OPTICAL RECOGNITION OF STATISTICAL PATTERNS

Sing H. Lee
Electrical Engineering & Computer Sciences
University of California, San Diego 92093

SUMMARY

This paper describes the optical implementation of the Fukunaga-Koontz transform (FKT) and the Least-Squares Linear Mapping Technique (LSLMT). The FKT is a linear transformation which 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. The LSLMT is useful for performing a transform from large dimensional feature space to small dimensional decision space for separating multiple image classes by maximizing the interclass differences while minimizing the intraclass variations. The FKT and the LSLMT were optically implemented by utilizing a coded phase optical processor. Good experimental results were obtained, and they were compared with the performance of the matched filter and the average filter.

I. INTRODUCTION

Optical matched filtering, which is invariant to the translation of input, has been the main basis for optical pattern recognition for many years (refs. 1-3). Optical Mellin transform was later introduced to obtain scale invariant correlation (ref. 4). Currently, the subject of performing statistical pattern recognition and classification optically is found to be of considerable interest (refs. 5-9). For the two-class problem, we shall show in this paper that the Fukunaga-Koontz transform (FKT) can be implemented optically (ref. 9). For the K-class problem where K is greater than two, the optical implementation of the Least-Squares Linear Mapping Technique (LSLMT) will be presented (ref. 10).

Figure 1 shows the hybrid system for optical implementation of FKT and LSLMT. The key element in the hybrid system is the computer generated spatial filter. The filter is synthesized from K training sets of images using the FKT or LSLMT algorithm. When the LSLMT filter is inserted in the filter plane and a test image in the input plane of a coherent optical processor (COP), the output from the COP will be a bright spot of light in one of K predetermined locations, provided that the test image belongs to one of the K image classes in the statistical sense and it is illuminated with a random phase wavefront. As fast as one can input a new test image through a real-time interface device to the COP, a new bright spot of light indicating classification will appear among the K predetermined locations in the output. The COP effectively performs a matrix-vector multiplication in real time, where the matrix is $(K \times N)$, the vector is $(N \times 1)$, $N(= n \times n)$ is the space-bandwidth product or the number of pixels in an image, and K is the number of image classes.

When the FKT filter (instead of the LSLMT filter) is used, the output from the COP will contain a number of light spots whose intensities correspond to the squares of FKT coefficients. The FKT coefficients are associated with basis functions (features) which possess the interesting property that the most important basis

function for one class is also the least important basis function of the other class. Image classification can become as simple as comparing two FKT coefficients associated with the two basis functions of high separating power.

II. THE FUKUNAGA-KOONTZ TRANSFORM

A. Procedures for Calculating the Basis Function

The procedures for calculating the basis functions of the F-K transform are summarized as follows. (Details can be found in ref. 9.)

(a) Given a training set of images $\{\underline{X}_j^{(i)}\}$ consisting of M_1 image samples from class 1 ($i = 1$; $j = 1, \ldots, M_1$) and M_2 samples from class 2 ($i = 2$; $j = 1, \ldots, M_2$), we first represent each sample image as a column vector of length N, and define matrices W_1, W_2 and W_t as follows:

$$W_1 = \left[\underline{X}_1^{(1)}\ \underline{X}_2^{(1)}\ \underline{X}_3^{(1)}\ \cdots\ \underline{X}_{M_1}^{(1)}\right], \tag{1}$$

$$W_2 = \left[\underline{X}_1^{(2)}\ \underline{X}_2^{(2)}\ \underline{X}_3^{(2)}\ \cdots\ \underline{X}_{M_2}^{(2)}\right],$$

and

$$W_t = \left[\underline{X}_1^{(1)}\ \underline{X}_2^{(1)}\ \cdots\ X_{M_1}^{(1)} \vdots \underline{X}_1^{(2)}\ \underline{X}_2^{(2)}\ \cdots\ \underline{X}_{M_2}^{(2)}\right]$$

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

$$W_t W_t^+ = W_1 W_1^+ + W_2 W_2^+, \tag{2}$$

since

$$W_i W_i^+ = \sum_{j=1}^{M_i} \underline{X}_j^{(i)}\, \underline{X}_j^{(i)+}, \quad i = 1, 2$$

is the sample correlation matrix for the i^{th} class. Finding the basis functions normally involves diagonalizing the $(N \times N)$ matrix of $(W_t W_t^+)$, which is a very time consuming task. But, when the intra-class variation in each class is small, this time consuming task can be simplified to the following steps.

(b) We find the $(M_1 + M_2)$ eigenvectors E and eigenvalues Λ of $(W_t^+ W_t)$,

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

It can be seen from multiplying Eq. (3) from the left on both sides by W_t and regrouping that diagonalizing $(W_t W_t^+)$ gives $(W_t E)$ and Λ as the eigenvectors and

eigenvalues

$$(W_t W_t^+)\,(W_t E) = (W_t E)\,\Lambda \tag{4}$$

Therefore, the time consuming task of diagonalizing $(W_t W_t^+)$ has become the much easier problem of diagonalizing the $[(M_1 + M_2) \times (M_1 + M_2)]$ matrix $(W_t^+ W_t)$ and calculating the product $(W_t E)$.

(c) By choosing the M_1 eigenvectors with the largest eigenvalues, we construct matrices E_c and Λ_c of dimensions $[(M_1 + M_2) \times M_1]$ and $[M_1 \times M_1]$, respectively.

(d) The eigenvectors Ψ, Θ and eigenvalues Γ, M of the transformed single class correlation matrices are given by

$$\left[\Lambda_c^{-1} E_c^+ W_t^+ (W_1 W_1^+) W_t E_c \Lambda_c^{-1}\right]\Psi = \Psi\Gamma \tag{5a}$$

$$\left[\Lambda_c^{-1} E_c^+ W_t^+ (W_2 W_2^+) W_t E_c \Lambda_c^{-1}\right]\Theta = \Theta M \tag{5b}$$

It is shown in reference 9 that the eigenvectors of equations (5a) and (5b) are identical,

$$\Theta = \Psi \tag{6a}$$

and that the eigenvalues of equations (5a) and (5b) are related by

$$M = 1 - \Gamma \tag{6b}$$

Thus, the eigenvectors with large eigenvalues applied to one class will have small eigenvalues when applied to the other class.

(e) The basis functions for the F-K transform are $(\Psi_i^+ \Lambda_c^{-1} E_c^+ W_t^+)$, where Ψ_i are the eigenvectors in Ψ.

B. Optical Implementation of the F-K Transform

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

$$h^*(-x,-y) = \sum_{pq}^{M} f_{pq}^*(x + p\Delta,\ y + q\Delta) \exp\left[-i\phi_r(x + p\Delta,\ y + q\Delta)\right] \tag{7}$$

where $f_{pq}(x,y)$ is the p,q^{th} basis function generated from displaying the eigenvector

$(\underline{\psi}_i^+ \Lambda_c^{-1} E_c^{-1} W_t^+)$ in a two-dimensional format, $\phi_r(x,y)$ is a random function uniformly distributed from 0 to 2 , and Δ is a shift constant (see refs. 11-12 for more details of the coded-phase optical processor). Since the coded-phase function has an auto-correlation of a delta function, a coherent optical correlator produces an output consisting of the transform coefficients space at intervals of Δ

$$C(x',\ y') = \int\int g(x,y) \exp\left[i\phi_r(x,y)\right] \sum_{p,q}^{M} f_{pq}^*(x - x' + p\Delta,\ y - y' + q\Delta)$$

$$\exp\left[-\phi_r(x - x' + p\Delta,\ y - y' + q\Delta)\right] dxdy$$

$$= \sum_{p,q}^{M} \left[\int\int g(x,y) f_{pq}^*(x,y)\, dxdy\right] \delta(x' - p\Delta,\ y' - q\Delta) \qquad (8)$$

Both the coded-phase distribution which illuminates the input and the spatial filter can be obtained by computer generated holograms.

Experimentally the optical Fukunaga-Koontz transformation was applied to the problem of distinguishing birds from fish. Ten images of song birds were input to the computer through a T.V. 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 figure 2.

The F-K basis functions are shown in figure 3(a) along with the test images consisting of five new birds and five new fish in figure 3(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 grey level which 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 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 4 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 grey level corresponding to zero is shown in the small square under each basis function.

A filter was generated which contained six of the basis functions shown in figure 3(a) in phase-coded form. Basis functions numbers 3, 4, and 1 were chosen to represent fish-like features, and 8, 9 and 10 to represent bird-like 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 figure 5. The output was detected with a T.V. camera, digitized, and displayed on a T.V. monitor. The digital computer can measure these six coefficients, and use them as input data to a linear or non-linear classifier for best classification.

The results of using three of the ten test images in the coded-phase processor

are shown in figures 6(a)-(c). The lower right-hand corner of the T.V. monitor contains a sampled version of the actual light field detected by the T.V. 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.

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 which is located by the mark "x". Points which 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 7 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 figure 2 and ten are from the test set of figure 3(b). A linear decision surface is able to separate the class "birds" from the class "fish" with no errors.

III. THE LEAST-SQUARES LINEAR MAPPING TECHNIQUE

The LSLMT is useful for performing a transform A, which maps an image vector $\underline{X}_j^{(i)}$ from the i^{th} class in the large N-dimensional feature space as close to one specific unit $\underline{V}_i$ as possible in the K-dimensional decision space such that the overall mean-square error incurred in the mapping is minimized (see figure 8)

$$A\underline{X}_j^{(i)} = \underline{V}_i + \underline{\varepsilon}_{ij} \quad (i = 1, 2, \ldots K;\ j = 1, 2, \ldots, M) \tag{9}$$

where $\underline{\varepsilon}_{ij}$ is the error vector for the i^{th} classifier and j^{th} image sample. The algorithm of LSLMT involves maximizing the interclass differences and minimizing the intraclass variations.

A. Procedures for Calculating the Basis Functions

The procedures for calculating the matrix A and the basis functions of the LSLMT are summarized as follows: (Details can be found in ref. 10.)

(a) Given a training set of images $\{\underline{X}_j^{(i)}\}$ which consists of K image classes and M sample images in each class (i = 1, ..., K; j = 1, ..., M), we define a (N × KM) matrix W as

$$W = \left[\underline{X}_1^{(1)}, \underline{X}_2^{(1)}, \ldots, \underline{X}_M^{(1)} \vdots \underline{X}_1^{(2)}, \underline{X}_2^{(2)}, \ldots, \underline{X}_M^{(2)} \vdots \cdots \vdots \underline{X}_1^{(K)}, \underline{X}_2^{(K)}, \ldots, \underline{X}_M^{(K)}\right] \tag{10}$$

(b) We choose the unit vectors $\underline{V}_i$ in the K-dimensional decision space to be orthonormal, e.g., when K = 3,

$$\underline{V}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \underline{V}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \underline{V}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{11}$$

(c) Knowing the vectors $\underline{X}_j^{(i)}$ and $\underline{V}_i$ from parts (a) and (b) respectively, we can calculate $\sum_{i=1}^{K} \sum_{j=1}^{M} \underline{V}_i \underline{X}_j^{(i)+}$.

(d) Knowing the matrix W from part (a), we can find the eigenvectors Ψ and eigenvalues Λ of W^+W

$$W^+W = \Psi \Lambda \Psi^+ \tag{12}$$

It is noted that W^+W is a $(KM \times KM)$ matrix.

(e) It can be shown that the inverse of the correlation matrix is given by

$$\left[\sum_{i=1}^{K} \sum_{j=1}^{M} \underline{X}_j^{(i)} \underline{X}_j^{(i)+} \right]^{-1} = W \Psi \Lambda^{-2} \Psi^+ W^+ \tag{13}$$

(f) Using the results of parts (c) and (e), we can calculate the matrix A for LSLMT.

$$A = \left[\sum_{i=1}^{K} \sum_{j=1}^{M} \underline{V}_i \underline{X}_j^{(i)+} \right] \left[\sum_{i=1}^{K} \sum_{j=1}^{M} \underline{X}_j^{(i)} \underline{X}_j^{(i)+} \right]^{-1} \tag{14}$$

The basis functions are given by the row vectors of matrix A.

B. Optical Implementation of LSLMT

To employ the coded-phase processor for performing the matrix-vector multiplication of equation (9), we first need to convert the column vector $\underline{X}_j^{(i)}$ corresponding to the input image g(x,y) back to its two-dimensional image format of $(n \times n)$ pixels and the row vectors of the matrix A to the pattern functions $f_{pq}(x,y)$. The total number of pattern functions is K, i.e., $p + q = 1, 2, \ldots, K$, because there are only K row vectors in the matrix A. The output from the coded-phase processor will then be in a two-dimensional form with a spacing Δ between different light spots whose magnitudes of brightness are V_i, $i = 1, 2, \ldots, K$. The brightest spot indicates which class the input image belongs to.

The key element required in optically implementing LSLMT is a computer generated hologram filter whose impulse response contains the pattern functions f_{pq} derived from the row vectors of matrix A. The advantages of optical classification of statistical patterns are parallel processing and real-time data rates. As an example, to recognize any unknown image of $(n \times n)$ pixels which has the same statistical properties as those images used in the training sets of K-classes, the amount of parallel computations involved is $(K \times n \times n)$ multiplications and $(K \times n \times n)$ additions

in real-time. When the real-time rate is 1/30 sec, K is 100, n is 500 and each pixel has 64 gray levels (or 6 bits), the computational rate is greater than 9×10^9 bits/sec.

Experimentally, the LSLMT was applied to design a classifier for hand-written letters in a variety of styles. Ten images of each of the hand-written characters m, t and a (as shown in figure 9) were input to the computer through a TV camera/digitizer system forming the training sets of three classes (detail description about our micro-computer based video-image analysis system are given in refs. 11-12). Using the thirty image samples for the three classes and following the procedures outlined in Section IIIA, we calculated the matrix A for LSLMT. The three row vectors in A yield the three pattern functions f_{pq} as shown in figure 10. Using these pattern functions we next generate a computer hologram filter whose impulse response is given by equation (7). The computer generated filter is shown in figure 11 and applied to the eight test images shown in figure 12.

Three of the eight results from optical implementation of LSLMT are illustrated in figure 13. On the right the outputs from the coded-phase processor are displayed, given the test patterns on the left. Each output contains three spots of light in pre-determined locations. The square roots of the spot intensities correspond to the absolute values of V_1, V_2 and V_3. The location of the brightest spot in the output plane clearly indicates classification. The measured values of V_i for all eight test patterns are summarily listed in table 2 and plotted in figure 14. Based on table 2 and figure 14, we can conclude that we have successfully demonstrated the optical implementation of LSLMT.

IV. COMPARISONS AMONG VARIOUS OPTICAL PATTERN RECOGNITION METHODS

We now turn to compare the LSLMT with other methods of optical pattern recognition and classification. The comparisons cover the matched filter, the average filter and the Fukunaga-Koontz transform. The first two comparisons were carried out digitally. The third comparison was optical.

A. Comparison with the Matched Filter

Two digital filters were generated with the impulse responses of the first letters m or t in the training sets shown in figure 9(a),(b). All the letters of m and t in the training sets (fig. 9(a),(b)) and the test set (fig. 15(a)) were used as the inputs g(x,y). Figure 15(b) shows the normalized correlations between the inputs and the matched filters, which are defined by

$$V_1 = \langle g m_1^* \rangle / \langle m_1 m_1^* \rangle$$

$$V_2 = \langle g t_1^* \rangle / \langle t_1 t_1^* \rangle \qquad (15)$$

where $\langle\rangle$ signifies an inner product. Using the dotted line as the classification boundary will obviously yield zero error of classification. However, a more restrictive classifier defined as the ratio between V_1 and V_2 is often preferred because it is insensitive to scalar multiplication of g(x,y) which may result from changes in input illumination. This more restrictive classification boundary passes through

the origin, as shown by the solid line in figure 15, and yields 30% error in recognition.

This digital matched filtering result is compared with LSLMT using the same set of letters in figures 9(a),(b), and 15(a). Figure 16 shows that the LSLMT can provide error free classification.

B. Comparison with an Average Filter

Digital average filters were calculated by averaging the characters of the training sets in figure 9. The digital filters are shown in figure 17, which can be compared with the pattern functions f_{pq} for LSLMT (fig. 10). While the average filters are formed by superposing only the images of the same class, to form pattern functions f_{pq} for LSLMT, it appears that one may superpose not only the images of one class but also subtract the superposed images from other classes.

Using the digital average filters, we obtain the correlations V_1 and V_2 between the averaged characters and the input g(x,y), which can be any image in the training (fig. 9(a),(b)) or test (fig. 15(b)) sets.

$$V_1 = \langle g \sum_j m_j^* \rangle / \langle | \sum_j m_j |^2 \rangle$$

$$V_2 = \langle g \sum_j t_j^* \rangle / \langle | \sum_j t_j |^2 \rangle \qquad (16)$$

Figure 18 shows the results of using digital average filters. The error rate is 20%, which is a little better than the 30% rate associated with digital matched filters. But, both digital average and matched filters yield results inferior to the LSLMT, which has a zero error rate.

C. Comparison with Fukunaga-Koontz Transform

This comparison was performed optically using the coded-phase optical processor and computer generated hologram filters. Two computer hologram filters were generated using the same training setsof images of song birds and fish, one based on the principles of the FKT described in reference 9 and the other based on the LSLMT (ref. 10). The results of optically implementing the LSLMT are shown in figure 19. Two of the six test images used as input are shown in the left column. The corresponding outputs each containing two spots of light for this two-class classification problem are shown in the middle column. The linear classification results are shown in the right column. These linear classification results together with those for four other test images are shown in figure 20(a), which are to be compared with the linear classification results based on the FKT as shown in figure 20(b). It is obvious that both FKT and LSLMT are good classifiers. The FKT has the advantage that it can provide many features for each image. These features can be used as the basis for a nonlinear classification routine. However, it is useful only for two-class problems. The LSLMT provides only K features (one feature for each class), but it can be applied to multiclass problems (K > 2).

V. SUMMARY

We have shown that the Fukunaga-Koontz transform can be used as a feature extractor in a two-class classification application. It was used for classifying birds and fish. After the F-K basis functions were calculated, those most useful for classification were incorporated into a computer generated hologram. A coherent optical processor was designed 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 T.V. camera, digitized, and fed into a micro-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.

The Least-Squares Linear Mapping Technique has been optically implemented to classify large images also by utilizing the phase-coded optical processor and computer generated hologram filters. The principles of LSLMT are incorporated into the computer hologram filters. Although we have demonstrated here only two- and three-class problems, the LSLMT is useful for optical classification problems of many classes.

A method was developed which simplified the computation of the FKT or LSLMT basis functions for large dimensional imagery, and was found to work well when the intraclass variation in each class was small, and the correlation matrix could be approximated by a sample correlation matrix of low rank.

The advantages of optically implementing the FKT and LSLMT are parallel and real-time processing. In comparisons with the matched filter and the average filter, the LSLMT is clearly superior for classifying statistical patterns because it maximizes the interclass differences and minimizes the intraclass variations.

ACKNOWLEDGEMENTS

The research was performed in collaboration with J. Leger and Z. Gu. The support of this research from the National Science Foundation and the Air Force Office of Scientific Research is gratefully acknowledged.

REFERENCES

1a. Vander Lugt, A.: Signal Detection by Complex Spatial Filtering. IEEE Trans. Inform. Theory, vol. IT-10, 1964, pp. 139-145.

1b. Vander Lugt, A., F. B. Rotz and A. Klooster: Character Reading by Optical Spatial Filtering. In Optical and Electrooptical Information Processing, Tippet et al., eds., MIT Press, Cambridge, Mass., 1965, p. 125.

2. Gara, A. D.: Real-time Tracking of Moving Objects by Optical Correlation. Appl. Opt., vol. 18, 1979, pp. 172-174.

3. Almeida, S. P. and J. K. T. Eu: Water Pollution Monitoring Using Matched Spatial Filters. Appl. Opt, vol. 15, 1976, pp. 510-515.

4. Casasent, D. and D. Psaltis: Position, Rotation, and Scale Invariant Optical Correlation. Appl. Opt., vol. 15, 1976, pp. 1795-1799.

5. Hester, C. F. and D. Casasent: Optical Pattern Recognition Using Average Filters to Produce Discriminant Hypersurfaces. Proc. SPIE, vol. 201, 1979, pp. 77-82.

6. Caulfield, H. J., R. Haimes and D. Casasent: Beyond Matched Filtering. Opt. Eng., vol. 19, 1980, pp. 152-156.

7. Leger, J. R. and S. H. Lee: A Hybrid Optical Processor for Pattern Recognition and Classification Using a Generalized Set of Pattern Functions. Appl. Opt., 1982 (to be published).

8. Duvernoy, J: Optical Pattern Recognition and Clustering: Karhunen Loéve Analysis, Appl. Opt., vol. 15, 1976, pp. 1584-1590.

9. Leger, J. R. and S. H. Lee: Optical Implementation of Fukunaga-Koontz Transform. J. Opt. Soc. Amer., 1982 (to be published).

10. Gu, Z. H., J. R. Leger and S. H. Lee: Optical Implementation of the Least-Squares Linear Mapping Technique for Image Classification. J. Opt. Soc. Amer., 1982 (to be published).

11. Leger, J. R. and S. H. Lee: Coherent Optical Implementation of Generalized Two-Dimensional Transforms. Opt. Eng., vol. 18, 1979, pp. 518-523.

12. Leger, J. R., J. Cederquist and S. H. Lee: Micro-computer Based Hybrid Processors at UCSD. Opt. Eng., 1982 (to be published).

Table 1. Eigenvalues Corresponding to Ten Basis Functions of FKT

Basis Function	Eigenvalue (λ_i)	$\lvert\lambda_i - 0.5\rvert$	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

Table 2. The measured values of V_1, V_2 and V_3 for the eight test patterns shown in Figure 12.

	V_1	V_2	V_3	Largest for which class?
m_1	1.58	0.02	0.12	m
m_2	0.93	0	0.12	m
t_1	0.02	1.1	0.07	t
t_2	0.07	1.52	0.01	t
t_3	0.05	0.95	0.02	t
a_1	0.02	0.01	0.88	a
a_2	0.03	0.01	1.43	a
a_3	0.02	0.01	1.22	a

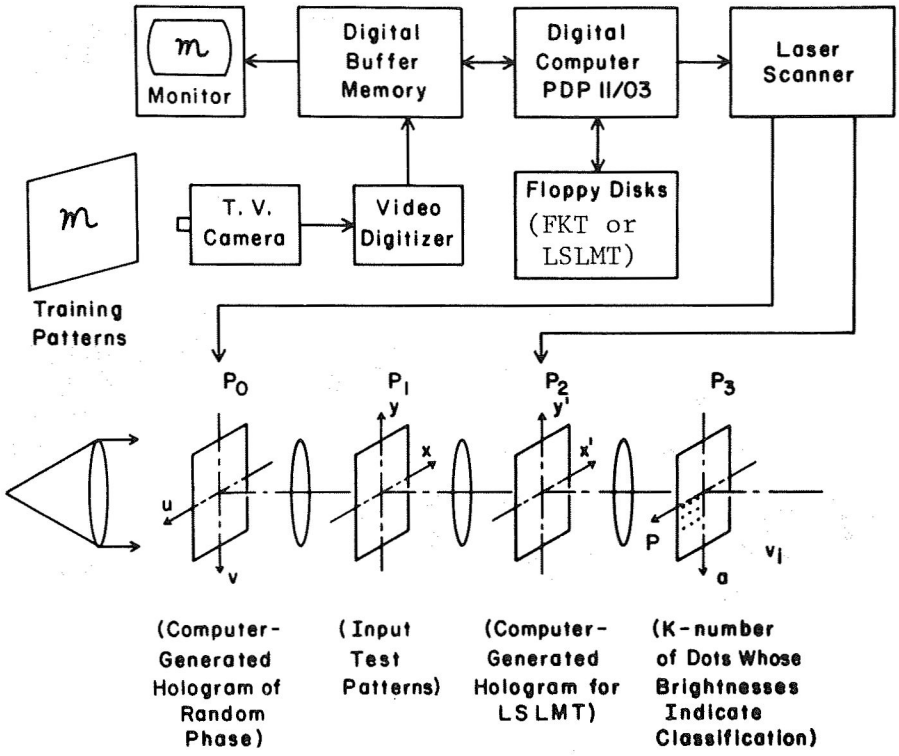
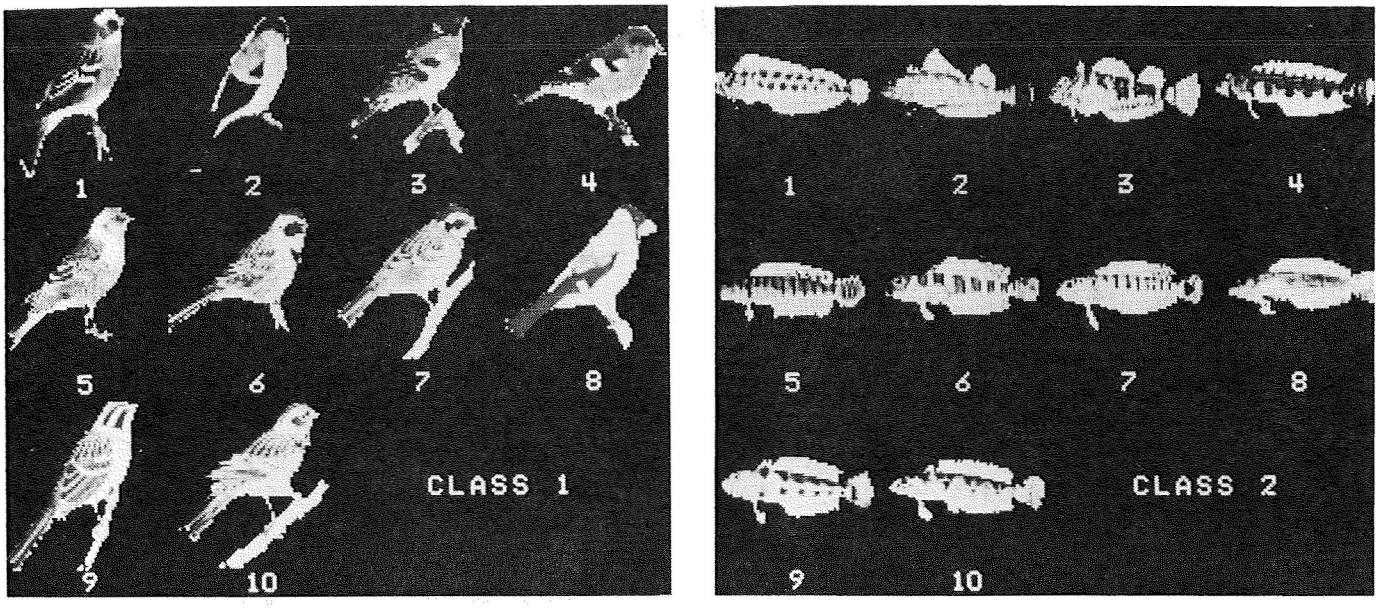
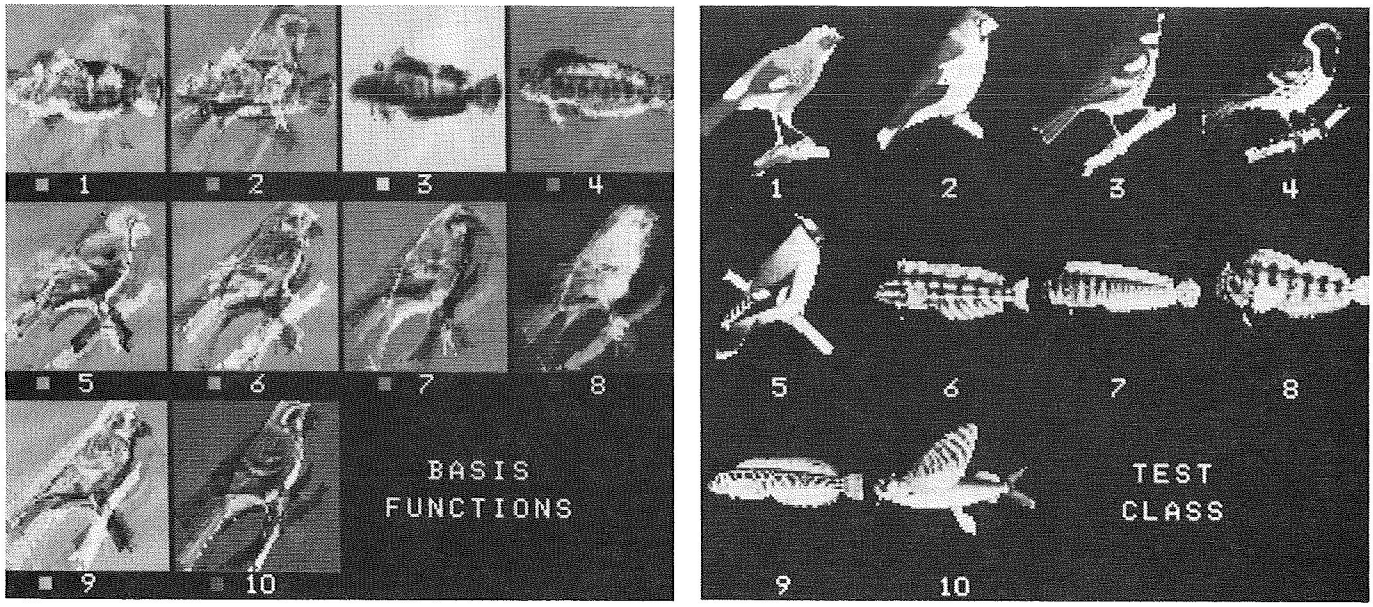


Figure 1.- The hybrid system for optical recognition of statistical patterns.



(a) (b)

Figure 2.- (a) Class 1 training set consisting of ten song birds. (b) Class 2 training set consisting of ten fish.



(a) (b)

Figure 3.- (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 grey level corresponding to zero. (b) Test images consisting of five new birds and five new fish.

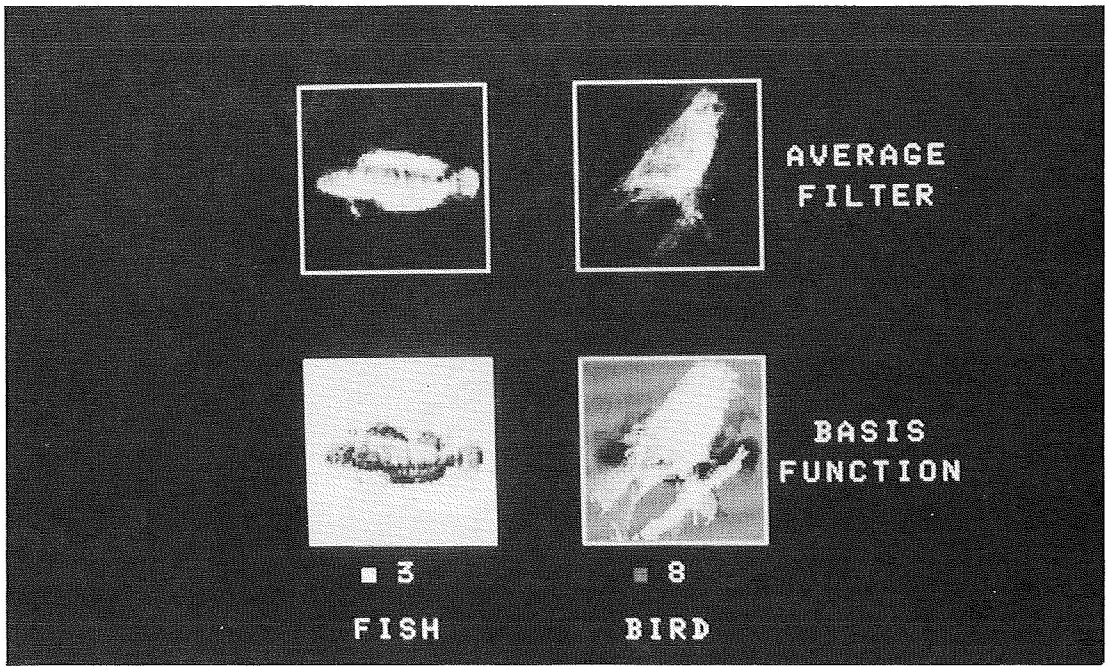


Figure 4.- Comparisons of basis functions 3 (fish) and 8 (birds) with filters formed by the arithmetic average of the training sets.

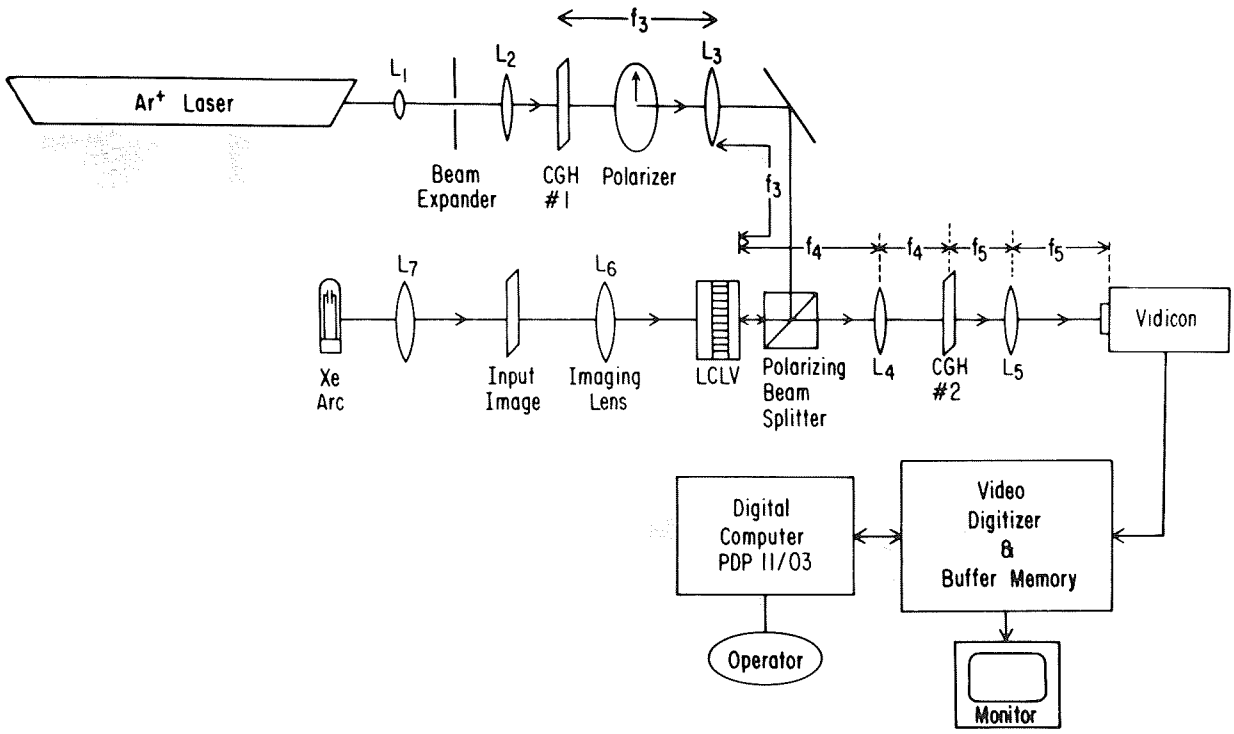


Figure 5.- Hybrid implementation of the coded-phase optical processor. A computer hologram of a coded-phase array is shown as CGH#1. A second computer hologram containing the F-K basis functions in coded-phase form is shown as CGH#2. LCLV is a liquid crystal light valve for converting an incoherent (test) 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.

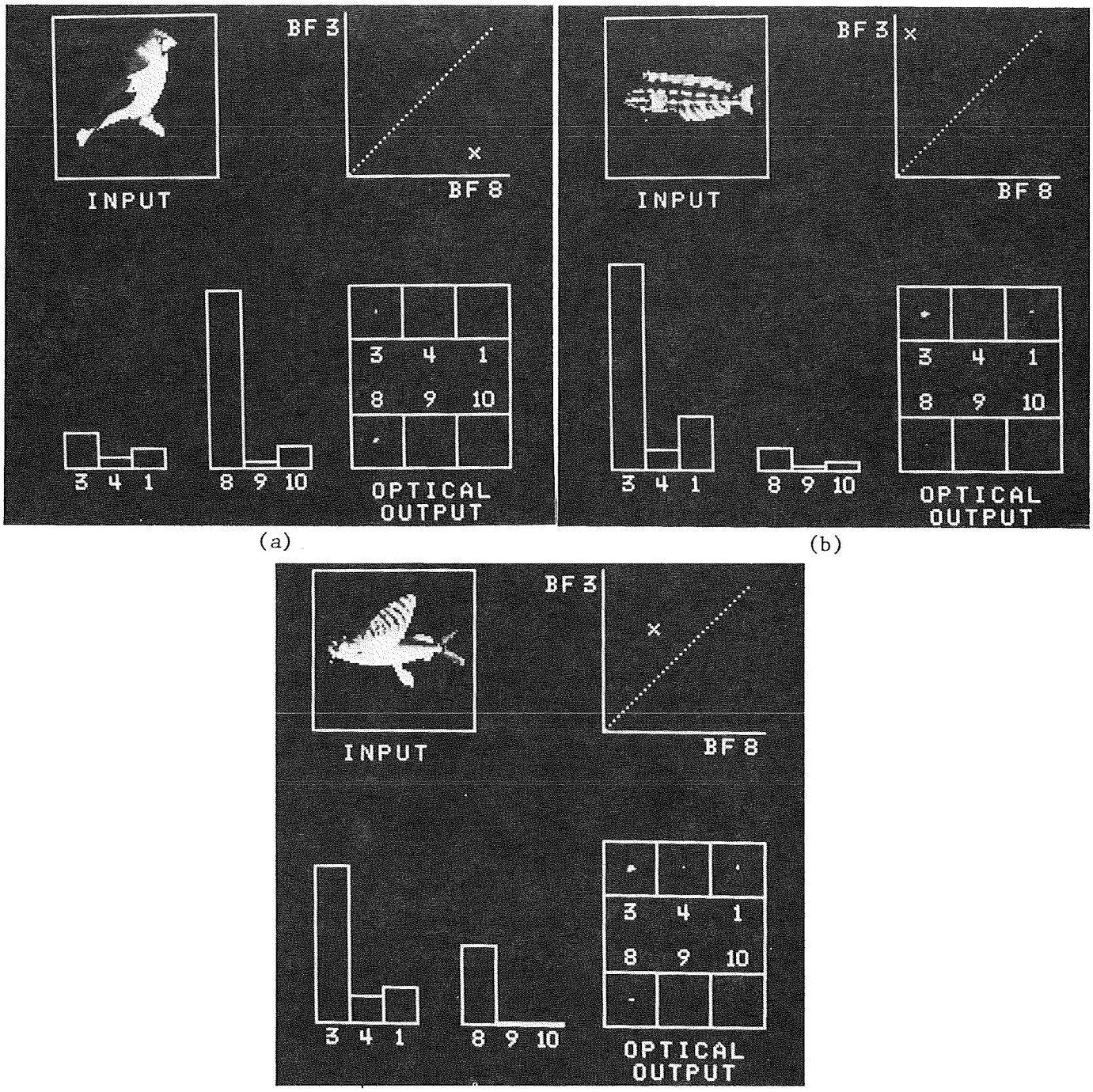


Figure 6.- Optical implementation of the F-K transform using the coded-phase optical processor. The six basis functions with greatest separation power were used. Basis functions 3, 4 and 1 are best for fish, and 8, 9 and 10 are best for birds. Three of the ten images are used as inputs in (a),(b)& (c), where the input is reproduced in the upper left-hand corner of the T.V. screen. The output of the optical processor is detected by a T.V. 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).

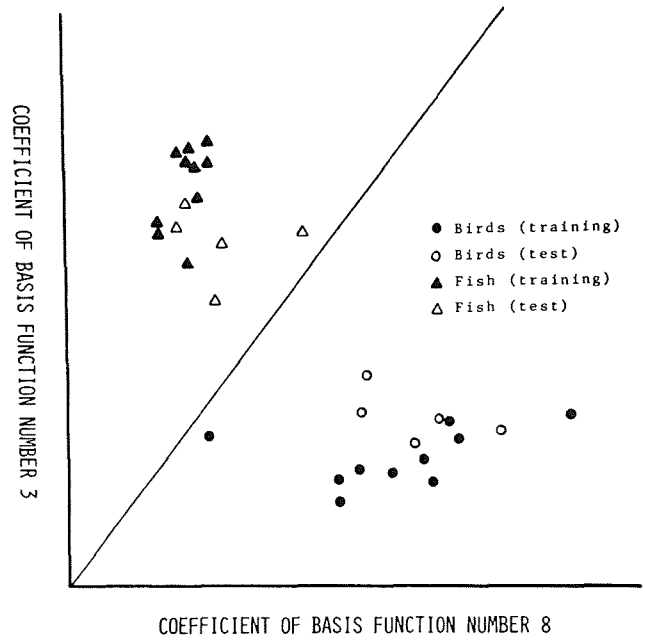


Figure 7.- Classification of birds and fish using coefficients of basis functions 3 and 8. Ten members of each training set, as well as ten members of the test set were classified.

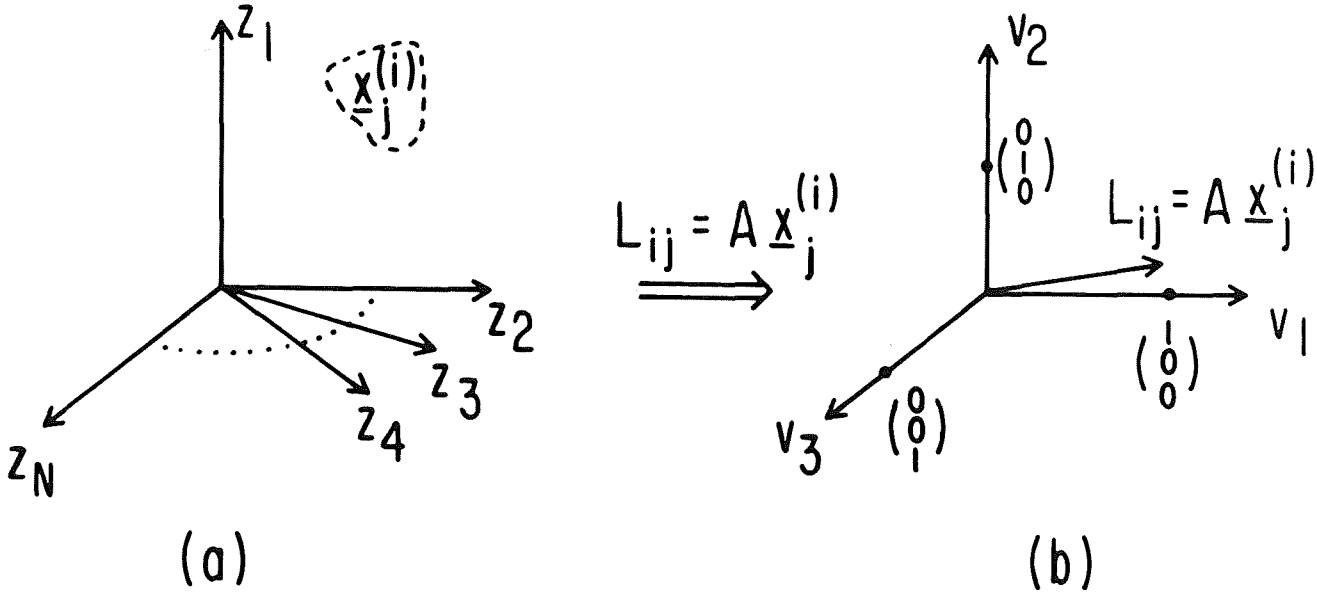


Figure 8.- The schematic diagrams of LSLMT for 3 classes. (a) N-dimensional feature space. (b) Three-dimensional decision space.

Figure 9.- (a) Class 1 training set consisting of ten characters of m. (b) Class 2 training set consisting of ten characters of t. (c) Class 3 training set consisting of ten characters of a.

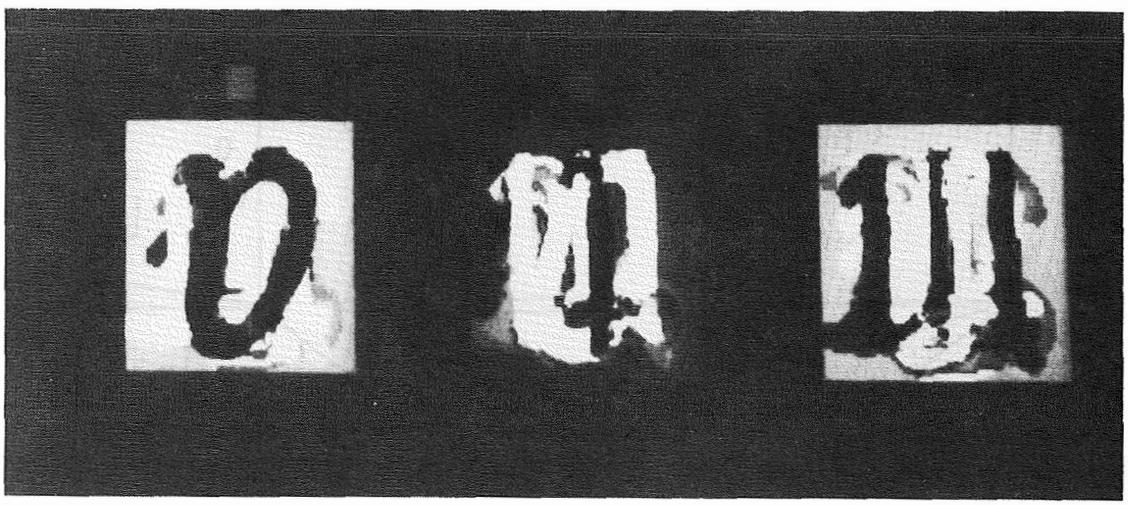
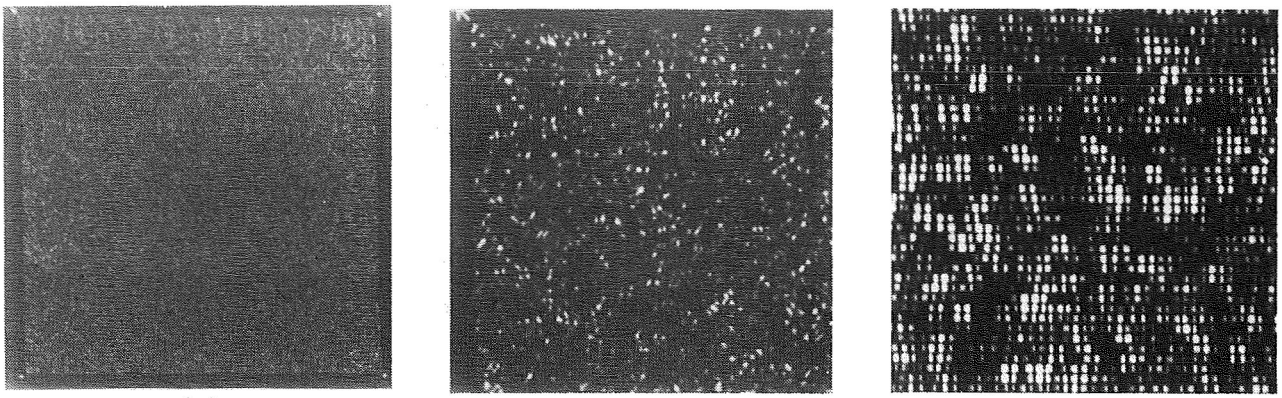


Figure 10.- Three pattern functions f_{pq} for LSLMT. Since the pattern functions f_{pq} contain both positive and negative values in general, bias levels (indicated by the small grey squares below f_{pq}) are added to f_{pq} to display them.



(a) (b) (c)

Figure 11.- A computer-generated spatial filter for LSLMT made with a laser scanning system. (a) 4 × 4 filter array, (b) one filter, (c) the central part of (b).

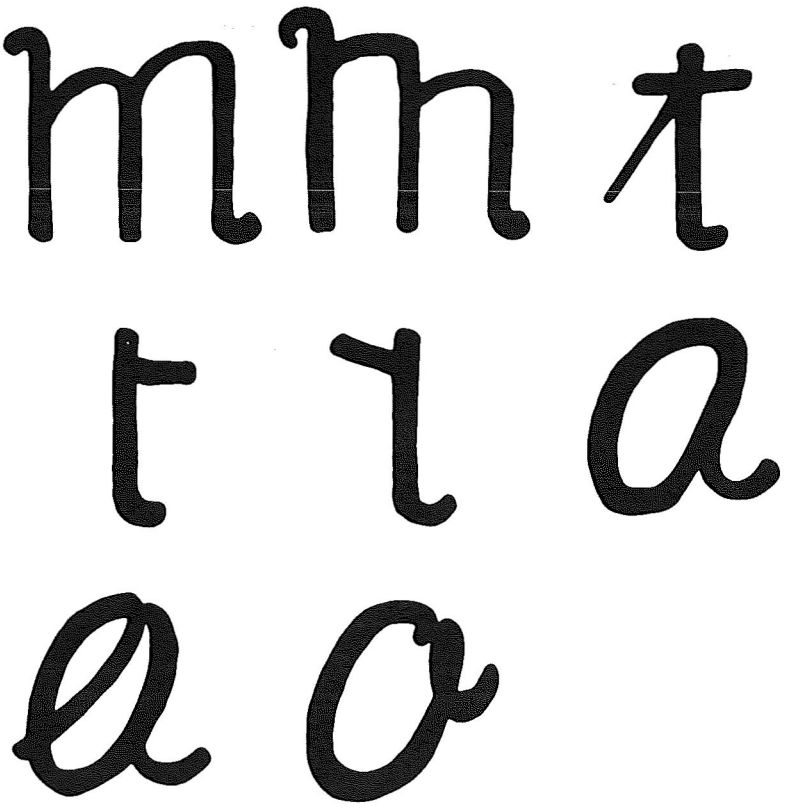


Figure 12.- A test set of images consisting of two new m-characters, three new t-characters and three new a-characters.

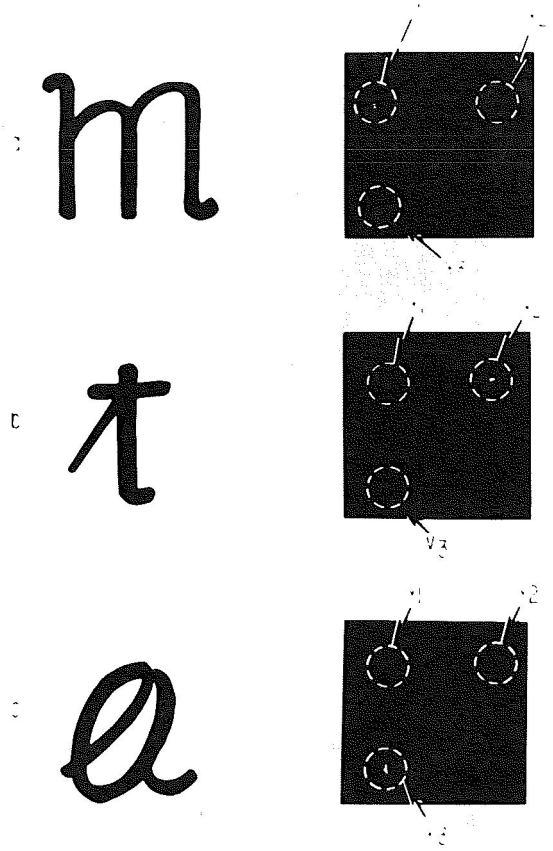


Figure 13.- The result of optical implementation of the LSLMT for 3 classes. Three of the eight test images used are shown on the left. The outputs of the optical processor are displayed on the right.

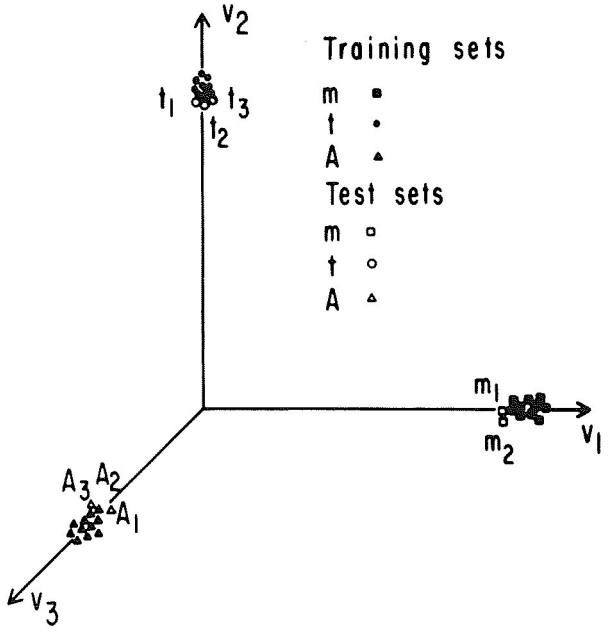
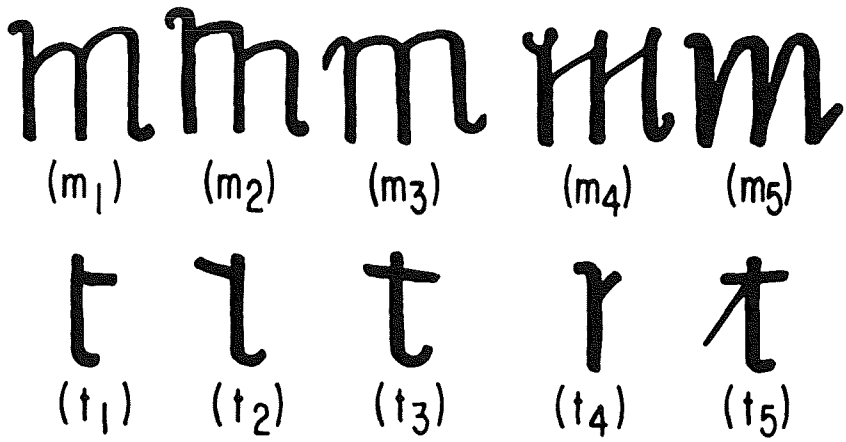
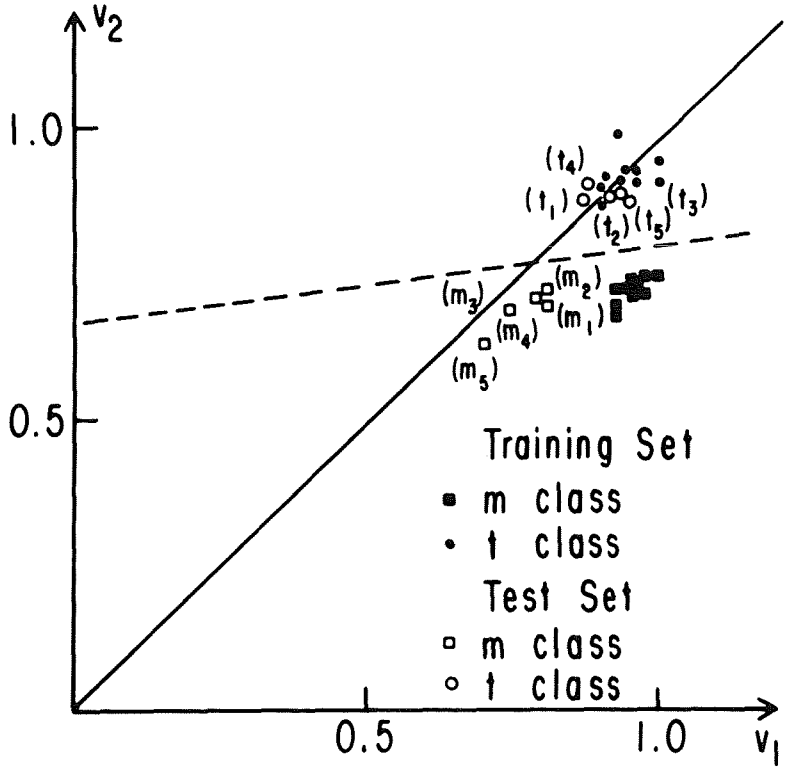


Figure 14.- Classification result with LSLMT for three classes (m,t,a). The solid symbols of triangle, circle and square are for the training sets. The hollow symbols of triangle, circle and square are for the test images.



(a)



(b)

Figure 15.- (a) A test class of m and t characters. (b) Classification using the matched filter for 2 classes (m,t). The solid line represents a restrictive linear classification boundary. The error rate is 9 out of 30, or 30%.

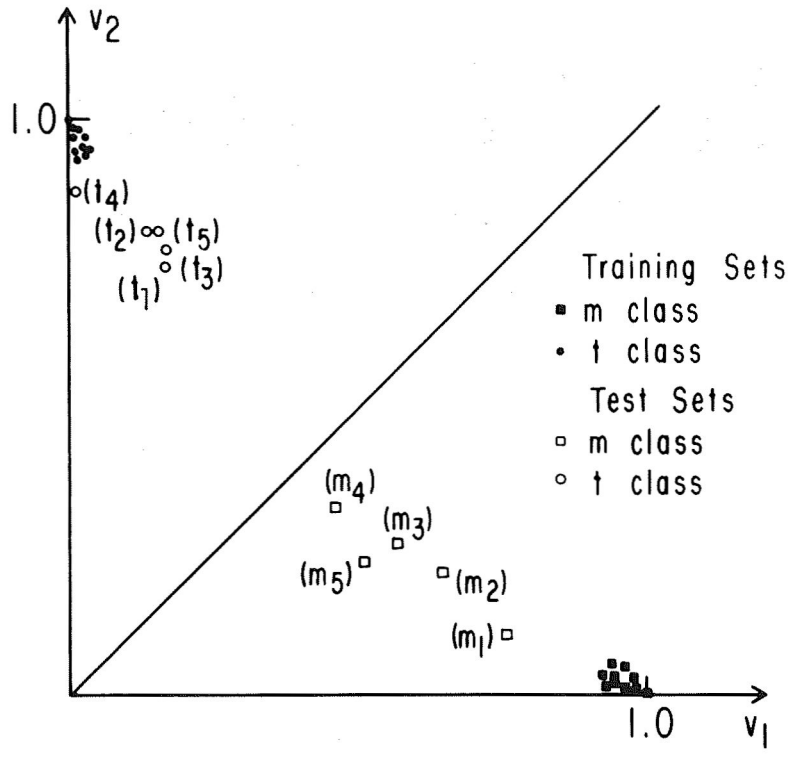


Figure 16.- Classification result with LSLMT of 2 classes (m and t characters).

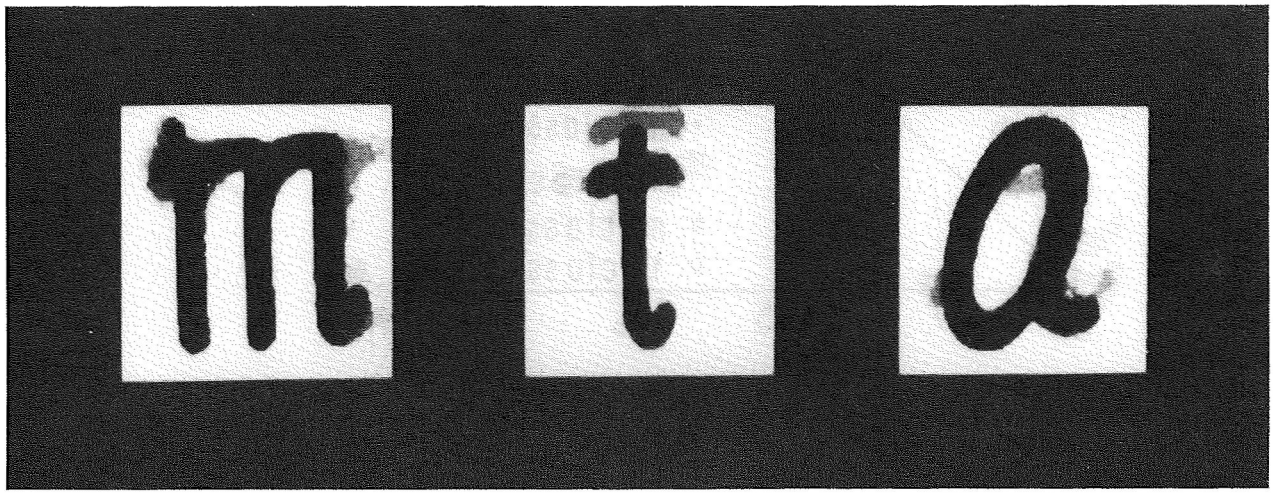


Figure 17.- 3 average filters of the training sets (m,t,a).

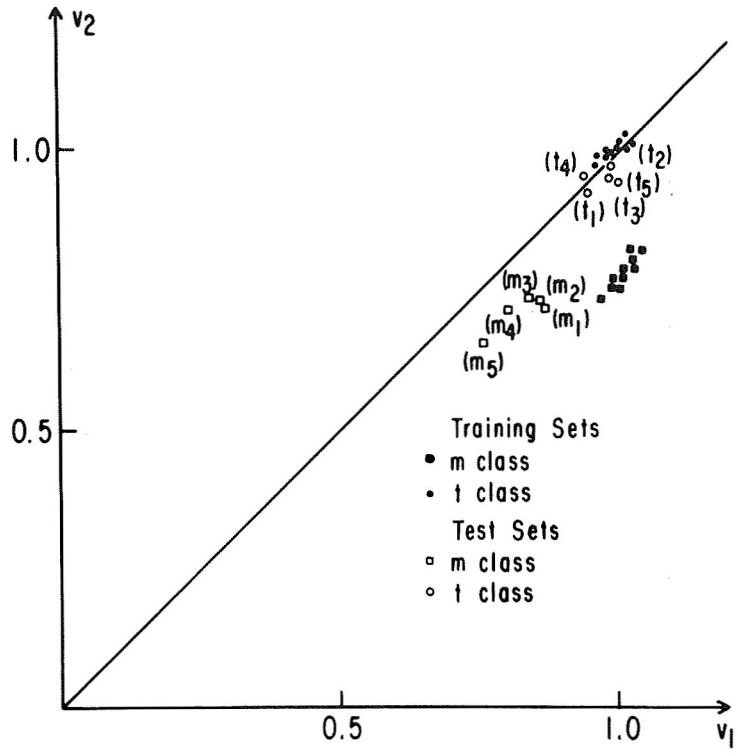


Figure 18.- Classification using average filters for 2 classes (m,t). The error rate is 6 out of 30, or 20%.

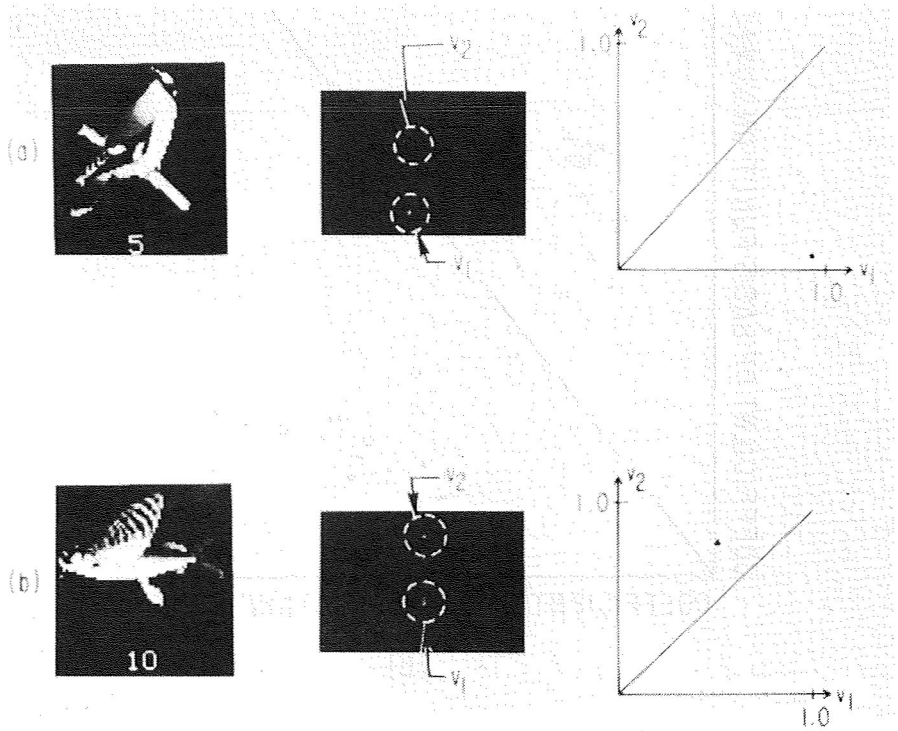


Figure 19.- The optical implementation of the LSLMT for two classes (Song Bird and Fish). The left column shows two of the six test images as input, the middle column shows the outputs, the linear classifier is shown in the right column.

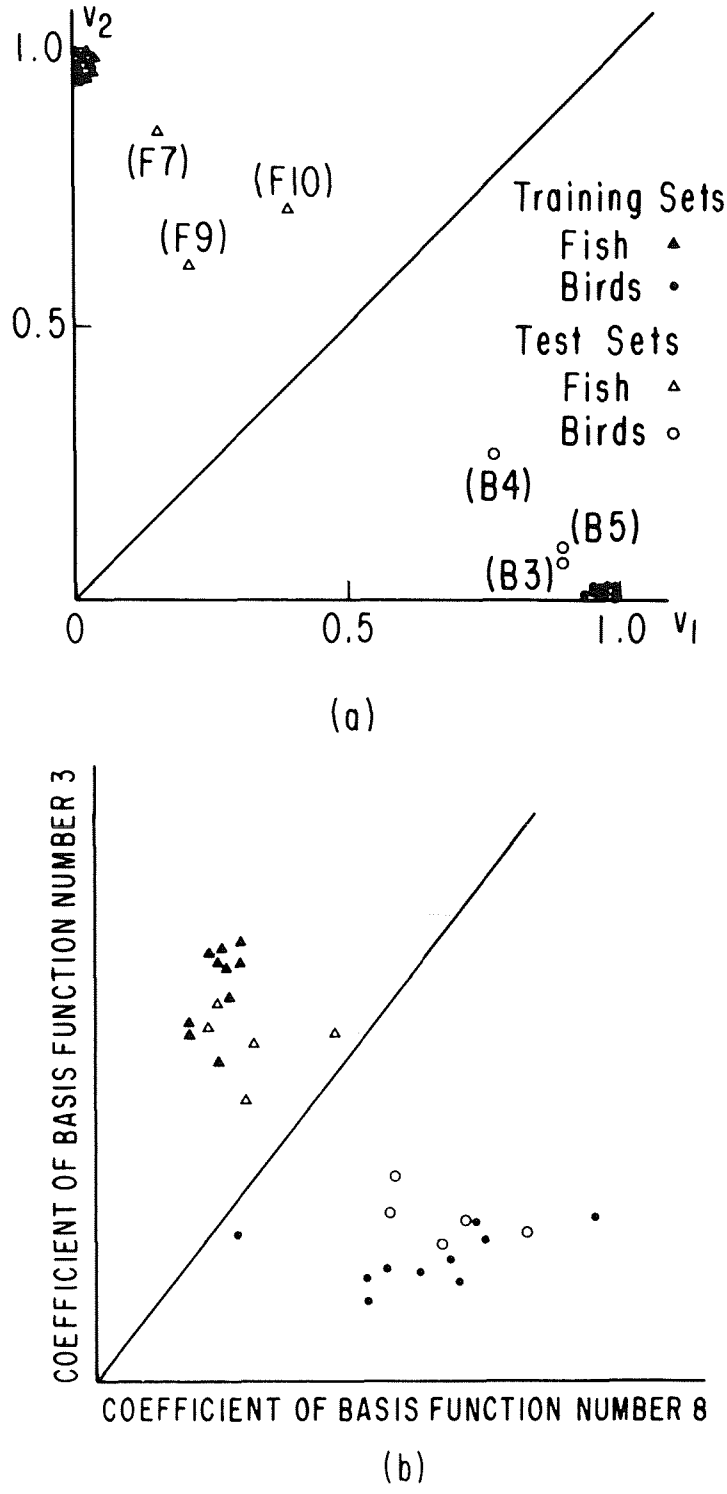


Figure 20.- (a) Results of optically implementing the LSLMT for two classes (Song Bird and Fish). (b) Results of optically implementing the Fukunage-Koontz transform using coefficients of basis functions 3 and 8.