Image classification by an optical implementation of the Fukunaga–Koontz transform

James R. Leger* and Sing H. Lee

Electrical Engineering and Computer Sciences, University of California, San Diego, La Jolla, California 92093

Received July 27, 1981

The Fukunaga–Koontz (F–K) transform is a linear transformation that performs image-feature extraction for a two-class image classification problem. It has the property that the most important basis functions for representing one class of image data (in a least-squares sense) are also the least important for representing a second image class. We present a new method of calculating the F–K basis functions for large dimensional imagery by using a small digital computer, when the intraclass variation can be approximated by correlation matrices of low rank. Having calculated the F–K basis functions, we use a coherent optical processor to obtain the coefficients of the F–K transform in parallel. Finally, these coefficients are detected electronically, and a classification is performed by the small digital computer.

1. INTRODUCTION

Image classification is often performed in two steps. First, a feature extractor is applied that retains only the data useful for classification purposes. Next, a classifier is applied to these features. Thus the classifier operates on a reduced number of data points and can be more sophisticated without resulting in excessive processing time.

Since the feature extractor is applied to the raw image, it must operate on a large number of points. Digital techniques are limited because of the time it takes to perform even the simplest operations on large images. Optical techniques, however, can process large numbers of image points simultaneously because of the parallel processing inherent in an optical system. One of the feature-extraction techniques that has been extensively employed is the optical Fourier power spectrum.[1] However, the Fourier power spectrum is not always well suited to identifying important features. For example, if the relative positions of objects in an image are important for classification, it is clear that Fourier techniques are not suitable because the Fourier power spectrum contains information only about the autocorrelation of the image [see Figs. 10(a) and 10(b) in Appendix A]. Fortunately, it has been shown that optical processors are also capable of performing general linear transformations[2] and therefore have the potential to perform a more powerful feature extraction on large images. In this paper, we describe the optical implementation of the Fukunaga–Koontz (F–K) transform[3] to extract features that are useful for separating two classes of images. In order to provide physical insight about the F–K transform, we derive the transform in Section 2, with special attention given to a new method of calculating the F–K basis function for the large images by small computers. In Section 3 the optical implementation of the F–K transform is described with experimental results. Finally, Appendix A contains some examples that use computer synthesized images to illustrate the effectiveness of the F–K transform.

2. DERIVATION OF THE TECHNIQUE

The F–K transform selects features that are useful for separating two classes by using basis functions with the property that the most important basis function for one class is also the least important for the other class. The F–K transform involves performing a Karhunen–Loève (K–L) transform, followed by a whitening transform, and finally another K–L transform. (This procedure is identical to a simultaneous diagonalization.[4])

We start out by representing an image of $(n \times n)$ pixels as an $N\ (=n^2)$ dimensional vector $\mathbf{X}$. The two classes of images are described by training sets, with the first training set consisting of M_1 sample images from class 1 and the second set consisting of M_2 sample images from class 2. Given $\mathbf{X}_j{}^{(i)}$ as the jth image vector from the training set of the ith class, we can define the training sets in matrix form as follows:

$$W_1 = \left[\left(\frac{P_1}{M_1}\right)^{1/2}\mathbf{X}_1{}^{(1)}\left(\frac{P_1}{M_1}\right)^{1/2}\mathbf{X}_2{}^{(1)}\left(\frac{P_1}{M_1}\right)^{1/2}\mathbf{X}_3{}^{(1)}\ldots\left(\frac{P_1}{M_1}\right)^{1/2}\mathbf{X}_{M_1}{}^{(1)}\right],$$

$$W_2 = \left[\left(\frac{P_2}{M_2}\right)^{1/2}\mathbf{X}_1{}^{(2)}\left(\frac{P_2}{M_2}\right)\mathbf{X}_2{}^{(2)}\left(\frac{P_2}{M_2}\right)^{1/2}\mathbf{X}_3{}^{(2)}\ldots\left(\frac{P_2}{M_2}\right)^{1/2}\mathbf{X}_{M_2}{}^{(2)}\right],$$

and

$$W_t = \left[\left(\frac{P_1}{M_1}\right)^{1/2}\mathbf{X}_1{}^{(1)}\left(\frac{P_1}{M_1}\right)^{1/2}\mathbf{X}_2{}^{(1)}\ldots\left(\frac{P_1}{M_1}\right)^{1/2}\mathbf{X}_{M_1}{}^{(1)}\quad\left(\frac{P_2}{M_2}\right)^{1/2}\mathbf{X}_1{}^{(2)}\left(\frac{P_2}{M_2}\right)^{1/2}\mathbf{X}_2{}^{(2)}\ldots\left(\frac{P_2}{M_2}\right)^{1/2}\mathbf{X}_{M_2}{}^{(2)}\right], \tag{1}$$

where W_1 is an $(N \times M_1)$ matrix, W_2 is an $(N \times M_2)$ matrix, W_t is an $[N \times (M_1 + M_2)]$ matrix containing sample images from both classes, and p_i is the *a priori* probability of the ith class. The sample correlation matrix of the whole process (computed with data from both classes) is then given by

$$W_tW_t^+ = W_1W_1^+ + W_2W_2^+, \tag{2}$$

since

$$W_iW_i^+ = \frac{p_i}{M_i}\sum_{j=1}^{M_i} \mathbf{X}_j^{(i)}\mathbf{X}_j^{(i)+}, \qquad i = 1, 2$$

is the sample correlation matrix for the ith class.

A. First K-L Transform

The first step in calculating the basis functions of the F–K transform is to diagonalize $W_tW_t^+$ by a linear transformation. This transformation eliminates all statistical correlation that exists between points in images from the original two image classes. Since W_t is an $[N \times (M_1 + M_2)]$ matrix, $W_tW_t^+$ is an $(N \times N)$ matrix, where we recall that N is the dimension of the image vector. Since images typically contain a large number of pixels (N), diagonalizing $W_tW_t^+$ is a numerically impossible task. However, the number of members necessary in the training sets $(M_1 + M_2)$ is governed by the variation within the two classes and can be small if the intraclass variation is small (see Section 2.G). In this case, we can first diagonalize the $[(M_1 + M_2) \times (M_1 + M_2)]$ matrix $W_t^+W_t$ by solving

$$(W_t^+W_t)E = E\Lambda \tag{3}$$

for the $(M_1 + M_2)$ eigenvectors E and eigenvalues Λ. Multiplying Eq. (3) from the left on both sides by W_t and regrouping then produces

$$(W_tW_t^+)(W_tE) = (W_tE)\Lambda. \tag{4}$$

Thus we see that diagonalizing the two-class correlation matrix $(W_tW_t^+)$ gives (W_tE) and Λ as the eigenvectors and eigenvalues. The difficult problem of diagonalizing the $(N \times N)$ matrix $(W_tW_t^+)$ directly has become the much easier problem of diagonalizing the $[(M_1 + M_2) \times (\mathrm{M}_1 + \mathrm{M}_2)]$ matrix $(W_t^+W_t)$ and calculating the product (W_tE). [In general, an $(N \times N)$ matrix contains N eigenvectors. However, the rank of W_t^+ is at most $(M_1 + M_2)$ since it is composed of at most $(M_1 + M_2)$ linearly independent equations. Thus the product of the two matrices $(W_tW_t^+)$ has a maximum rank of $(M_1 + M_2)$. The eigenvalues of the remaining eigenvectors are zero, and we need not calculate them.]

B. Whitening Transform

The eigenvectors contained in (W_tE) are orthogonal but are not normalized. They can be normalized by multiplying both sides of Eq. (4) from the right-hand side by $\Lambda^{(-1/2)}$ to produce

$$(W_tW_t^+)(W_tE\Lambda^{-1/2}) = (W_tE\Lambda^{-1/2})\Lambda. \tag{5}$$

We can show that $(W_tE\Lambda^{-1/2})$ is orthonormal with the help of Eq. (3):

$$\begin{aligned}(W_tE\Lambda^{-1/2})^+(W_tE\Lambda^{-1/2}) &= \Lambda^{-1/2}E^+(W_t^+W_tE)\Lambda^{-1/2}\\ &= \Lambda^{-1/2}E^+E\Lambda\Lambda^{-1/2} = \Lambda^{-1/2}\Lambda\Lambda^{-1/2} = I.\end{aligned} \tag{6}$$

Therefore we can multiply both sides of Eq. (5) from the left by $(W_tE\Lambda^{-1/2})^+$ to produce

$$(\Lambda^{-1/2}E^+W_t^+)(W_tW_t^+)(W_tE\Lambda^{-1/2}) = \Lambda. \tag{7}$$

We can now scale each eigenvector by dividing by the square root of its eigenvalue:

$$\Lambda^{-1}E^+W_t^+(W_tW_t^+)W_tE\Lambda^{-1} = I. \tag{8}$$

This procedure is known as a whitening transform since it results in the identity matrix, i.e., a correlation matrix that is identical to that obtained from uncorrelated white noise.

C. Second K-L Transform

We have just shown that application of the transformation $(\Lambda^{-1}E^+W_t^+)$ to the sum-correlation matrix $(W_tW_t^+)$ results in the identity matrix. However, when this same transformation is applied to the single-class-correlation matrix $(W_1W_1^+)$ or $(W_2W_2^+)$, the resulting matrices are not necessarily diagonal matrices. We therefore can calculate the eigenvectors of the transformed single-class-correlation matrices:

$$[\Lambda^{-1}E^+W_t^+(W_1W_1^+)W_tE\Lambda^{-1}]\Psi = \Psi\Gamma, \tag{9a}$$

$$[\Lambda^{-1}E^+W_t^+(W_2W_2^+)W_tE\Lambda^{-1}]\Theta = \Theta M. \tag{9b}$$

Although the matrices $(W_1W_1^+)$ and $(W_2W_2^+)$ are of dimension $(N \times N)$, the transformation $(\Lambda^{-1}E^+W_t^+)$ reduces the dimension of the entire quantity in brackets to $[(M_1 + M_2) \times (M_1 + M_2)]$. Thus the eigenvectors of this quantity (Ψ) are again numerically easy to calculate. After Ψ is found, we can calculate $\psi i^+\Lambda^{-1}\mathrm{E}^+\mathrm{W}_t^+$, where ψ_i are eigenvectors in Ψ. $(\psi_i^+\Lambda^{-1}E^+W_t^+)$ are called the basis functions of the F–K transform.

In summary, the procedure for finding the basis functions for the F–K transform are as follows:

(1) Given $\{\mathbf{X}_j^{(i)}\}$, determine W_1, W_2, W_t by using Eqs. (1) and (2).

(2) Find the eigenvectors E and eigenvalues Λ of $(W_t^+W_t)$ by using Eq. (3) and calculate $(\Lambda^{-1}E^+W_t^+)$.

(3) Find Ψ and Γ by using Eq. (9a), or

$$\Psi^+\Lambda^{-1}E^+W_t^+(W_1W_1^+)W_tE\Lambda^{-1}\Psi = \Gamma.$$

(4) The basis functions for the F–K transform are

$$(\psi_i^+\Lambda^{-1}E^+W_t^+),$$

where ψ_i are the eigenvectors in Ψ.

D. Properties of Ψ, Θ, Λ and M.

Solving Eq. (2) for $W_1W_1^+$ and substituting it into Eq. (9a) yields

$$\Lambda^{-1}E^+W_t^+(W_tW_t^+ - W_2W_2^+)W_tE\Lambda^{-1}\Psi = \Psi\Gamma \tag{10}$$

and finally

$$\Lambda^{-1}E^+W_t^+(W_2W_2^+)W_tE\Lambda^{-1}\Psi = \Psi(I - \Gamma). \tag{11}$$

By comparing Eqs. (9b) and (11), we see that

$$\Theta = \Psi \tag{12}$$

and

$$M = I - \Gamma. \tag{13}$$

Thus the eigenvectors of the transformed correlation matrices (transformed versions of $W_1W_1^+$ and $W_2W_2^+$) are identical, and their corresponding eigenvalues are related by the equation

$$\mu_i = 1 - \gamma_i, \tag{14}$$

where μ_i are the diagonal components of M and γ_i are the diagonal components of Γ. Since the eigenvalues of each transformed class must be nonnegative, we see from Eq. (14) that they all must occur in the range [0, 1]. Thus the eigenvectors with large eigenvalues applied to one class will have small eigenvalues when applied to the other class. The eigenvalue is a measure of the importance of its corresponding eigenvector in the linear expansion, since it gives the mean-square error incurred when the eigenvector of the expansion is left out. Thus, if an image vector $\mathbf{X}$ is expanded by using basis vectors given by $(\psi_i{}^+\Lambda^{-1}E^+W_t{}^+)$, where ψ_i are the eigenvectors in Ψ with eigenvalues γ_i closest to 1, the resulting coefficients will be most important for describing class 1 and least important for describing class 2. Conversely, basis vectors formed from using ψ_i's with eigenvalues γ_i closest to 0 (and hence μ_i closest to 1) will be most important for describing class 2 and least important for class 1. The class-separation power of a basis vector can be defined as $|\gamma_i - 0.5|$, where γ_i is the eigenvalue of the eigenvector in the basis vector $(\psi_i{}^+\Lambda^{-1} E^+W_t{}^+)$. The sign of $(\gamma_i - 0.5)$ indicates the class that is most important for the basis vector. Feature selection consists of projecting the image vector onto a few of these basis vectors with high separating power. Eigenvectors with eigenvalues close to 0.5 are the least important for class-separation purposes.

E. Restoring the Ordering of Basis Functions

A straightforward application of Eq. (9a) will provide $(M_1 + M_2)$ basis functions with equal separation power. Thus a selection of a subset of the basis functions with highest separation power is impossible to perform. It is important that we restore the ordering of the basis functions so that the dimensionality reduction resulting from using a subset of basis functions is optimized.

The initial equal separation power can be seen by observing the rank and dimension of the term in brackets in Eq. (9a). The dimension of the term is $(M_1 + M_2)$ since E and Λ are $[(M_1 + M_2) \times (M_1 + M_2)]$ matrices. Thus the eigenvector matrix Ψ will contain $(M_1 + M_2)$ eigenvectors and Γ will contain $(M_1 + M_2)$ eigenvalues. The rank of $(W_1W_1{}^+)$ is only M_1, however, since W_1 contains only M_1 linearly independent equations. This implies that only M_1 of the eigenvalues in Γ will be nonzero, leaving M_2 zero values. By Eq. (11), these M_2 zero eigenvalues from class 1 must have a value of one when applied to class 2, since they must satisfy Eq. (14). The rank of the class 2 correlation matrix $W_2W_2{}^+$ is M_2, however, implying that it has only M_2 nonzero eigenvalues. Since we have just shown that it must have M_2 unity eigenvalues, this implies that it contains M_2 values equal to 1 and M_1 values equal to zero. Hence all eigenvectors have an equal class separation power of 0.5.

We can restore the eigenvalue ranking by performing a dimensionality reduction on the E matrix in Eq. (3). By choosing the M_1 eigenvectors with the largest eigenvalues λ_i, we construct matrices E' and Λ' of dimensions $[(M_1 + M_2) \times M_1]$ and $[M_1 \times M_1]$, respectively. When these are substituted for E and Λ in Eq. (9a), the dimension of the term in brackets is reduced to $(M_1 \times M_1)$, and it becomes full rank. This amounts to adding a dimensionality reduction step between steps (2) and (3) in the summary provided in Section 2.C.

F. Modified F-K Transform

One drawback of using correlation matrices to define the two classes is that the information contained in the difference of the mean vectors is not fully exploited. A modification to the F–K transform has been proposed to enhance the mean-difference information.[5] It consists of subtracting off the mean of *one* of the classes from both training sets. Thus one of the classes becomes zero mean and its correlation matrix becomes a covariance matrix, whereas the mean of the other class now represents the mean difference between the two original classes.

G. Determining the Size of the Training Set

The method described in Sections 2.A–2.C. allows one to calculate the F–K basis vectors by using large dimensional images if the number of members in the training set is small. This requirement implies that the intraclass variation must be fairly small so the sample-correlation matrices accurately represent the real correlation functions of the classes.[4]

One method for determining whether the training set is large enough is to obtain eigenvectors and eigenvalues for $2M$ training samples. If the summation of any M of the $2M$ eigenvalues is small compared with the summation of all $2M$ eigenvalues, then M samples is sufficient. Once it is determined that there are M dominant eigenvectors in a particular distribution, the accuracy of these M eigenvectors (that is, how closely they approximate the first M eigenvectors of a large training set) can be improved by calculating the eigenvectors several times from different M-member training sets. An arithmetic mean of the individual estimates of eigenvectors and eigenvalues results in estimates with improved accuracy.

Another more direct method of determining training set size is to choose a size M for a specific problem and generate the F–K basis vectors. Then, with a suitable classifier, test the ability of the system to classify both members of the training set and new members from outside the training set. If the system can accurately classify members of the training set, yet fails to classify new members correctly, the training set is not providing an accurate description of the statistical class and should be expanded.

3. OPTICAL IMPLEMENTATION OF THE F–K TRANSFORM

A. Method for Performing Transforms

Since the F–K transform is a linear transformation, the coefficient corresponding to a specific basis function is found by calculating the inner product between an input image function and the complex conjugate of the basis function. A coherent optical processor can be designed to calculate these coefficients in parallel by multiplying the input image by a coded-phase function $\exp[i\phi_r(x, y)]$ and designing a filter whose impulse response $h^*(-x, -y)$ consists of the complex conjugate of a summation of shifted products of the coded-phase function and a particular basis function:

$$h^*(-x,-y) = \sum_{pq}^{n} f_{pq}{}^*(x + p\Delta, y + q\Delta) \times \exp[-i\phi_r(x + p\Delta, y + q\Delta)], \quad (15)$$

where $f_{pq}(x, y)$ is the p,qth basis function generated from displaying the eigenvector $(\psi_i{}^+\Lambda^{-1}E^+W_t{}^+)$ in a two-dimensional format, $\phi_r(x, y)$ is a random function uniformly distributed from 0 to 2π, and Δ is a shift constant (see Ref. 2 for more details of the coded-phase optical processor). Since the coded-phase function has an autocorrelation of a delta function, a coherent optical correlator produces an output consisting of the transform coefficients spaced at intervals of Δ:

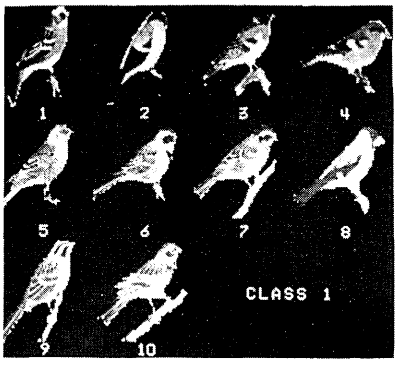
$$\begin{aligned} C(x', y') = & \iint g(x, y) \exp[i\phi_r(x, y)] \\ & \times \sum_{p,q}^{M} f_{pq}{}^*(x - x' + p\Delta, y - y' + q\Delta) \\ & \times \exp[-i\phi_r(x - x' + p\Delta, y - y' + q\Delta)]\mathrm{d}x\mathrm{d}y \\ = & \sum_{p,q}^{M} \left[\iint g(x, y) f_{pq}{}^*(x, y)\mathrm{d}x\mathrm{d}y\right] \\ & \times \delta(x' - p\Delta, y' - q\Delta). \end{aligned} \tag{16}$$

Both the coded-phase distribution that illuminates the input and the spatial filter can be obtained by computer-generated holograms. These computer holograms need be generated only once; they are used over and over again in the coded-phase processor with new test images to produce their F–K coefficients for classifications.

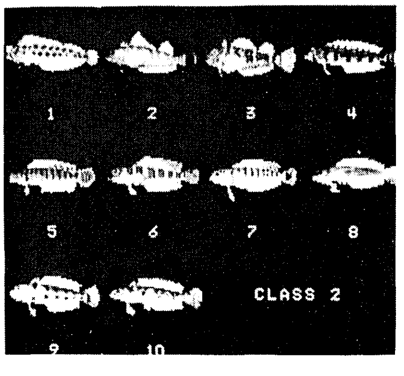
B. Experimental Results

The optical F–K transformation was applied to the problem of distinguishing birds from fish. In this experiment, the correlation matrices were used directly without subtracting the mean image of one of the classes. Ten images of song birds were input to the computer through a television digitizer system. These images formed the training set for class 1. Ten images of fish were also input to the computer to form the training set for class 2. These two sets are shown in Figs. 1(a) and 1(b).

The F–K basis functions are shown in Fig. 2(a) along with the test images consisting of five new birds and five new fish in Fig. 2(b). Since the basis functions can contain negative values, the pictures have been scaled so that black equals the most negative and white the most positive. The gray level that corresponds to a value of zero is shown in the small square below each basis function. The eigenvalues corresponding to the ten basis functions are given in Table 1. We see that

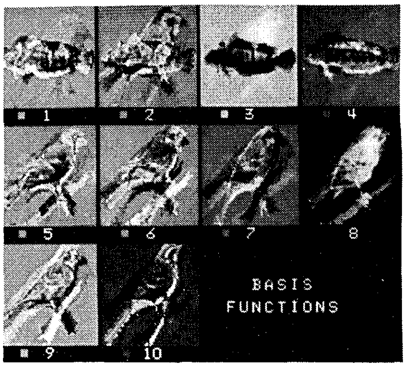


(a)

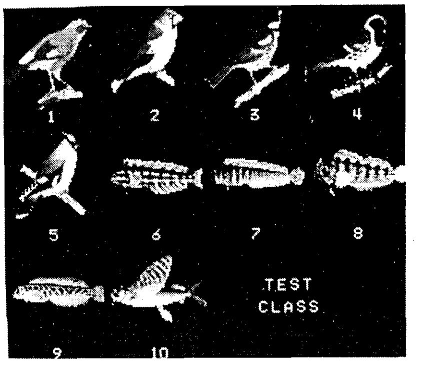


(b)

Fig. 1. (a) Class 1 training set consisting of 10 song birds; (b) class 2 training set consisting of 10 fish.



(a)



(b)

Fig. 2. (a) F–K basis functions. Basis function 3 is best for class 2 (fish), and basis function 8 is best for class 1 (birds). Small square to the left of image number indicates gray level corresponding to zero. (b) Test images consisting of 5 new birds and 5 new fish.

Table 1. Eigenvalues Corresponding to Ten Basis Functions

Basis Function	Eigenvalue (γ_i)	$\lvert\gamma_i - 0.5\rvert$	Best for Which Class?
1	0.1354	0.3646	fish
2	0.4846	0.0154	fish
3	0.0008	0.4992	fish
4	0.0034	0.4966	fish
5	0.9416	0.4416	birds
6	0.9883	0.4883	birds
7	0.9966	0.4966	birds
8	0.9998	0.4998	birds
9	0.9992	0.4992	birds
10	0.9985	0.4985	birds

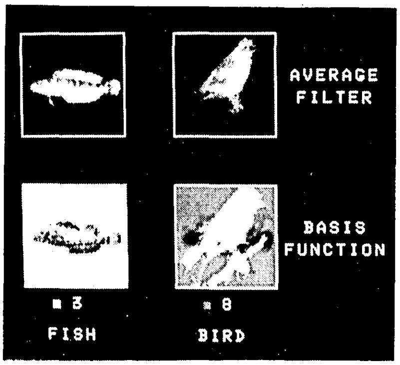


Fig. 3. Comparisons of basis functions 3 (fish) and 8 (birds) with filters and formed by the arithmetic average of the training sets.

the best basis function with bird-type features is number 8. The best basis function with fish-type features is number 3.

It is interesting to compare the third and eighth basis functions with the arithmetic average of each training set. Figure 3 shows that the most important F–K basis function for a class is similar, but not identical, to the average filter for that class. This is expected since both means were retained in the training sets. However, the F–K basis functions are not all positive. The gray level corresponding to zero is shown in the small square under each basis function.

A filter was generated that contained six of the basis functions shown in Fig. 2(a) in phase-coded form. Basis functions numbered 3, 4, and 1 were chosen to represent fishlike features and 8, 9, and 10 to represent birdlike features. This filter was placed in the filter plane of an optical correlator, and a computer hologram of the coded-phase array was placed in front of the input plane so that its reconstruction illuminated the input. This is illustrated in Fig. 4. The output was detected with a television camera, digitized, and displayed on a television monitor. The digital computer can measure these six coefficients and use them as input data to a nonlinear classifier for best classification.

The results of using five of the ten test images in the coded-phase processor are shown in Figs. 5(a)–5(e). The lower-right-hand corner of the television monitor contains a sampled version of the actual light field detected by the television camera at the output of the coded-phase processor. It consists of six points of light, where the square root of the intensity of each point corresponds to the absolute value of the coefficient for a specific basis function. The computer has added the lines around these points and the corresponding basis function numbers to help the display to be more meaningful. The measured values of the coefficients are plotted by the computer in the lower-left-hand part of the screen. The input image has been reproduced in the upper-left-hand part of the screen.

A simple linear classifier based on the coefficients from basis functions 3 and 8 was performed by the digital computer. This is drawn by the computer in the upper-right-hand corner of the screen. The magnitudes of the two coefficients define a point in the two-dimensional space that is located by the crosses. Points that are located below and to the right of the dotted diagonal line are classified as birds, whereas points above and to the left of the line are classified as fish. It is clear from the figures that, in each case, the combination of an F–K transform and a simple linear classifier leads to a correct classification of the input image. Figure 6 shows the combined result of 30 classifications, based on the magnitude of coefficients from basis functions 3 and 8. Twenty of the points are from the training set of Fig. 1, and ten are from the test set of Fig. 2(b). A linear decision surface is able to separate the class *birds* from the class *fish* with no errors.

Variations in orientation and scale can be accounted for by including rotated and scaled versions of the same scene in the training class. By choosing a suitable range of orientations and sizes in the training set, we can incorporate the *a priori* knowledge of most likely orientation and size to optimize the transform for a specific problem. In the previous example, we have the *a priori* knowledge that fish usually swim horizontally and almost never are oriented straight up and down. This information can be included in the training set for fish by having most of the images of fish be nearly horizontal.

4. CONCLUSION

We have shown that the Fukunaga–Koontz transform can be used as a feature extractor in a two-class classification application. A method was developed that simplified the computation of the F–K basis functions for large dimensional imagery and was found to work well when the intraclass variation in each class was small, and the correlation could be approximated by a sample correlation matrix of low rank.

The F–K transform was used as a feature extractor for classifying birds and fish. After the F–K functions were calculated, those most useful for classification were incorporated into a computer-generated hologram. A coherent optical processor was designed by using this computer-generated hologram to perform the F–K transform in real time. The output of the optical processor, consisting of the squared magnitude of the F–K coefficients, was detected by a television camera, digitized, and fed into a digital computer for classification. A simple linear classifier based on only two F–K coefficients was able to separate the images into two classes, indicating that the F–K transform had chosen good features.

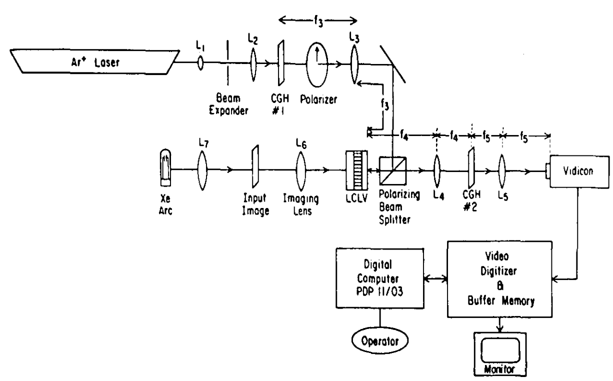
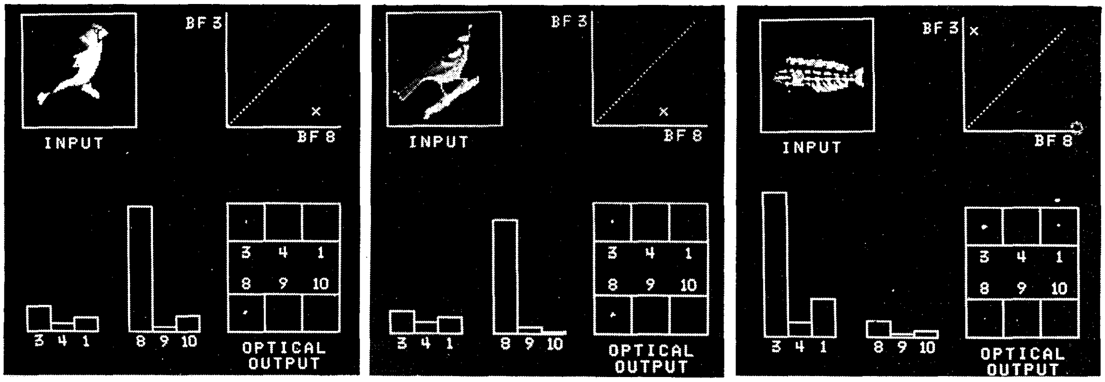
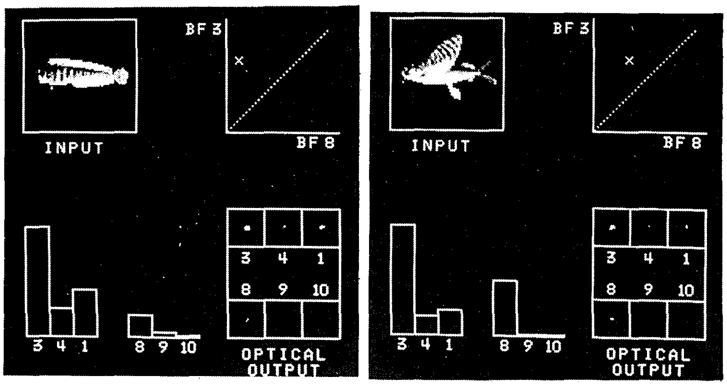


Fig. 4. Hybrid implementation of the coded-phase optical processor. A computer hologram of a coded-phase array is shown as CGH #1. A second computer hologram containing the F–K basis functions in coded-phase form is shown as CGH #2. LCLV is a liquid crystal light valve for converting an incoherent (test) image into a coherent image. The resultant F–K coefficients are detected by the vidicon and analyzed by a digital computer. With the same CGH #2 but new test images, new F–K coefficients are obtained and new classifications are achieved in real time.



(a) (b) (c)



(d) (e)

Fig. 5. Optical implementation of the F–K transform by using the coded-phase optical processor. The six basis functions with greatest separation power were used. Basis functions 3, 4, and 1 are best for fish, and 8, 9, and 10 are best for birds. Five of the ten images are used as inputs (a)–(e), in which the input is reproduced in the upper-left-hand corner of the television screen. The output of the optical processor is detected by a television camera, displayed in the lower-right-hand corner of the screen, measured by the video digitizer, and graphed in the lower left. A linear classifier using basis functions 3 and 8 is shown in the upper-right-hand corner of the screen, with the dotted line separating birds (below and to the right of the line) from fish (above and to the left of the line).

APPENDIX A: DIGITAL IMPLEMENTATION OF THE F-K TRANSFORM

In order to study the properties of the Fukunaga–Koontz transform for image classification, three experiments were performed using a digital computer. Two-dimensional images were generated by applying a two-dimensional set of random numbers to a digital linear filter. To obtain a sample image from a specific class, a 64 × 64 set of random numbers uniformly distributed between 0 and 255 was Fourier transformed, multiplied by the transfer function corresponding to the class, and inverse Fourier transformed. The training sets consisted of ten computer-generated images for each class. By varying the transfer function, the intraclasss variation and interclass separation could be controlled. Ten basis funtions for separating two classes were generated by using the method presented in Section 2, in which the mean image of the first training set was subtracted from images in both training sets. The members of the training set were then transformed using these ten basis functions.

A simple linear classifier based on the magnitude of the coefficients of two of the basis functions was again used for classification. The two basis functions were chosen so that the eigenvalue of one was closest to 1 while the eigenvalue of the other was closest to zero. These two coefficients were displayed on a two-dimensional graph, with the axes corresponding to the magnitudes of the coefficients of the two basis functions (basis vectors). A decision boundary was chosen to separate the two classes best in this two-dimensional space. New images (not members of the training set) were then classified by applying the transformation and observing on which side of the decision boundary they occurred.

In the first experiment, class 1 was defined by a low-pass transfer function with a cutoff $(f_x)_c = 4/L$ and $(f_y)_c = 4/L$, where L is the image size in both the x and y directions. Class 2 was defined by a high-pass transfer function, where $(f_x)_c = 28/L$ and $(f_y)_c = 28/L$. The maximum spatial frequency that

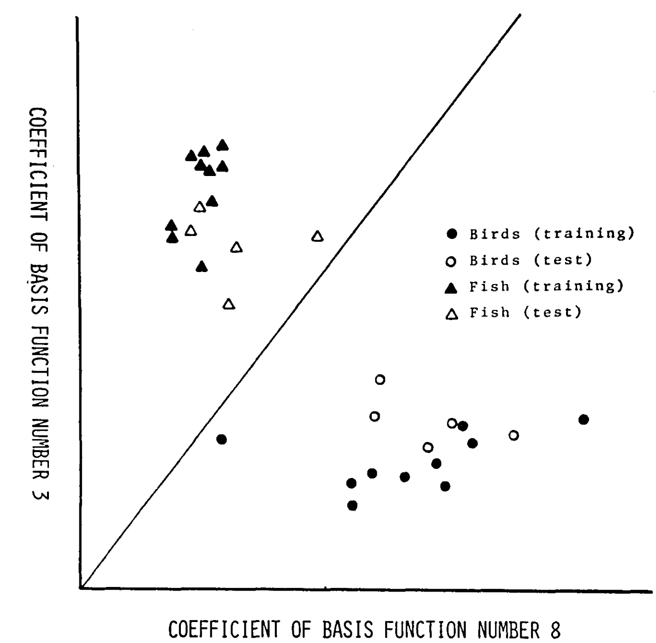
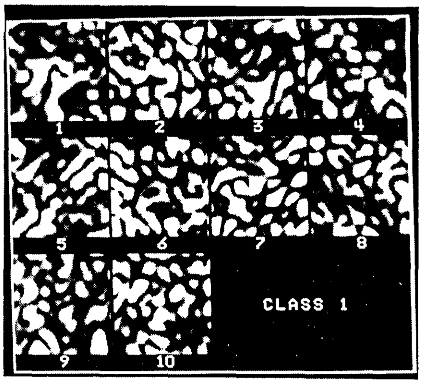
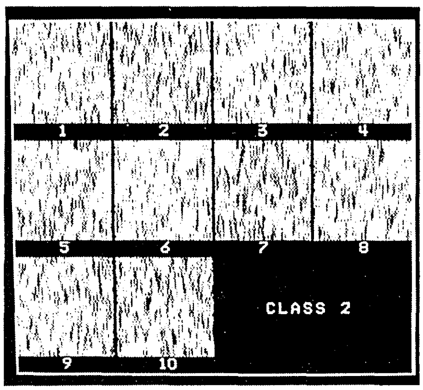


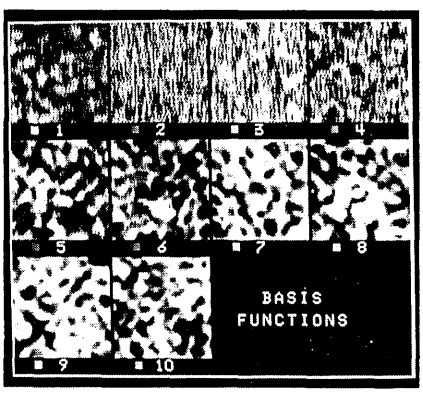
Fig. 6. Classification of birds and fish by using coefficients of basis functions 3 and 8. Ten members of each training set, as well as ten members of the test set, were classified.



(a)



(b)



(c)

Fig. 7. (a) Class 1 training set consisting of random numbers run through a low-pass filter. (b) Class 2 training set, in which a high-pass filter was used. (c) F–K basis functions. Basis function 1 is best for class 2 images, and basis function 9 is best for class 1 images.

can be represented by a 64 × 64 point image is $32/L$ in both the x and y directions. The training sets for the two classes and the F–K basis functions are shown in Figs. 7(a)–7(c). By examining the eigenvalues, we can determine that the first basis function contains the most important features of class 2 and the least important features of class 1. Basis function 9 contains the most important features of class 1 and the least important features of class 2. Figure 8 displays the coefficients of these two basis functions for members of the training class, as well as new test members. The solid line separates the two classes in this two-dimensional space, and the classifier (which simply assigns the class according to which side of the line the point occurs on) is seen to perform the task with

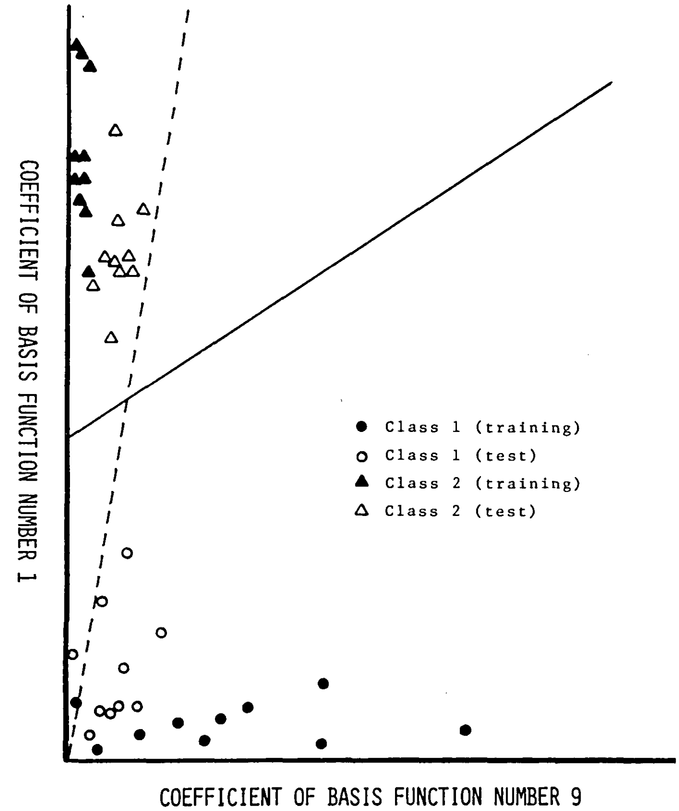


Fig. 8. Classification using coefficients of basis functions 1 and 9. Member of the training sets, as well as new class members, were transformed and plotted. The dotted line represents a classification boundary based on the ratio of the two coefficients. The solid line represents a general linear classification boundary.

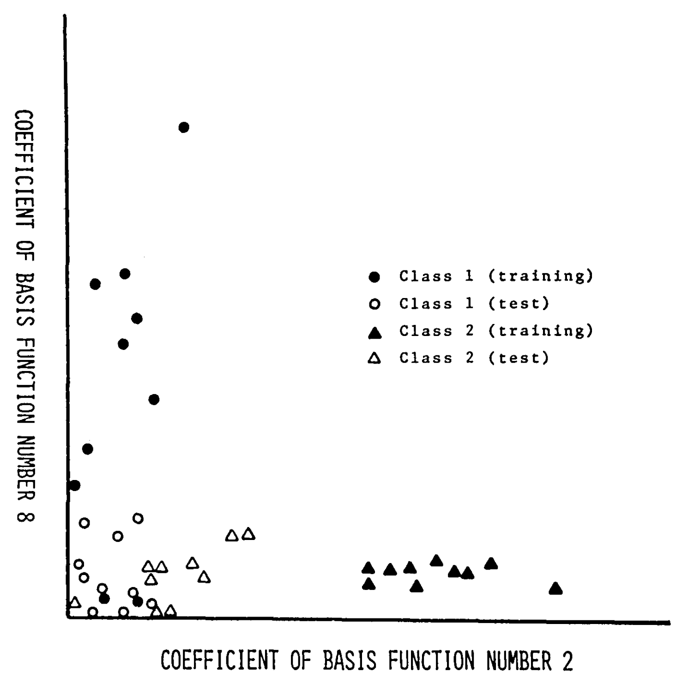
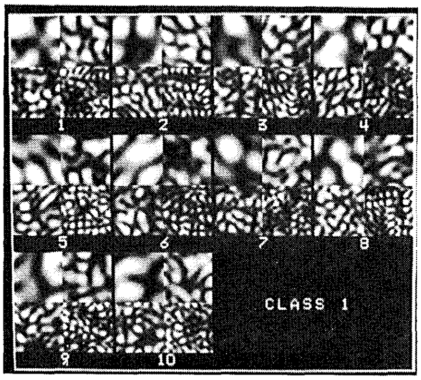


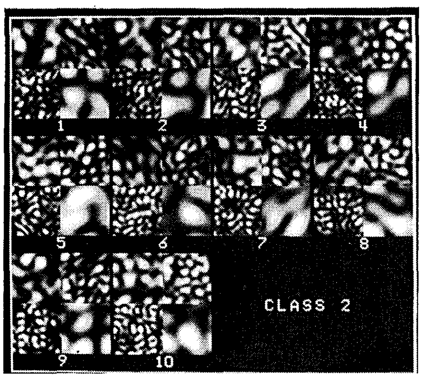
Fig. 9. Failure of classifier resulting from excessive intraclass variation. Members of the training set are linearly separable, but test members are not.

Table 2. Frequency Cutoffs of the Four Filters

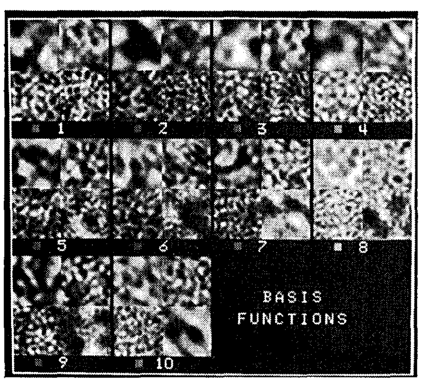
	f_x Coordinate		f_y Coordinate	
Filter	Low-Frequency Cutoff ($1/L$)	High-Frequency Cutoff ($1/L$)	Low-Frequency Cutoff ($1/L$)	High-Frequency Cutoff ($1/L$)
1	0	2	0	2
2	2	4	2	4
3	4	6	4	6
4	6	8	6	8



(a)



(b)



(c)

Fig. 10. (a) Class 1 training set. Four bandpass filters were applied to different corners of a two-dimensional array of random numbers. (b) Class 2 training set. The same bandpass filters were used but were applied to different corners of the random array. (c) F–K basis functions. Basis function 2 is best for class 1, and basis function 8 is best for class 2.

no errors. A slightly more restrictive classifier can be defined as the ratio of the two transform coefficients, restricting the classification boundary to pass through the origin. This classifier has the advantage that it is insensitive to a scaler multiplication of the original input data (which may result, for example, if the illumination of the scene changes).

If the intraclass variation is increased without increasing the number of training images, we expect to reach a point at which the sample correlation matrix is not an accurate estimate of the true correlation function of a class. This effect is seen in another experiment summarized in Fig. 9. Two low-pass filters were used for each class. The cutoff frequencies were $(f_x)_c = 4/L$, $(f_y)_c = 4/L$ for class 1 and $(f_x)_c = 10/L$, $(f_y)_c = 10/L$ for class 2, respectively. It is apparent that the training images are well separated and can easily be classified with no errors. However, test images perform poorly. Thus it appears that the variation within the classes is too great, and the chances of presenting the classifier with an image that is unlike any from the training set are high.

In the previous two examples, dimensionality reduction could have been performed by using specific coefficients of the Fourier transform. In the next experiment, the two classes were designed to contain the same general spatial frequencies. Instead of applying a single filter to the entire image, four filters with bandpass transfer functions given in Table 2 were applied, each to a different quarter of the image. Class 1 was formed by applying filter 1 to the upper-left part of the image, filter 2 to the upper right, filter 3 to the lower left, and filter 4 to the lower right. Class 2 was formed by applying filter 1 to the lower-right part of the image, filter 2 to the upper left, filter 3 to the upper right, and filter 4 to the lower left. Figure 10(a) and (b) shows the training sets of the two classes. Their classification by using Fourier transforms is expected to be unsatisfactory. The F–K basis functions are shown in Fig. 10(c). Basis functions 2 and 8 are the most important for class 1 and class 2, respectively. Figure 11 shows a plot of the coefficients of these two basis functions for members of the training set and new test images. The training set can be linearly separated with no errors. The test set classifies all

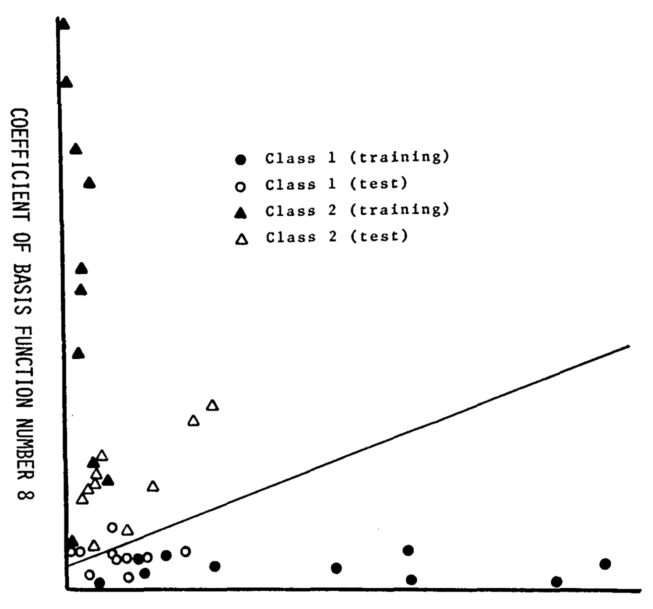


Fig. 11. Classification using coefficients of basis functions 2 and 8.

elements of class 2 correctly and misclassifies 3 of the 10 elements of class 1. Thus a total of 3 errors was made using 20 test images (and 40 images overall).

ACKNOWLEDGMENTS

J. Leger would like to thank Jacques Duvernoy for many helpful discussions. The research was supported by the National Science Foundation under grant no. ENG78-07462.

* Present address, The 3M Company, 518-1 3M Center, St. Paul, Minnesota 55144.

REFERENCES

1. J. R. Leger and S. H. Lee, "Signal processing using hybrid systems," in *Applications of the Optical Fourier Transform*, H. Stark, ed. (Academic, New York, 1982).
2. J. R. Leger and S. H. Lee, "Coherent optical implementation of generalized two-dimensional transforms," Opt. Eng. **18**, 518–523 (1979).
3. K. Fukunaga and W. L. G. Koontz, "Application of the Karhunen–Loève expansion to feature selection and ordering," IEEE Trans. Comput. **C-19**, 311–318 (1970).
4. K. Fukunaga, *Introduction to Statistical Pattern Recognition* (Academic, New York, 1972).
5. P. Sanyal and D. H. Foley, "Feature selection by a modified Fukunaga–Koontz transform and its graphical interpretation," presented at the Milwaukee Symposium on Automatic Computation and Control, April 24, 1976.