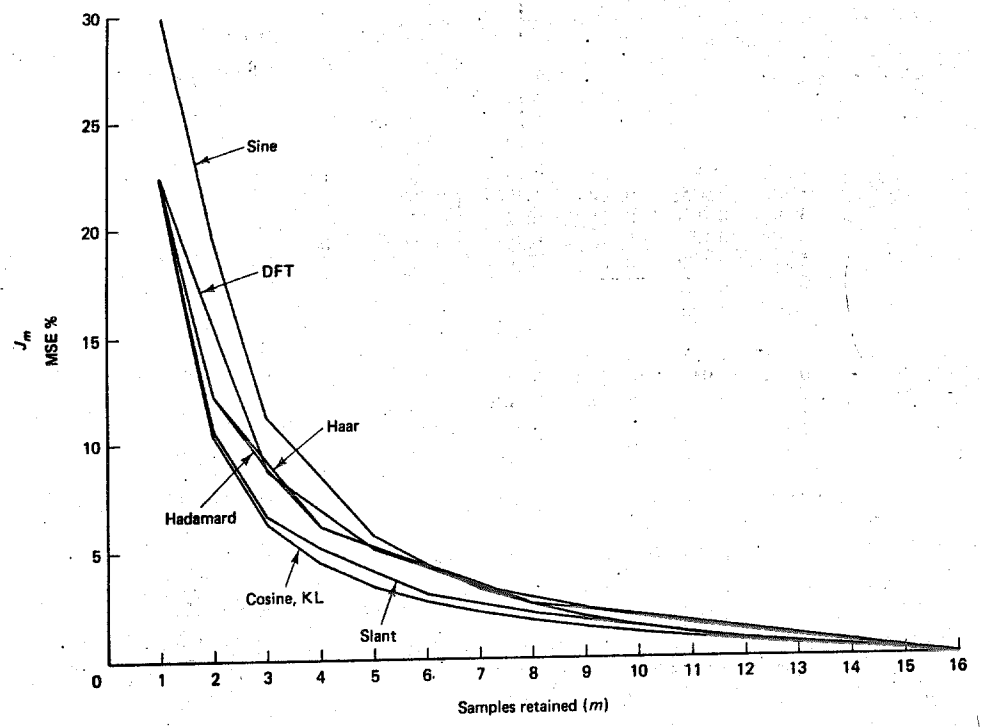


Figure 5.19 Performance of different unitary transforms with respect to basis restriction errors (J_m) versus the number of basis (m) for a stationary Markov sequence with $N = 16$, $\rho = 0.95$.

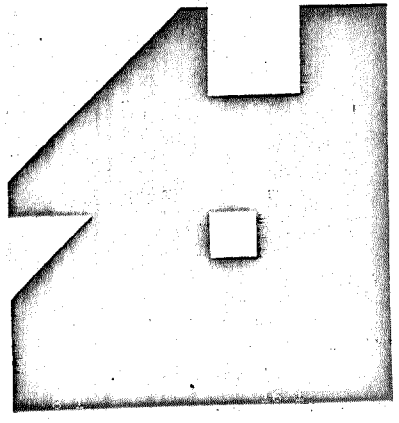
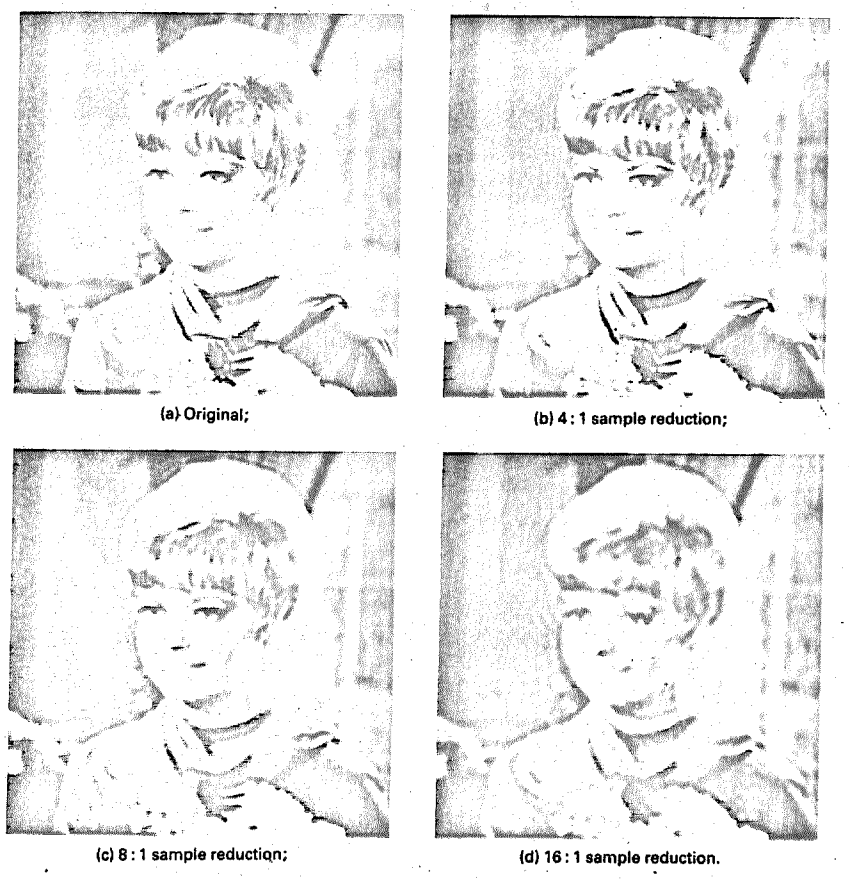


Figure 5.20 Zonal filters for 2:1, 4:1, 8:1, 16:1 sample reduction. White areas are passbands, dark areas are stopbands.

The image transform is filtered by a *zonal mask* (Fig. 5.20) such that only a fraction of the transform coefficients are retained and the remaining ones are set to zero. Define the normalized mean square error

$$J_s \triangleq \frac{\sum\sum_{k,l \in \text{stopband}} |\nu_{k,l}|^2}{\sum_{k,l=0}^{N-1}\sum |\nu_{k,l}|^2} = \frac{\text{energy in stopband}}{\text{total energy}}$$

Figure 5.21 shows an original image and the image obtained after cosine transform zonal filtering to achieve various sample reduction ratios. Figure 5.22 shows the zonal filtered images for different transforms at a 4:1 sample reduction ratio. Figure 5.23 shows the mean square error versus sample reduction ratio for different transforms. Again we find the cosine transform to have the best performance.



(a) Original; (b) 4:1 sample reduction; (c) 8:1 sample reduction; (d) 16:1 sample reduction.

Figure 5.21 Basis restriction zonal filtered images in cosine transform domain.

Figure 5.22 Basis restriction zonal filtering using different transforms with 4:1 sample reduction.

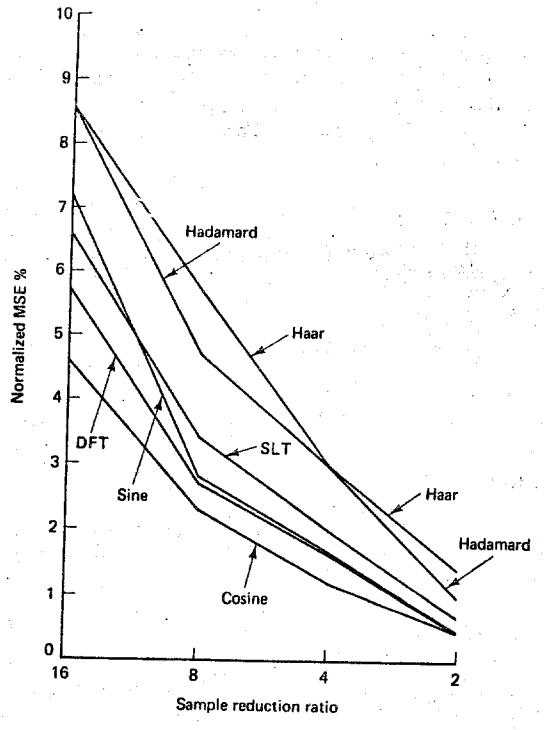


Figure 5.23 Performance comparison of different transforms with respect to basis restriction zonal filtering for 256×256 images.

5.12 A SINUSOIDAL FAMILY OF UNITARY TRANSFORMS

This is a class of complete orthonormal sets of eigenvectors generated by the parametric family of matrices whose structure is similar to that of $\mathbf{R}^{-1}$ [see (5.96)],

$$\mathbf{J} = \mathbf{J}(k_1, k_2, k_3) = \begin{bmatrix} 1-k_1\alpha & -\alpha & & & k_3\alpha \\ -\alpha & 1 & 0 & & \\ & 0 & 1 & -\alpha & \\ k_3\alpha & & -\alpha & 1-k_2\alpha & \end{bmatrix} \tag{5.180}$$

In fact, for $k_1 = k_2 = \rho$, $k_3 = 0$, $\beta^2 = (1-\rho^2)/(1+\rho^2)$, and $\alpha = \rho/(1+\rho^2)$, we have

$$\mathbf{J}(\rho, \rho, 0) = \beta^2 \mathbf{R}^{-1} \tag{5.181}$$

Since $\mathbf{R}$ and $\beta^2\mathbf{R}^{-1}$ have an identical set of eigenvectors, the KL transform associated with $\mathbf{R}$ can be determined from the eigenvectors of $\mathbf{J}(\rho, \rho, 0)$. Similarly, it

can be shown that the basis vectors of previously discussed cosine, sine, and discrete Fourier transforms are the eigenvectors of $\mathbf{J}(1,1,0)$, $\mathbf{J}(0,0,0)$, and $\mathbf{J}(1,1,-1)$, respectively. In fact, several other fast transforms whose basis vectors are sinusoids can be generated for different combinations of k_1, k_2, and k_3. For example, for $0 \le m, k \le N-1$, we obtain the following transforms:

1. Odd sine − 1: $k_1 = k_3 = 0, k_2 = 1$

$$\phi_m(k) = \frac{2}{\sqrt{2N+1}} \sin \frac{(k+1)(2m+1)\pi}{2N+1} \tag{5.182}$$

2. Odd cosine − 1: $k_1 = 1, k_2 = k_3 = 0$

$$\phi_m(k) = \frac{2}{\sqrt{2N+1}} \cos \frac{(2k+1)(2m+1)\pi}{2(2N+1)} \tag{5.183}$$

Other members of this family of transforms are given in [34].

Approximation to the KL Transform

The J matrices play a useful role in performance evaluation of the sinusoidal transforms. For example, two sinusoidal transforms can be compared with the KL transform by comparing corresponding J-matrix distances

$$\Delta(k_1, k_2, k_3) \triangleq \|\mathbf{J}(k_1, k_2, k_3) - \mathbf{J}(\rho, \rho, 0)\|^2 \tag{5.184}$$

This measure can also explain the close performance of the DCT and the KLT. Further, it can be shown that the DCT performs better than the sine transform for $0.5 \le \rho \le 1$ and the sine transform performs better than the cosine for other values of ρ. The J matrices are also useful in finding a fast sinusoidal transform approximation to the KL transform of an arbitrary random sequence whose covariance matrix is **A**. If **A** commutes with a J matrix, that is, **AJ** = **JA**, then they will have an identical set of eigenvectors. The best fast sinusoidal transform may be chosen as the one whose corresponding J matrix minimizes the *commuting distance* $\|\mathbf{AJ} - \mathbf{JA}\|^2$. Other uses of the J matrices are (1) finding fast algorithms for inversion of banded Toeplitz matrices, (2) efficient calculation of transform coefficient variances, which are needed in transform domain processing algorithms, and (3) establishing certain useful asymptotic properties of these transforms. For details see [34].

5.13 OUTER PRODUCT EXPANSION AND SINGULAR VALUE DECOMPOSITION

In the foregoing transform theory, we considered an $N \times M$ image **U** to be a vector in an NM-dimensional vector space. However, it is possible to represent any such image in an r-dimensional subspace where r is the rank of the matrix **U**.

Let the image be real and $M \le N$. The matrices $\mathbf{UU}^T$ and $\mathbf{U}^T\mathbf{U}$ are nonnegative, symmetric and have the identical eigenvalues, $\{\lambda_m\}$. Since $M \le N$, there are at most $r \le M$ nonzero eigenvalues. It is possible to find r orthogonal, $M \times 1$ eigenvectors $\{\phi_m\}$ of $\mathbf{U}^T\mathbf{U}$ and r orthogonal $N \times 1$ eigenvectors $\{\psi_m\}$ of $\mathbf{UU}^T$, that is,

$$\mathbf{U}^T\mathbf{U}\phi_m = \lambda_m \phi_m, \qquad m = 1, \ldots, r \tag{5.185}$$

$$\mathbf{UU}^T\psi_m = \lambda_m \psi_m, \qquad m = 1, \ldots, r \tag{5.186}$$

The matrix **U** has the representation

$$\mathbf{U} = \Psi \Lambda^{1/2} \Phi^T \tag{5.187}$$

$$= \sum_{m=1}^{r} \sqrt{\lambda_m}\, \psi_m \phi_m^T \tag{5.188}$$

where Ψ and Φ are $N \times r$ and $M \times r$ matrices whose mth columns are the vectors ψ_m and ϕ_m, respectively, and $\Lambda^{1/2}$ is an $r \times r$ diagonal matrix, defined as

$$\Lambda^{1/2} = \begin{bmatrix} \sqrt{\lambda_1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sqrt{\lambda_r} \end{bmatrix} \tag{5.189}$$

Equation (5.188) is called the *spectral representation*, the *outer product expansion*, or the *singular value decomposition* (SVD) of **U**. The nonzero eigenvalues (of $\mathbf{U}^T\mathbf{U}$), λ_m, are also called the singular values of **U**. If $r \ll M$, then the image containing NM samples can be represented by $(M+N)r$ samples of the vectors $\{\lambda_m^{1/4}\psi_m, \lambda_m^{1/4}\phi_m; m = 1, \ldots, r\}$.

Since Ψ and Φ have orthogonal columns, from (5.187) the *SVD transform* of the image **U** is defined as

$$\Lambda^{1/2} = \Psi^T \mathbf{U} \Phi \tag{5.190}$$

which is a separable transform that diagonalizes the given image. The proof of (5.188) is outlined in Problem 5.31.

Properties of the SVD Transform

1. Once $\phi_m, m = 1, \ldots, r$ are known, the eigenvectors ψ_m can be determined as

$$\psi_m \triangleq \frac{1}{\sqrt{\lambda_m}} \mathbf{U}\phi_m, \qquad m = 1, \ldots, r \tag{5.191}$$

It can be shown that ψ_m are orthonormal eigenvectors of $\mathbf{UU}^T$ if ϕ_m are the orthonormal eigenvectors of $\mathbf{U}^T\mathbf{U}$.

2. The SVD transform as defined by (5.190) is not a unitary transform. This is because Ψ and Φ are rectangular matrices. However, we can include in Φ and Ψ additional orthogonal eigenvectors ϕ_m and ψ_m, which satisfy $\mathbf{U}\phi_m = 0$, $m = r+1, \ldots, M$ and $\mathbf{U}^T\psi_m = 0, m = r+1, \ldots, N$ such that these matrices are unitary and the *unitary SVD transform* is

$$\begin{bmatrix} \Lambda^{1/2} \\ \mathbf{0} \end{bmatrix} = \Psi^T \mathbf{U} \Phi \tag{5.192}$$

3. The image $\mathbf{U}_k$ generated by the partial sum

$$\mathbf{U}_k \triangleq \sum_{m=1}^{k} \sqrt{\lambda_m}\,\psi_m \phi_m^T, \qquad k \le r \tag{5.193}$$

is the best least squares rank-k approximation of $\mathbf{U}$ if λ_m are in decreasing order of magnitude. For any $k \le r$, the least squares error

$$\epsilon_k^2 = \sum_{m=1}^{M}\sum_{n=1}^{N} |u(m,n) - u_k(m,n)|^2, \qquad k = 1,2,\ldots,r \tag{5.194}$$

reduces to

$$\epsilon_k^2 = \sum_{m=k+1}^{r} \lambda_m \tag{5.195}$$

Let $L \triangleq NM$. Note that we can always write a two-dimensional unitary transform representation as an outer product expansion in an L-dimensional space, namely,

$$\mathbf{U} = \sum_{l=1}^{L} w_l \mathbf{a}_l \mathbf{b}_l^T \tag{5.196}$$

where w_l are scalars and $\mathbf{a}_l$ and $\mathbf{b}_l$ are sequences of orthogonal basis vectors of dimensions $N \times 1$ and $M \times 1$, respectively. The least squares error between $\mathbf{U}$ and any partial sum

$$\hat{\mathbf{U}}_k \triangleq \sum_{l=1}^{k} w_l \mathbf{a}_l \mathbf{b}_l^T \tag{5.197}$$

is minimized for any $k \in [1, L]$ when the above expansion coincides with (5.193), that is, when $\hat{\mathbf{U}}_k = \mathbf{U}_k$.

This means *the energy concentrated in the transform coefficients $w_l, l = 1,\ldots,k$ is maximized by the SVD transform for the given image.* Recall that the KL transform, maximizes *the average energy in a given number of transform coefficients, the average being taken over the ensemble for which the autocorrelation function is defined.* Hence, on an image-to-image basis, the SVD transform will concentrate more energy in the same number of coefficients. But the SVD has to be calculated for *each* image. On the other hand the KL transform needs to be calculated only once for the whole image ensemble. Therefore, while one may be able to find a reasonable fast transform approximation of the KL transform, no such fast transform substitute for the SVD is expected to exist.

Although applicable in image restoration and image data compression problems, the usefulness of SVD in such image processing problems is severely limited because of large computational effort required for calculating the eigenvalues and eigenvectors of large image matrices. However, the SVD is a fundamental result in matrix theory that is useful in finding the generalized inverse of singular matrices and in the analysis of several image processing problems.

Example 5.11

Let

$$\mathbf{U} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 3 \end{bmatrix}$$

The eigenvalues of $\mathbf{U}^T\mathbf{U}$ are found to be $\lambda_1 = 18.06$, $\lambda_2 = 1.94$, which give $r = 2$, and the SVD transform of $\mathbf{U}$ is

$$\mathbf{\Lambda}^{1/2} = \begin{bmatrix} 4.25 & 0 \\ 0 & 1.39 \end{bmatrix}$$

The eigenvectors are found to be

$$\phi_1 = \begin{bmatrix} 0.5019 \\ 0.8649 \end{bmatrix}, \qquad \phi_2 = \begin{bmatrix} 0.8649 \\ -0.5019 \end{bmatrix}$$

(*continued on page 180*)

TABLE 5.3 Summary of Image Transforms

DFT/unitary DFT	Fast transform, most useful in digital signal processing, convolution, digital filtering, analysis of circulant and Toeplitz systems. Requires complex arithmetic. Has very good energy compaction for images.
Cosine	Fast transform, requires real operations, near optimal substitute for the KL transform of highly correlated images. Useful in designing transform coders and Wiener filters for images. Has excellent energy compaction for images.
Sine	About twice as fast as the fast cosine transform, symmetric, requires real operations; yields fast KL transform algorithm which yields recursive block processing algorithms, for coding, filtering, and so on; useful in estimating performance bounds of many image processing problems. Energy compaction for images is very good.
Hadamard	Faster than sinusoidal transforms, since no multiplications are required; useful in digital hardware implementations of image processing algorithms. Easy to simulate but difficult to analyze. Applications in image data compression, filtering, and design of codes. Has good energy compaction for images.
Haar	Very fast transform. Useful in feature extracton, image coding, and image analysis problems. Energy compaction is fair.
Slant	Fast transform. Has "image-like basis"; useful in image coding. Has very good energy compaction for images.
Karhunen-Loeve	Is optimal in many ways; has no fast algorithm; useful in performance evaluation and for finding performance bounds. Useful for small size vectors e.g., color multispectral or other feature vectors. Has the best energy compaction in the mean square sense over an ensemble.
Fast KL	Useful for designing fast, recursive-block processing techniques, including adaptive techniques. Its performance is better than independent block-by-block processing techniques.
Sinusoidal transforms	Many members have fast implementation, useful in finding practical substitutes for the KL transform, analysis of Toeplitz systems, mathematical modeling of signals. Energy compaction for the optimum-fast transform is excellent.
SVD transform	Best energy-packing efficiency for any given image. Varies drastically from image to image; has no fast algorithm or a reasonable fast transform substitute; useful in design of separable FIR filters, finding least squares and minimum norm solutions of linear equations, finding rank of large matrices, and so on. Potential image processing applications are in image restoration, power spectrum estimation and data compression.

From above ψ_1 is obtained via (5.191) to yield

$$\mathbf{U}_1 = \sqrt{\lambda_1}\,\psi_1\,\phi_1^T = \begin{bmatrix} 1.120 & 1.94 \\ 0.935 & 1.62 \\ 1.549 & 2.70 \end{bmatrix}$$

as the best least squares rank-1 approximation of **U**. Let us compare this with the two dimensional cosine transform **U**, which is given by

$$\mathbf{V} = \mathbf{C}_3\mathbf{U}\mathbf{C}_2^T = \frac{1}{\sqrt{12}}\begin{bmatrix} \sqrt{2} & \sqrt{2} & \sqrt{2} \\ \sqrt{3} & 0 & -\sqrt{3} \\ 1 & -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{\sqrt{12}}\begin{bmatrix} 10\sqrt{2} & -2\sqrt{2} \\ -\sqrt{3} & \sqrt{3} \\ -1 & -5 \end{bmatrix}$$

It is easy to see that $\sum\sum_{k,l} v^2(k, l) = \lambda_1 + \lambda_2$. The energy concentrated in the K samples of SVD, $\sum_{m=1}^{K}\lambda_m$, $K = 1, 2$, is greater than the energy concentrated in any K samples of the cosine transform coefficients (show!).

5.14 SUMMARY

In this chapter we have studied the theory of unitary transforms and their properties. Several unitary tranforms, DFT, cosine, sine, Hadamard, Haar, Slant, KL, sinusoidal family, fast KL, and SVD, were discussed. Table 5.3 summarizes the various transforms and their applications.

PROBLEMS

5.1 For given P, Q show that the error σ_ϵ^2 of (5.8) is minimized when the series coefficients $v(k, l)$ are given by (5.3). Also show that the basis images must form a complete set for σ_ϵ^2 to be zero for $P = Q = N$.

5.2 *(Fast transforms and Kronecker separability)* From (5.23) we see that the number of operations in implementing the matrix-vector product is reduced from $O(N^4)$ to $O(N^3)$ if $\mathcal{A}$ is a Kronecker product. Apply this idea inductively to show that if $\mathcal{A}$ is $M \times M$ and

$$\mathcal{A} = \mathbf{A}_1 \otimes \mathbf{A}_2 \otimes \cdots \otimes \mathbf{A}_m$$

where $\mathbf{A}_k$ is $n_k \times n_k$, $M = \prod_{k=1}^{m} n_k$, then the transformation of (5.23) can be implemented in $O(M\sum_{k=1}^{m} n_k)$, which equals $nM \log_n M$ if $n_k = n$. Many fast algorithms for unitary matrices can be given this interpretation which was suggested by Good [9]. Transforms possessing this property are sometimes called *Good transforms*.

5.3 For the 2×2 transform **A** and the image **U**

$$\mathbf{A} = \tfrac{1}{2}\begin{bmatrix} \sqrt{3} & 1 \\ -1 & \sqrt{3} \end{bmatrix}, \qquad \mathbf{U} = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$$

calculate the transformed image **V** and the basis images.

5.4 Consider the vector **x** and an orthogonal transform **A**

$$\mathbf{x} = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix}, \qquad \mathbf{A} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

Let $\mathbf{a}_0$ and $\mathbf{a}_1$ denote the columns of $\mathbf{A}^T$ (that is, the basis vectors of **A**). The transformation $\mathbf{y} = \mathbf{A}\mathbf{x}$ can be written as $y_0 = \mathbf{a}_0^T\mathbf{x}$, $y_1 = \mathbf{a}_1^T\mathbf{x}$. Represent the vector **x** in Cartesian coordinates on a plane. Show that the transform **A** is a rotation of the coordinates by θ and y_0 and y_1 are the projections of **x** in the new coordinate system (see Fig. P5.4).

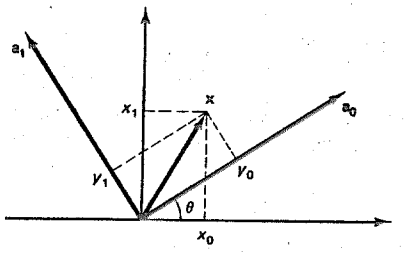


Figure P5.4

5.5 Prove that the magnitude of determinant of a unitary transform is unity. Also show that all the eigenvalues of a unitary matrix have unity magnitude.

5.6 Show that the entropy of an $N \times 1$ Gaussian random vector **u** with mean $\boldsymbol{\mu}$ and covariance $\mathbf{R}_u$ given by

$$H(\mathbf{u}) = \frac{N}{2}\log_2(2\pi e|\mathbf{R}_u|^{1/N})$$

is invariant under any unitary transformation.

5.7 Consider the zero mean random vector **u** with covariance $\mathbf{R}_u$ discussed in Example 5.2. From the class of unitary transforms

$$\mathbf{A}_\theta = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}, \qquad \mathbf{v} = \mathbf{A}_\theta\mathbf{u}$$

determine the value of θ for which (a) the average energy compressed in v_0 is maximum and (b) the components of **v** are uncorrelated.

5.8 Prove the two-dimensional energy conservation relation of (5.28).

5.9 *(DFT and circulant matrices)*

a. Show that (5.60) follows directly from (5.56) if ϕ_k is chosen as the kth column of the unitary DFT **F**. Now write (5.58) as a circulant matrix operation $\mathbf{x}_2 = \mathbf{H}\mathbf{x}_1$. Take unitary DFT of both sides and apply (5.60) to prove the circular convolution theorem, that is, (5.59).

b. Using (5.60) show that the inverse of an $N \times N$ circulant matrix can be obtained in $O(N \log N)$ operations via the FFT by calculating the elements of its first column as

$$[\mathbf{H}^{-1}]_{n,0} = \frac{1}{N}\sum_{k=0}^{N-1} W_N^{-kn}\lambda_k^{-1} = \text{inverse DFT of } \{\lambda_k^{-1}\}$$

c. Show that the $N \times 1$ vector $\mathbf{x}_2 = \mathbf{T}\mathbf{x}_1$, where $\mathbf{T}$ is $N \times N$ Toeplitz but not circulant, can be evaluated in $O(N \log N)$ operations via the FFT.

5.10 Show that the N^2 complex elements $v(k, l)$ of the unitary DFT of a real sequence $\{u(m, n), 0 \le m, n \le N-1\}$ can be determined from the knowledge of the partial sequence

$$\left\{v(k, 0), 0 \le k \le \frac{N}{2}\right\}, \left\{v\left(k, \frac{N}{2}\right), 0 \le k \le \frac{N}{2}\right\},$$

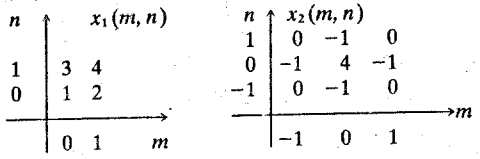
$$\left\{v(k, l), 0 \le k \le N-1, 1 \le l \le \frac{N}{2}-1\right\}, \qquad (N \text{ even})$$

which contains only N^2 nonzero real elements, in general.

5.11 **a.** Find the eigenvalues of the 2×2 doubly block circulant matrix

$$\mathcal{H} = \left[\begin{array}{cc|cc} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \\ \hline 3 & 4 & 1 & 2 \\ 4 & 3 & 2 & 1 \end{array}\right]$$

b. Given the arrays $x_1(m, n)$ and $x_2(m, n)$ as follows:



Write their convolution $x_3(m, n) = x_2(m, n) \circledast x_1(m, n)$ as a doubly block circulant matrix operating on a vector of size 16 and calculate the result. Verify your result by performing the convolution directly.

5.12 Show that if an image $\{u(m, n), 0 \le m, n \le N-1\}$ is multiplied by the checkerboard pattern $(-1)^{m+n}$, then its unitary DFT is centered at $(N/2, N/2)$. If the unitary DFT of $u(m, n)$ has its region of support as shown in Fig. P5.12, what would be the region of support of the unitary DFT of $(-1)^{m+n}u(m, n)$? Figure 5.6 shows the magnitude of the unitary DFTs of an image $u(m, n)$ and the image $(-1)^{m+n}u(m, n)$. This method can be used for computing the unitary DFT whose origin is at the center of the image matrix. The frequency increases as one moves away from the origin.

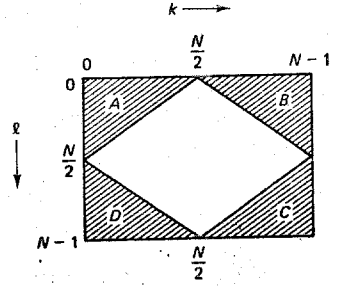


Figure P5.12

5.13 Show that the real and imaginary parts of the unitary DFT matrix are not orthogonal matrices in general.

5.14 Show that the $N \times N$ cosine transform matrix $\mathbf{C}$ is orthogonal. Verify your proof for the case $N = 4$.

5.15 Show that the $N \times N$ sine transform is orthogonal and is the eigenmatrix of $\mathbf{Q}$ given by (5.102). Verify your proof for the case $N = 3$.

5.16 Show that the cosine and sine transforms of an $N \times 1$ sequence $\{u(0), \ldots, u(N-1)\}$ can be calculated from the DFTs of the $2N \times 1$ symmetrically extended sequence $\{u(N-1), u(N-2), \ldots, u(1), u(0), u(0), u(1), \ldots, u(N-1)\}$ and of the $(2N+2) \times 1$ antisymmetrically extended sequence $\{0 - u(N-1), \ldots, -u(1), -u(0), 0, u(0), u(1), \ldots, u(N-1)\}$, respectively.

5.17 Suppose an $N \times N$ image $\mathbf{U}$ is mapped into a row-ordered $N^2 \times 1$ vector $\boldsymbol{u}$. Show that the $N^2 \times N^2$ one-dimensional Hadamard transform of $\boldsymbol{u}$ gives the $N \times N$ two-dimensional Hadamard transform of $\mathbf{U}$. Is this true for the other transforms discussed in the text? Give reasons.

5.18 Using the Kronecker product recursion (5.104), prove that a $2^n \times 2^n$ Hadamard transform is orthogonal.

5.19 Calculate and plot the energy packed in the first 1, 2, 4, 8, 16 sequency ordered samples of the Hadamard transform of a 16×1 vector whose autocorrelations are $r(k) = (0.95)^k$.

5.20 Prove that an $N \times N$ Haar transform matrix is orthogonal and can be implemented in $O(N)$ operations on an $N \times 1$ vector.

5.21 Using the recursive formula for generating the slant transforms prove that these matrices are orthogonal and fast.

5.22 If the KL transform of a zero mean $N \times 1$ vector $\mathbf{u}$ is $\boldsymbol{\Phi}$, then show that the KL transform of the sequence

$$\hat{u}(n) = u(n) + \mu, \qquad 0 \le n \le N-1$$

where μ is a constant, remains the same only if the vector $\mathbf{1} \triangleq (1, 1, \ldots, 1)^T$ is an eigenvector of the covariance matrix of $\mathbf{u}$. Which of the fast transforms discussed in the text satisfy this property?

5.23 If $\mathbf{u}_1$ and $\mathbf{u}_2$ are random vectors whose autocorrelation matrices commute, then show that they have a common KL transform. Hence, show that the KL transforms for autocorrelation matrices $\mathbf{R}$, $\mathbf{R}^{-1}$, and $f(\mathbf{R})$, where $f(\cdot)$ is an arbitrary function, are identical. What are the corresponding eigenvalues?

5.24* The autocorrelation array of a 4×1 zero mean vector $\mathbf{u}$ is given by $\{0.95^{|m-n|}, 0 \le m, n \le 3\}$.

a. What is the KL transform of $\mathbf{u}$?

b. Compare the basis vectors of the KL transform with the basis vectors of the 4×4 unitary DFT, DCT, DST, Hadamard, Haar, and Slant transforms.

c. Compare the performance of the various transforms by plotting the basis restriction error J_m versus m.

5.25* The autocorrelation function of a zero mean random field is given by (5.150), where $\rho = 0.95$. A 16×16 segment of this random field is unitarily transformed.

a. What is the maximum energy concentrated in 16, 32, 64, and 128 transform coefficients for each of the *seven* transforms, KL, cosine, sine, unitary DFT, Hadamard, Haar, and Slant?

b. Compare the performance of these transforms for this random field by plotting the mean square error for sample reduction ratios of 2, 4, 8, and 16. (*Hint:* Use Table 5.2.)

5.26 *(Threshold representation)* Referring to Fig. 5.16, where $u(n)$ is a Gaussian random sequence, the quantity

$$C_m = \frac{1}{N}\sum_{n=0}^{N-1}[u(n)-z(n)]^2 = \frac{1}{N}\sum_{n=0}^{N-1}\left[u(n)-\sum_{j=0}^{m-1}b(n,j)v(j)\right]^2$$

is a random variable with respect to m. Let m be such that

$$C_{m-1} > \epsilon^2, \qquad C_m \le \epsilon^2 \quad \text{for any fixed} \quad \epsilon^2 > 0$$

If $\mathbf{A}$ and $\mathbf{B}$ are restricted to be unitary transforms, then show that $E[m]$ is minimized when $\mathbf{A} = \mathbf{\Phi}^{*T}$, $\mathbf{B} = \mathbf{A}^{-1}$, where $\mathbf{\Phi}^{*T}$ is the KL transform of $u(n)$. For details see [33].

5.27 *(Minimum entropy property of the KL transform)* [30] Define an entropy in the $\mathbf{A}$-transform domain as

$$H[\mathbf{A}] = -\sum_{k=0}^{N-1}\sigma_k^2 \log_e \sigma_k^2$$

where σ_k^2 are the variances of the transformed variables $v(k)$. Show that among all the unitary transforms the KL transform minimizes this entropy, that is, $H[\mathbf{\Phi}^{*T}] \le H[\mathbf{A}]$.

5.28 **a.** Write the $N \times N$ covariance matrix $\mathbf{R}$ defined in (2.68) as

$$\beta^2\mathbf{R}^{-1} = \mathbf{J}(k_1, k_2, k_3) - \Delta\mathbf{J}$$

where $\Delta\mathbf{J}$ is a sparse $N \times N$ matrix with nonzero terms at the four corners. Show that the above relation yields

$$\mathbf{R} = \beta^2\mathbf{J}^{-1} + \beta^2\mathbf{J}^{-1}\Delta\mathbf{J}\mathbf{J}^{-1} + \mathbf{J}^{-1}(\Delta\mathbf{R})\mathbf{J}^{-1}$$

where $\Delta\mathbf{R} \triangleq \Delta\mathbf{J}\mathbf{R}\Delta\mathbf{J}$ is also a sparse matrix, which has at most four (corner) nonzero terms. If $\mathbf{\Phi}$ diagonalizes $\mathbf{J}$, then show that the variances of the transform coefficients are given by

$$\sigma_k^2 \triangleq [\mathbf{\Phi}^{*T}\mathbf{R}\mathbf{\Phi}]_{k,k} = \frac{\beta^2}{\lambda_k} + \frac{\beta^2}{\lambda_k^2}[\mathbf{\Phi}^{*T}\Delta\mathbf{J}\mathbf{\Phi}]_{k,k} + \frac{1}{\lambda_k^2}[\mathbf{\Phi}^{*T}\Delta\mathbf{R}\mathbf{\Phi}]_{k,k} \qquad \text{P5.28-1}$$

Now verify the formulas

$$\sigma_k^2(\text{DCT}) = \frac{\beta^2}{\lambda_k} - \frac{4(1-\rho)^2\alpha^2}{N\lambda_k^2\rho}[1-(-1)^k\rho^N]\left[\cos^2\left(\frac{k\pi}{2N}\right) - \frac{1}{2}\delta(k)\right] \qquad \text{P5.28-2}$$

where $\lambda_k = 1 - 2\alpha\cos k\pi/N$,

$$\sigma_k^2(\text{DST}) = \frac{\beta^2}{\lambda_k} + \frac{4\alpha^2}{(N+1)\lambda_k^2}[1+(-1)^k\rho^{N+1}]\sin^2\left(\frac{(k+1)\pi}{N+1}\right), \qquad 0 \le k \le N-1 \qquad \text{P5.28-3}$$

where $\lambda_k = 1 - 2\alpha\cos(k+1)\pi/(N+1)$, and

$$\sigma_k^2(\text{DFT}) = \frac{\beta^2}{\lambda_k} - \frac{2(1-\rho^N)\alpha[\cos(2\pi k/N) - 2\alpha]}{N\lambda_k^2} \qquad \text{P5.28-4}$$

where $\lambda_k = 1 - 2\alpha\cos 2\pi k/N$, $0 \le k \le N-1$.

b. Using the formulas P5.28-2–P5.28-4 and (5.120), calculate the fraction of energy packed in $N/2$ transform coefficients arranged in decreasing order by the cosine, sine, unitary DFT, and Hadamard transforms for $N = 4, 16, 64, 256, 1024$, and 4096 for a stationary Markov sequence whose autocorrelation matrix is given by $\mathbf{R} = \{\rho^{|m-n|}\}$, $\rho = 0.95$.

5.29 **a.** For an arbitrary real stationary sequence, its autocorrelation matrix, $\mathbf{R} \triangleq \{r(m-n)\}$, is Toeplitz. Show that $\mathbf{A}$-transform coefficient variances denoted by $\sigma_k^2(\mathbf{A})$, can be obtained in $O(N\log N)$ operations via the formulas

$$\sigma_k^2(\mathbf{F}) = \frac{1}{N}\sum_{n=-N+1}^{N-1}(N-|n|)r(n)W_N^{nk}, \qquad \mathbf{F} = \text{unitary DFT}$$

$$\sigma_k^2(\text{DST}) = r(0) + \frac{1}{N+1}\left[2a(k) + b(k)\cot\left(\frac{\pi(k+1)}{N+1}\right)\right],$$

$$a(k) + jb(k) \triangleq \sum_{n=1}^{N-1}[r(n) + r(-n)]\exp\left[\frac{j\pi n(k+1)}{N+1}\right]$$

where $0 \le k \le N-1$. Find a similar expression for the DCT.

b. In two dimensions, for stationary random fields, (5.36) implies we have to evaluate

$$\sigma_{k,l}^2(\mathbf{A}) = \sum_m\sum_n\sum_{m'}\sum_{n'}a(k,m)a^*(k,m')r(m-m',n-n')a(l,n)a^*(l,n')$$

Show that $\sigma_{k,l}^2(\mathbf{A})$ can be evaluated in $O(N^2\log N)$ operations, when $\mathbf{A}$ is the FFT, DST, or DCT.

5.30 Compare the maximum energy packed in k SVD transform coefficients for $k = 1, 2$, of the 2×4 image

$$\mathbf{U} = \begin{pmatrix} 1 & 2 & 5 & 6 \\ 3 & 4 & 7 & 8 \end{pmatrix}$$

with that packed by the cosine, unitary DFT, and Hadamard transforms.

5.31 *(Proof of SVD representation)* Define $\boldsymbol{\phi}_m$ such that $\mathbf{U}\boldsymbol{\phi}_m = \mathbf{0}$ for $m = r+1, \ldots, M$ so that the set $\boldsymbol{\phi}_m$, $1 \le m \le M$ is complete and orthonormal. Substituting for $\boldsymbol{\psi}_m$ from (5.191) in (5.188), obtain the following result:

$$\sum_{m=1}^{r}\sqrt{\lambda_m}\boldsymbol{\psi}_m\boldsymbol{\phi}_m^T = \mathbf{U}\left[\sum_{m=1}^{r}\boldsymbol{\phi}_m\boldsymbol{\phi}_m^T\right] = \mathbf{U}\left[\sum_{m=1}^{M}\boldsymbol{\phi}_m\boldsymbol{\phi}_m^T\right] = \mathbf{U}$$

BIBLIOGRAPHY

Sections 5.1, 5.2

General references on image transforms:

1. H. C. Andrews. *Computer Techniques in Image Processing*. New York: Academic Press, 1970, Chapters 5, 6.
2. H. C. Andrews. "Two Dimensional Transforms." in *Topics in Applied Physics: Picture Processing and Digital Filtering*, vol. 6, T. S. Huang (ed)., New York: Springer Verlag, 1975.
3. N. Ahmed and K. R. Rao. *Orthogonal Transforms for Digital Signal Processing*. New York: Springer Verlag, 1975.

4. W. K. Pratt. *Digital Image Processing*. New York: Wiley Interscience, 1978.
5. H. F. Harmuth. *Transmission of Information by Orthogonal Signals*. New York: Springer Verlag, 1970.
6. Proceedings Symposia on *Applications of Walsh Functions*, University of Maryland, *IEEE-EMC* (1970–73) and Cath. U. of Ameri., 1974.
7. D. F. Elliott and K. R. Rao. *Fast Transforms, Algorithms and Applications*, New York: Academic Press, 1983.

Section 5.3

For matrix theory description of unitary transforms:

8. R. Bellman. *Introduction to Matrix Analysis*, New York: McGraw-Hill, 1960.

Sections 5.4, 5.5

For DFT, FFT, and their applications:

9. I. J. Good. "The Interaction Algorithm and Practical Fourier Analysis." *J. Royal Stat. Soc.* (London) B20 (1958): 361.
10. J. W. Cooley and J. W. Tukey. "An Algorithm for the Machine Calculation of Complex Fourier Series," *Math. Comput.* 19, 90 (April 1965): 297–301.
11. *IEEE Trans. Audio and Electroacoustics*. Special Issue on the Fast Fourier Transform AU-15 (1967).
12. G. D. Bergland. "A Guided Tour of the Fast Fourier Transform." *IEEE Spectrum* 6 (July 1969): 41–52.
13. E. O. Brigham. *The Fast Fourier Transform*. Englewood Cliffs, N.J.: Prentice-Hall, 1974.
14. A. K. Jain. "Fast Inversion of Banded Toeplitz Matrices Via Circular Decomposition." *IEEE Trans. ASSP* ASSP-26, no. 2 (April 1978): 121–126.

Sections 5.6, 5.7

15. N. Ahmed, T. Natarajan, and K. R. Rao. "Discrete Cosine Transform." *IEEE Trans. on Computers* (correspondence) C-23 (January 1974): 90–93.
16. A. K. Jain. "A Fast Karhunen Loeve Transform for a Class of Random Processes." *IEEE Trans. Communications*, Vol. COM-24, pp. 1023–1029, Sept. 1976.
17. A. K. Jain. "Some New Techniques in Image Processing," *Proc. Symposium on Current Mathematical Problems in Image Science*, Monterey, California, November 10–12, 1976.
18. W. H. Chen, C. H. Smith, and S. C. Fralick. "A Fast Computational Algorithm for the Discrete Cosine Transform." *IEEE Trans. Commun.* COM-25 (September 1977): 1004–1009.
19. M. J. Narasimha and A. M. Peterson. "On the Computation of the Discrete Cosine Transform. *IEEE Trans. Commun.* COM-26, no. 6 (June 1978): 934–936.

20. P. Yip and K. R. Rao. "A Fast Computational Algorithm for the Discrete Sine Transform." *IEEE Trans. Commun.* COM-28, no. 2 (February 1980): 304–307.

Sections 5.8, 5.9, 5.10

For Walsh functions and Hadamard, Haar, and slant transforms, see [1–6] and:

21. J. L. Walsh. "A Closed Set of Orthogonal Functions." *American J. of Mathematics* 45 (1923): 5–24.
22. R. E. A. C. Paley. "A Remarkable Series of Orthogonal Functions." *Proc. London Math. Soc.* 34 (1932): 241–279.
23. H. Kitajima. "Energy Packing Efficiency of the Hadamard Transform." *IEEE Trans. Comm.* (correspondence) COM-24 (November 1976): 1256–1258.
24. J. E. Shore. "On the Applications of Haar Functions," *IEEE Trans. Communications* COM-21 (March 1973): 209–216.
25. W. K. Pratt, W. H. Chen, and L. R. Welch. "Slant Transform Image Coding." *IEEE Trans. Comm.* COM-22 (August 1974): 1075–1093. Also see W. H. Chen, "Slant Transform Image Coding." Ph.D. Thesis, University of Southern California, Los Angeles, California, 1973.

Section 5.11

For theory of KL transform and its historic development:

26. H. Hotelling. "Analysis of a Complex of Statistical Variables into Principle Components." *J. Educ. Psychology* 24 (1933): 417–441 and 498–520.
27. H. Karhunen. "Uber Lineare Methoden in der Wahrscheinlich-Keitsrechnung." *Ann. Acad. Science Fenn*, Ser. A.I. 37, Helsmki, 1947. (also see translation by I. Selin in the Rand Corp., Doc. T-131, August 11, 1960).
28. M. Loeve. "Fonctions Aleatoires de Seconde Ordre," in P. Levy, *Processus Stochastiques et Mouvement Brownien*. Paris, France: Hermann, 1948.
29. J. L. Brown, Jr. "Mean Square Truncation Error in Series Expansion of Random Functions," *J. SIAM* 8 (March 1960): 28–32.
30. S. Watanabe. "Karhunen-Loeve Expansion and Factor Analysis, Theoretical Remarks and Applications." *Trans. Fourth Prague Conf. Inform. Theory*, Statist. Decision Functions, and Random Processes, Prague, 1965, pp. 635–660.
31. H. P. Kramer and M. V. Mathews. "A Linear Coding for Transmitting a Set of Correlated Signals." *IRE Trans. Inform. Theory* IT-2 (September 1956): 41–46.
32. W. D. Ray and R. M. Driver. "Further Decomposition of the Karhunen-Loeve Series Representation of a Stationary Random Process." *IEEE Trans. Info. Theory* IT-11 (November 1970): 663–668.

For minimum mean square variance distribution and entropy properties we follow, primarily, [30] and [31]. Some other properties of the KL transform are discussed in: