

AD 740620

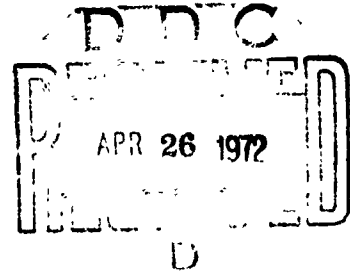
AFOSR - TR - 70 - 0 - 517

FINAL REPORT

REPORT FOR THE PERIOD JULY 1, 1970 - AUGUST 31, 1971
OF WORK PERFORMED UNDER AIR FORCE OFFICE OF SCIENTIFIC
RESEARCH CONTRACT NUMBER F44620-70-C-0113

PREPARED FOR:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
1400 WILSON BLVD.
ARLINGTON, VIRGINIA



BY: THE DEPARTMENT OF BIOLOGICAL SCIENCES AND
THE IMAGE PROCESSING LABORATORY, ELECTRONIC
SCIENCES LABORATORY, DEPARTMENT OF ENGINEERING
UNIVERSITY OF SOUTHERN CALIFORNIA
UNIVERSITY PARK
LOS ANGELES, CALIFORNIA 90007
FEBRUARY, 1972

Approved for public release;
distribution unlimited.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) University of Southern California Biological Sciences & Electrical Engineering Los Angeles, California 90007	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED
	2b. GROUP

3. REPORT TITLE
RESEARCH ON OPTICAL IMAGE PROCESSING

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Scientific Final

5. AUTHOR(S) (First name, middle initial, last name)
William K. Pratt L. R. Welch
H. C. Andrews
L. A. Liporace

6. REPORT DATE February 1972	7a. TOTAL NO. OF PAGES 28	7b. NO. OF REFS 29
---------------------------------	------------------------------	-----------------------

8a. CONTRACT OR GRANT NO. F44620-70-C-0113 b. PROJECT NO. 9777 61102F d. 681312	9a. ORIGINATOR'S REPORT NUMBER(S)
	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT
Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES TECH, OTHER	12. SPONSORING MILITARY ACTIVITY Air Force Office of Scientific Research 1400 Wilson Boulevard (NL) Arlington, Virginia 22209
---	--

13. ABSTRACT
Research on computer processing of optical information included coding, enhancement and detection using monochrome and color images. A bandwidth reduction and tolerance to channel errors for monochrome images have resulted from digital image transforms. Sample reductions on the order of the ratio 5 to 1, compared to conventional PCM coding, have been reported for transform coding of monochrome images. Spatial redundancy of color images and the limitations of human color vision can be exploited by transform coding to achieve a substantial bandwidth reduction for the digital transmission of color television and facsimile. The generalized Wiener filtering technique has been applied to the enhancement of images. The noise effects were substantially reduced. A new measure of similarity between images has been developed which is proposed as a useful detection criterion. The measure has information theoretic properties and, with proper image normalization, is the entropy function. The measure is convex upward reaching a peak when two images are identical and thus perfect image detection occurs at the maximum of this image similarity measure. Five manuscripts have been submitted for publication and eight papers were presented at various conferences/symposia.

Reproduced by
**NATIONAL TECHNICAL
INFORMATION SERVICE**
Springfield, Va 22151

FINAL REPORT

RESEARCH ON OPTICAL IMAGE PROCESSING

Dr. William K. Pratt, principal investigator
Dr. H. C. Andrews
Dr. L. A. Liporace
Dr. L. R. Weich

Research on Optical Image Processing

1. Introduction

The research performed on optical image processing falls into three rather broad categories: image coding, image enhancement, and image detection. Consideration has been given both to monochrome and color images. Also, a study has been undertaken to evaluate the implementation requirements of one of the more promising image processing techniques.

The following sections summarize the results of some of the research performed in this study. More detailed accounts of the various research projects are to be found in the published papers listed at the end of the section.

2. Spatial Transform Coding of Color Images

Digital image transforms have been applied quite successfully to obtain a bandwidth reduction and tolerance to channel errors for monochrome images [1-4]. The potential for monochrome image bandwidth compression with transform coding arises from the spatial correlation within natural images. As a result of this spatial correlation, the image energy within the transform domain tends to be clustered toward a relatively few number of transform samples. Low magnitude samples may either be discarded completely or coded with a few number of bits without seriously effecting image quality. Sample reductions on the order of 5:1, compared to conventional PCM coding, have been reported for transform coding of monochrome images [1-3].

Preliminary studies have indicated that the spatial redundancy of color images and the limitations of human color vision can be exploited by transform coding to achieve a substantial bandwidth reduction for the digital transmission of color television and facsimile [4-8]. Figure 1 contains a block diagram of a color image transform coding system. In the system the color image is represented by three source tristimulus signals $R(x,y)$, $G(x,y)$, $B(x,y)$ that specify the red, green, and blue content of an image element at coordinates (x,y) , according to the N.T.S.C. receiver phosphor primary system [9].

The source tristimulus signals are then converted to a new three dimensional space through some linear or nonlinear, invertible coordinate conversion process. The objective of the coordinate conversion is to produce three planes of data, $f_1(x,y)$, $f_2(x,y)$, $f_3(x,y)$, to be called color signals, that are most amenable to transform coding. Next, a spatial, unitary transformation is performed on each color signal plane resulting in three transform domain planes $F_1(u,v)$, $F_2(u,v)$, $F_3(u,v)$. Quantization and coding to achieve a bandwidth reduction is then performed on the three transform domains. At the receiver, the channel output is decoded, and the inverse spatial transforms and inverse coordinate conversion operations are performed to reconstruct the source tristimulus signals.

Color Image Transform Coding

Consider the color image transform coding system of Figure 1 in which the color coordinate conversion is linear. The three color signal planes can be expressed as

$$\begin{aligned}\underline{f1} &= m_{11}\underline{R} + m_{12}\underline{G} + m_{13}\underline{B} \\ \underline{f2} &= m_{21}\underline{R} + m_{22}\underline{G} + m_{23}\underline{B} \\ \underline{f3} &= m_{31}\underline{R} + m_{32}\underline{G} + m_{33}\underline{B}\end{aligned}\tag{1}$$

Each of these color signal planes is then separately transformed, by perhaps a different unitary transform ($\underline{A}_k, \underline{A}_k$), to produce three transform planes

$$\begin{aligned}\underline{F1} &= \underline{A1} \underline{f1} \underline{A1} = m_{11}\underline{A1} \underline{R} \underline{A1} + m_{12}\underline{A1} \underline{G} \underline{A1} + m_{13}\underline{A1} \underline{B} \underline{A1} \\ \underline{F2} &= \underline{A2} \underline{f2} \underline{A2} = m_{21}\underline{A2} \underline{R} \underline{A2} + m_{22}\underline{A2} \underline{G} \underline{A2} + m_{23}\underline{A2} \underline{B} \underline{A2} \\ \underline{F3} &= \underline{A3} \underline{f3} \underline{A3} = m_{31}\underline{A3} \underline{R} \underline{A3} + m_{32}\underline{A3} \underline{G} \underline{A3} + m_{33}\underline{A3} \underline{B} \underline{A3}\end{aligned}\tag{2}$$

Table 1 defines several unitary transforms that have been examined for color image transform coding. From the above it is apparent that the order of the color coordinate conversion and two dimensional forward transformation processes is immaterial. In fact, the color coordinate conversion can be considered as the third dimension of a separable three dimensional data transformation.

Next, each transform domain plane is quantized and coded. The quantization operation is a nonlinear process that replaces each transform domain sample by its best estimate $\hat{\underline{F1}}, \hat{\underline{F2}}, \hat{\underline{F3}}$. Then, after reception over the

Table 1

Unitary Data Transforms

Fourier: Discrete sinusoidal basis vectors

$$\underline{A} = \left[\exp - \left| \frac{2\pi i}{N} ux \right| \right] \quad \underline{A}^{-1} = \underline{A}^{*T}$$

where N = data vector length
 x = data vector element index
 u = transformed data vector element index

Computation requires about $N \log_2 N$ complex multiply and add operations.

Hadamard: Sequence ordered orthogonal rectangular waveform basis vectors

$$\underline{A} = \underline{H}_N \underline{P} \quad \underline{A}^{-1} = \underline{A}$$

where \underline{P} = row permutation index

$$\underline{H}_N = \begin{bmatrix} \underline{H}_{N/2} & \underline{H}_{N/2} \\ \underline{H}_{N/2} & -\underline{H}_{N/2} \end{bmatrix}$$

$$\underline{H}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Computation requires about $N \log_2 N$ add operations

Note: Many other Hadamard matrices of dimension $N \bmod 4$ are known to exist.

Karhunen-Loeve: Basis vectors are eigenvectors of data covariance matrix.

$$\underline{A} = \underline{K} \quad \underline{A}^{-1} = \underline{K}^T$$

where $\underline{K} \underline{C}_f \underline{K}^{-1} = \text{diag} \left\{ \lambda(i) \right\}$
 \underline{C}_f = covariance matrix of data vector
 \underline{K} = eigenvectors of \underline{C}_f
 $\lambda(i)$ = eigenvalues of \underline{C}_f

Computation requires N^2 multiply and add operations

Note: If \underline{C}_f is a Markov process covariance matrix, i.e. $\underline{C}_f = [\rho^{|x_i - x_j|}]$ where ρ is the adjacent element correlation factor, the eigenvalues and eigenvectors can be found recursively

channel, assuming no channel errors, inverse unitary transforms are taken to reconstruct the color signals

$$\begin{aligned}
 \underline{\hat{f}}_1 &= \underline{A}_1^{-1} \underline{F}_1 \underline{A}_1^{-1} \\
 \underline{\hat{f}}_2 &= \underline{A}_2^{-1} \underline{F}_2 \underline{A}_2^{-1} \\
 \underline{\hat{f}}_3 &= \underline{A}_3^{-1} \underline{F}_3 \underline{A}_3^{-1}
 \end{aligned} \tag{3}$$

Finally, the inverse color coordinate conversion defined by $[\underline{n}_{ij}] \equiv [\underline{m}_{ij}]^{-1}$ is performed to give

$$\begin{aligned}
 \underline{\hat{R}} &= n_{11} \underline{\hat{f}}_1 + n_{12} \underline{\hat{f}}_2 + n_{13} \underline{\hat{f}}_3 \\
 \underline{\hat{G}} &= n_{21} \underline{\hat{f}}_1 + n_{22} \underline{\hat{f}}_2 + n_{23} \underline{\hat{f}}_3 \\
 \underline{\hat{B}} &= n_{31} \underline{\hat{f}}_1 + n_{32} \underline{\hat{f}}_2 + n_{33} \underline{\hat{f}}_3
 \end{aligned} \tag{4}$$

Again, the order of the inverse color coordinate conversion and the inverse spatial transformation is unimportant.

The design procedure for the color image transform coding system of Figure 1 consists of, (a), the selection of the color coordinate conversion matrix; (b), the choice of the unitary transform for each color signal plane; and (c), the specification of the quantization law for transform domain samples.

A considerable effort has been undertaken to analyze the expected performance of various color coordinate conversions and spatial transformations for color image transform coding [10, 11]. The results are summarized below:

- A. The best color coordinate conversion from the standpoint of color plane energy compaction is the Karhunen-Loeve color system which requires knowledge of the correlation between the (R, G, B) color planes.

- B. The (Y, I, Q) color coordinate system, which is the United States standard for color television transmission, provides almost as much energy compaction as the Karhunen-Loeve color coordinate system.
- C. The Karhunen-Loeve spatial transform provides the best energy compaction within a color plane. However, the Fourier and Hadamard transforms perform almost as well.

A series of experiments has been performed to subjectively determine the performance of the color image transform coding system using: a) various linear and nonlinear color coordinate systems; b) the Fourier, Hadamard, and Karhunen-Loeve spatial transforms; c) limitations in data block size; and, d) different quantization and coding techniques.

Summarizing, the experimental results indicate that:

- A) It is preferable to transform code the (Y, I, Q) color coordinates rather than the (R, G, B) coordinates of an image.
- B) Substantial low pass filtering can be performed on the chrominance signal of a color image without noticeable effect on resolution. There is some color desaturation caused by the low pass filtering, but a linear masking operation can be employed to correct for color shifts. For both Fourier and Hadamard transform coding of color images, three bits per element were assigned to the luminance component, Y, and 0.375 bits per element were allotted to the I and Q chrominance components each. Thus, the color image was coded with an average of 3.75 bits per element compared to the original which required 18 bits per element.
- C) With a hybrid coding scheme using a 16 x 16 element Karhunen-Loeve transform for the luminance signal and 16 x 16 Hadamard transforms for the (I, Q) signals, a color image can be coded with only 1.75 bits per element. Preliminary studies indicate that such a system could be implemented for operation approaching real time rates.

References

- 1.. A. Habibi and P.A. Wintz "Image Coding by Linear Transformations and Block Quantization," IEEE Trans. on Communication Technology, Vol. COM-19, No. 1 (February, 1971) pp. 50-62.
- 2.. G.B. Anderson and T.S. Hunag "Piecewise Fourier Transformation for Picture Bandwidth Compression," IEEE Trans. on Communication Technology, Vol. COM-19, No. 2 (April, 1971) pp. 133-140.
- 3.. W.K. Pratt, J. Kane, and H.C. Andrews "Hadamard Transform Image Coding," Proc. IEEE, Vol. 57, No. 1 (January, 1969) pp. 58-68.
4. L.C. Wilkins and P.A. Wintz "Bibliography on Data Compression, Picture Properties, and Picture Coding," IEEE Trans. on Information Theory, Vol. IT-17, No. 2 (March, 1971) pp. 180-197.
- 5.. U.F. Gronemann "Coding Color Pictures," M.I.T. Research Laboratory of Electronics, Technical Report 422, (June 17, 1964).
6. R. Schaphorst "Frame-to-Frame Coding of NTSC Color TV," Symposium on Picture Bandwidth Compression, M.I.T., Cambridge, Massachusetts (April, 1969).
7. W.K. Pratt "Application of Transform Coding to Color Images," Symposium on Picture Coding, North Carolina State University, Raleigh, North Carolina, (September, 1970).
8. J.O. Limb, C.B. Rubinstein, and K.A. Walsh "Coding of Color Pictures by Differential Quantization," Symposium on Picture Coding, North Carolina State University, Raleigh, North Carolina, (September, 1970).
- 9.. D.G. Fink, ed. Television Engineering Handbook, McGraw-Hill, New York (1957).

10. W. K. Pratt "Digital Color Image Coding and Transmission," University of Southern California, Report No. USCEE 403 (June, 1971).
11. W. K. Pratt "Spatial Transform Coding of Color Images," IEEE Transactions on Communication Technology (December, 1971).

3. 2. Image Enhancement

One of the basic problems of image enhancement is to diminish the effect of image noise arising from the optical or electronic sensor that generated the image. If the image noise can be statistically characterized, then it is possible to apply Wiener filtering techniques of signal estimation for image enhancement.

Wiener filtering is a classical technique of signal estimation that has been applied, primarily, to one dimensional, continuous signals, with analysis and implementation based upon continuous Fourier signal theory [1]. It is possible, of course, to perform Wiener filtering operations on time sampled signals, and extend the technique to two dimensions. Furthermore, the filtering operation can be implemented by any unitary transformation, rather than the Fourier transform [2-4]. Finally, it is possible to significantly reduce the computational requirements without severely affecting performance by a technique of selective computation.

One Dimensional Filtering. - Figure 1 is a block diagram of a generalized one dimensional Wiener filtering system. A zero mean, M element, data column vector, f , composed of additive and uncorrelated zero mean signal, s , and noise, n , components is the input to the system. The signal and noise are assumed uncorrelated.

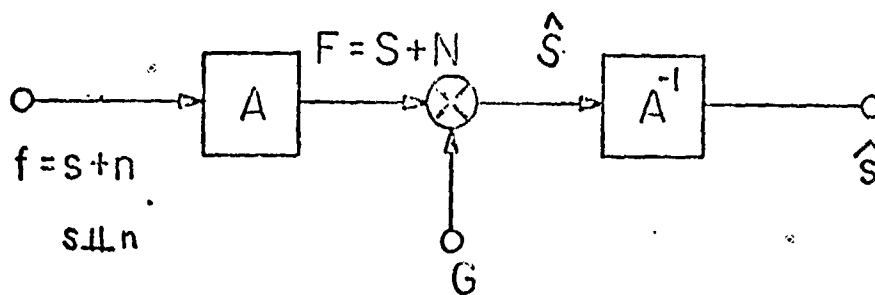


FIGURE 1. Generalized Wiener Filtering

A unitary transform operation is performed on the input yielding

$$F = Af = As + An \cong S + N \quad (1)$$

Table 1 defines several unitary transforms that have been used for digital signal processing applications. Next, the transformed input vector is multiplied by the filter matrix, G , and an inverse unitary transform operation is performed. The resultant

$$\hat{S} = A^{-1}GF = A^{-1}GAf \quad (2)$$

is considered the estimate of the signal, with the filter matrix, G , chosen so that the mean square error between the signal and its estimate is minimized.

Statistical Representation. - For many classes of data, a data vector f can be considered a sample of a random process of known mean \bar{f} and with a covariance matrix

$$C_f = \overline{(f - \bar{f})(f - \bar{f})^*}^T \quad (3)$$

The covariance matrix of a unitary transform of data vector, $F = Af$, likewise is given by

$$C_F = \overline{(F - \bar{F})(F - \bar{F})^*}^T \quad (4)$$

which reduces to [5]

$$C_F = AC_f A^*T = AC_f A^{-1} \quad (5)$$

There is no general closed form solution for eq. (5) except for special cases. However, it should be recognized that eq. (5) is simply a separable two dimensional unitary transform of C_f .

Generalized Wiener Filtering. - In the design of the generalized Wiener filtering system of Figure 1 a filter matrix, G , is chosen to minimize the mean square error

$$e = \text{Tr} \left\{ \overline{(s - \hat{s})(s - \hat{s})^*}^T \right\} \quad (6)$$

between the signal, s , and its estimate, \hat{s} . The expression for the m. s. e. can be rewritten in terms of transform domain quantities as

$$e = \text{Tr} \overline{(A^{-1}S - A^{-1}G(S+N))(A^{-1}S - A^{-1}G(S+N))^*}^T \quad (7)$$

Making use of the fact that

$$\begin{aligned} \overline{SS^*}^T &= C_S = AC_s A^*T \\ \overline{NN^*}^T &= C_N = AC_n A^*T \\ \overline{SN^*}^T &= NS^*T = 0 \end{aligned} \quad (8)$$

one obtains

$$e = \text{Tr} \left\{ C_S - 2GC_S + GG(C_S + C_N) \right\} \quad (9)$$

Now, a straightforward minimization of eq. (9) yields the optimum filter matrix

$$G_O = C_S (C_S + C_N)^{-1} \quad (10)$$

Thus, the optimum filter matrix may be determined from the transform spectral densities of the signal and noise vectors. Alternatively, the optimum filter can be found by a two dimensional transformation

$$G_O = A g_O A^{*T} \quad (11)$$

of a matrix

$$g_O = C_s (C_s + C_n)^{-1} \quad (12)$$

based upon the covariance matrices of the signal and noise. The matrix g_O will be called the response matrix, in analogy with the impulse response of a linear system.

Substitution of the optimum filter matrix into eq. (9) gives the minimum m. s. e.

$$e_{\min} = \text{Tr} \left\{ C_S C_N (C_S + C_N)^{-1} \right\} \quad (13)$$

Alternatively, the m. s. e. can be written in terms of the data domain covariance matrices as

$$e_{\min} = \text{Tr} \left\{ C_s C_n (C_s + C_n)^{-1} \right\} \quad (14)$$

Note that from eq. (14), the minimum m. s. e. is independent of the type of unitary transform employed. However, the character of the filter matrix, G , is dependent upon the unitary transform. The conclusion is that one is free to choose the transform that will minimize the computational processes entailed in filter generation and filter operation.

Table 2 lists the number of multiplications required to perform the Wiener filtering operations of eq. (2). For long vector lengths the multiplication requirements of the Fourier and Hadamard transforms approach those of the identity operator. The Karhunen-Loeve transform requires more than twice as many multiplications as the identity operator. The computational requirements for filter generation are considered in the next section.

Transform	Approximate number of multiplication/addition operations
Identity	M^2
Karhunen-Loeve	$2M^2 + M$
Fourier/Hadamard	$M^2 + 2M \log_2 M$

$M = 2^m$ where m is an integer

Suboptimal Wiener Filtering. - When considering only the amount of computation required to perform the Wiener filtering operation, as opposed to the computation requirements for the filter generation, it is obvious that the bulk of the computation for the Fourier, Hadamard, and identity transform systems can be attributed to the filter matrix multiplication operation. The Karhunen-Loeve transform system only requires a scalar filter multiplication, but entails two matrix multiplication operations for the transformations. As a compromise, consideration has been given to the development of filter matrices that contain a relatively large number of zero entries. With such matrices the filter operation could be performed with a reduced number of computations. The goal in the filter design, of course, is to maintain the m. s. e. performance as close to the optimum filter level as possible.

The suboptimal Wiener filtering design problem can be formulated as an optimization problem whereby the filter matrix, G , is chosen to minimize the m. s. e.

$$e = \text{Tr} \left\{ C_S - 2GC_S + GG(C_S + C_N) \right\} \quad (15)$$

under the constraint that certain selected elements of G are zero. The zero elements of G are so placed that fast computational algorithms can be employed in the filter matrix multiplication.

As a special case of this technique consider the filter design when the filter matrix is constrained to be a diagonal matrix. The optimum filter is found to be [3]

$$G = \text{diag} \left[\frac{C_S(i, i)}{C_S(i, i) + C_N(i, i)} \right] \quad (16)$$

and the resultant m. s. e. is

$$e = N \sum_{i=1}^N \frac{C_S^2(i, i)}{C_S(i, i) + C_N(i, i)} \quad (17)$$

Conceivably, one could determine the filter design when the filter matrix contains two elements per row, and continue with additional matrix elements until the desired m. s. e. is achieved. Unfortunately, the optimization problem rapidly becomes quite complex.

One alternative to the optimization is to use selected elements of the optimum vector Wiener filter matrix as the suboptimal filter matrix. For example, only the diagonal and near diagonal terms of the optimum filter matrix could be retained. Another technique is to select only those terms of large magnitude. For a magnitude threshold filter it has been found that the identity filter performs quite poorly compared to the other transforms because of its poor energy compaction properties. Also, the performance of the Fourier and Hadamard transforms for as few as about ten percent non-zero terms is quite close to the optimum performance.

Image Enhancement Applications:

The generalized Wiener filtering technique has been applied to the enhancement of images as a test of its validity. In this application each image line and column has been characterized as a Markov process with additive white noise. The image was divided into 16 by 16 element blocks for computational simplicity. Each block was filtered using an optimal Hadamard domain Wiener filter and a Wiener filter with the smallest 75% of the terms removed. In both cases the noise effects were substantially reduced compared to the original.

References

1. W. B. Davenport, Jr. and W. L. Root, Random Signals and Noise, McGraw-Hill, New York, 1968.
2. J. Pearl, "Application of Walsh Transform to Statistical Analysis," IEEE Trans. on System Science and Cybernetics, April, 1971.
3. J. Pearl, "Walsh Processing of Random Signals," Symposium on Applications of Walsh Functions, Washington D. C., April, 1971.
4. W. K. Pratt, "Linear and Nonlinear Filtering in the Walsh Domain," Symposium on Application of Walsh Functions, Washington, D. C., April, 1971.
5. W. K. Pratt, "A Comparison of Digital Image Transforms," Mervin J. Kelly Communications Conference, University of Missouri at Rolla, September, 1970.

⁴3. Image Detection

The research results fall into the five major categories listed below:

- a. Detection and Multidimensional Rotation
- b. Detection and the Minimum Spanning Tree
- c. Image Similarity
- d. Variance of Bayes Estimates
- e. Detection and Lateral Inhibition

The first two topics can be best described as feature selection mechanisms for dimensionality reduction in the image detection and classification problem.

In these cases, an N dimensional vector space approach is developed and utilized to obtain significant discriminatory features for input to a decision algorithm for detection purposes. The third approach is one of developing a measure of image similarity. Such a measure will then allow a decision to be made based upon a degree of similarity between images. This could be considered in analogy with the two dimensional matched filter. The fourth result includes some statistical analyses developing new properties of the variance of Bayes estimates. While not directly couched in the notation of imagery, the results, nonetheless, have bearing in detection problems utilizing Bayesian Estimators. The fifth topic is one in which the concept of lateral inhibition is utilized in an attempt to obtain a linear transformation which best implements the inhibition process for image detection.

DETECTION AND MULTIDIMENSIONAL ROTATIONS

In the context of an image detection and recognition environment, it becomes immediately obvious that for machine implementation, the dimensionality of the problem is far too large. Specifically, if one analyses the N^2 degrees of freedom which comprise an $N \times N$ sampled image, and attempts to develop decision algorithms in the N^2 dimensional space defined by the image, computational power suddenly becomes a premium. In order to reduce the dimensionality of the problem to a set of features which retain detection significance while removing those dimensions which provide little or nothing to the decision process, it is necessary to consider the image as a point in N^2 dimensional space. Then by appropriate transformations, be they linear or nonlinear, it will be possible to retain a fewer set of coordinates or degrees of freedom while maintaining detection integrity.

One method of particular interest for dimensionality reduction is a linear transformation provided by unitary matrices and often referred to as multidimensional rotations. Five such rotations were investigated as to their potential for discriminating significant features in an image detection problem involving the recognition of hand written characters. The five rotations of interest were the Karhunen-Loeve, Fourier, Walsh, Haar, and Identity. The results of the experiment suggested that large dimensionality reduction factors were indeed possible with little or no degradation in detection performance and the order of the most efficient rotations is as listed above. The papers at the end of this report present the details and published results on this topic.

DETECTION AND THE MINIMUM SPANNING TREE

The use of multidimensional rotations discussed above is restricted to linear techniques and will not cover the cases where dimensionality reduction requires nonlinear operations. As an example, consider a class of images in N^2 space whose intrinsic discriminatory dimensionality lies on a nonlinear surface undetectable by linear rotations. Specifically, data points on a sphere are intrinsically two dimensional although distributed in 3 space. In order to be able to handle such situations, (which are more likely than linear intrinsic dimensions) nonlinear recursive search techniques must be utilized. Research in this direction is currently underway where considerable emphasis is placed on the concept of the minimum spanning tree (MST). The MST is a unique connection of images (points in N^2 space) such that they are linked with the shortest total distance. The tree has the potential for maintaining local measures at the expense of global properties and, thus, provides a clue to measuring intrinsic dimensionality. As research on this topic was just initiated it is too early to present conclusive results, but the subject is under current investigation with indications of developing a promising nonlinear discriminatory feature selector for image detection.

IMAGE SIMILARITY

A new measure of similarity between images has been developed which is proposed as a useful detection criterion. The measure has information theoretic properties and, with proper image normalization, is the entropy function. The measure is convex upward reaching a peak when two images are identical and thus perfect image detection occurs at the maximum of this image similarity measure. The results of this research are well documented in the papers at the end of this report. However, in addition to providing a similarity measure, techniques for

dealing with magnifications and rotations are also discussed in that document and provide a type of prenormalization as input to the similarity measure.

VARIANCE OF BAYES ESTIMATES

This portion of the image detection research concerned itself with developing tighter bounds on the estimates obtained from Bayesian techniques. Specifically, exponential bounds are obtained and provide confidence measures as more and more samples are taken. The results of the research are far more widespread than just applying to image detection. In fact, any Bayesian estimator falls within the bounds obtained in this research. The paper describing this work is included in the papers section and goes into considerable detail on the subject.

DETECTION AND LATERAL INHIBITION

The contents of the research on this subject examines neural net properties, in particular lateral inhibition as a possible mechanism for extracting features important to image detection. Spatial filters have been investigated which emphasize edges in a pattern. Biologically, this spatial process is called lateral inhibition in that a neuron inhibits surrounding neurons directly proportional to its firing frequency and inversely proportional to distance. Lateral inhibition is a spatial filter that emphasizes edges or boundaries between contrasting light levels and may play a part in identifying objects. Linear transformations were developed mathematically to model such processes and extensive biological and engineering literature searches were undertaken to solidify the model. There is still controversy over whether a visual pattern is recognized as a whole or whether it is perceived through the combination of several elements or features. Just how meaningful features are extracted

from incoming sensory information is not known, except for lateral inhibition which enhances edges or boundaries.

5 A. Linear Image Transformation

Linear transformations of picture segments play a role in bandwidth compression, image enhancement and pattern recognition. The two dimensional structure of the data leads to rather large linear spaces and makes computation of a transform via its describing matrix impractical. A square picture segment, whose linear dimension is N , has N^2 values and the computation of a linear transformation by matrix multiplication requires N^4 multipliers. Clearly faster algorithms are needed.

The Fourier transform and Hadamard transform each possess algorithms which require $M \ln_2 M$ operations on M points. However, these are only two of a large class, called group character transforms, which have fast algorithms.

The above algorithms are most efficient when M is a power of 2. This class can be further generalized as follows: Let L_N be the matrix of a linear transform in N dimensional space which has a fast algorithm. If A is a $2N$ by $2N$ matrix with at most λ non-zero entries in each row then

$$L_{2N} = A \begin{pmatrix} L_N & 0 \\ 0 & L_N \end{pmatrix}$$

has a fast algorithm: If V_1, V_2 are of dimension N and W_N is the number of additions required to compute $L_N V_i$ then

$$L_{2N} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = A \begin{bmatrix} L_N V_1 \\ L_N V_2 \end{bmatrix}$$

requires $2W_N + 2(\lambda - 1)N$ additions. Repetitive use of this structure yields algorithms whose computation time grow as $N \ln_2 N$. Such constructions can be used to design transforms with fast algorithms and some additional property. For instance, the following modification of the Hadamard construction produces a matrix with the linear function $\{ah + b : h = 1, \dots, N\}$ as one of its rows.

Let

$$L_2 = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

$$a_1 = 1$$

and define recursively for N a power of 2 with I_N as N by N Identity matrix

$$L_{4N} = \begin{pmatrix} I_N & 0 & I_N & 0 \\ 0 & I_N & 0 & -I_N \\ -a_{2N} I_N & b_{2N} I_N & a_{2N} I_N & b_{2N} I_N \\ b_{2N} I_N & -a_{2N} I_N & -b_{2N} I_N & -a_{2N} I_N \end{pmatrix} \begin{pmatrix} L_{2N} & 0 \\ 0 & L_{2N} \end{pmatrix}$$

$a_{2N} = \frac{2a_N}{\sqrt{1+4a_N^2}}, b_{2N} = \frac{1}{\sqrt{1+4a_N^2}}$

The $2N^{\text{th}}$ row of L_{4N} is the linear function. The matrix also has the property that $\frac{1}{\sqrt{N}} L_N$ is an orthogonal matrix.

Whether or not this class of transforms is useful for picture analysis has not yet been studied.

The Karhunen-Loeve transform plays an important role in picture analysis but does not have a fast algorithm. This transform is defined as follows. Let V be an N -dimensional random vector and compute $R_{ij} = E \{v_i v_j\}$, where $E\{ \}$ denotes ensemble average (expected value). Let the rows of L_N be the unit eigenvectors of R in order of decreasing values of corresponding eigenvalue. This transform has the property that most of the "energy" of the random vector is in the first few coordinates. When the vector is formed by scanning the elements of a picture in linear order, the eigenvectors tend to be sines and cosines of increasing frequency. The frequencies are not integer multiple of a fundamental, so the fast Fourier transform does not apply. However, with the exception of the first few terms, cosines and sines alternate as eigenvectors and the frequencies differ by one half the fundamental. Thus, the transform can be approximated by two FFT's with some extra work involving projection of the vector onto the first few rows of the matrix. A fast algorithm for the inverse transform has not yet been found.

Transform Implementation

The Hadamard transform, requiring no multiplications, seem to be the simplest for which to design a hardware implementation. Let

(U_0, \dots, U_{2^n-1}) be the Hadamard transform of (V_0, \dots, V_{2^n-1}) . Then

$$(1) \quad U_i = \sum_j (-1)^{(i,j)} V_j$$

where

$$i = i_1 + 2i_2 + \dots + 2^{n-1}i_n$$

$$j = j_1 + 2j_2 + \dots + 2^{n-1}j_n$$

and

$$(i,j) = \sum_{n,n} i_n j_n$$

The above equation can be written to exhibit the structure of its fast algorithm.

$$(2) \quad U_i = \sum_{j_n=0,1} (-1)^{j_n i_n} \sum_{j_{n-1}} (-1)^{j_{n-2} i_{n-2}} \dots \sum_{j_1=0,1} (-1)^{j_1 i_1} V_j$$

The recursion described by

$$(3) \quad S_0(j_1, \dots, j_n) = V_j$$

and

$$S_k(i_1, \dots, i_k, j_{k+1}, \dots, j_n) = S_{k-1}(i_1, \dots, i_{k-1}, 0, j_{k+1}, \dots, j_n)$$

$$+ (-1)^{i_k} S_{k-1}(i_1, \dots, i_{k-1}, 1, j_{k+1}, \dots, j_n)$$

computes the successive summations and S_n is the transformed sequence:

$$U_i = S_n(i_1, \dots, i_n), \quad i = \sum_{e=1}^n i_e 2^{e-1}$$

If the S_k are regarded as functions of a single variable, time, running from 0 to $2^n - 1$, equation (3) becomes

$$S_k(t) = S_{k-1}(t) + S_{k-1}(t+2^{k-1}) \quad \text{for } t_k = 0$$

$$S_k(t) = S_{k-1}(t-2^{k-1}) - S_{k-1}(t) \quad \text{for } t_k = 1$$

where t_k is the coefficient of 2^{k-1} in the dyadic representation of t . The second of these equations ($t_k = 1$) may be replaced by

$$S_k(t) = S_k(t-2^{k-1}) - S_{k-1}(t)$$

These equations then lead directly to the logical diagram of Figure 1.

When two consecutive stages are connected, the final delay of the k^{th} stage may be incorporated into the initial delay of the $(k+1)^{\text{st}}$ stage. Therefore, a transform on 2^n points requires $2^n + 2^{n-1}$ units of delay, n adders and the associated gating circuits. It is estimated a sixteen point transform with eight-bit input samples and twelve-bit output samples, using medium scale integrated circuits at five megahertz would require less than \$1500 in components.

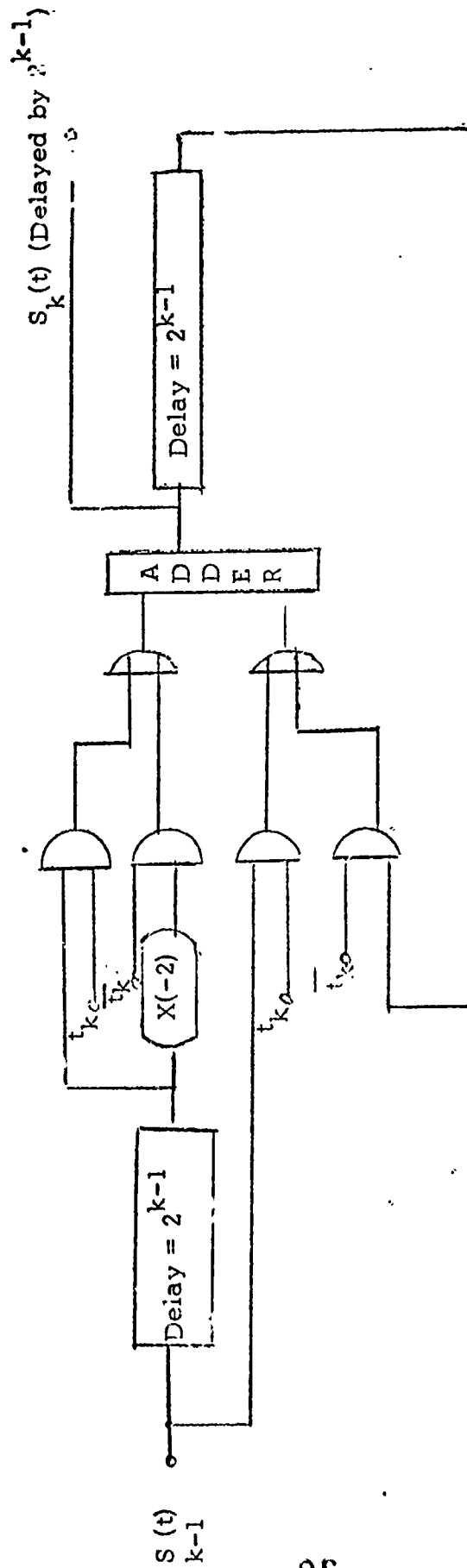


Figure 1. Block diagram of Hadamard transform computation

6. LIST OF PAPERS AND PRESENTATIONS SUPPORTED BY THE GRANT

1. Andrews, H. C., "Some Unitary Transformations in Pattern Recognition and Image Processing," IFIP Congress '71, Ljubjana, Yugoslavia (August, 1971).
2. Andrews, H. C., "Orthogonal Transforms and Feature Selection in Pattern Recognition," 4th Annual Hawaii International Conference on Systems Science, PP. 675-677. (January, 1971).
3. Andrews, H. C., "Walsh Functions in Image Processing, Feature Selection and Pattern Recognition," IEEE Transactions on Electromagnetic Compatibility, Vol. EMC-13, PP. 26-32. (August, 1971).
4. Liporace, L. A., "Variance of Bayes Estimates," accepted for publication in the IEEE Transactions on Information Theory.
5. Liporace, L. A. and I. S. Reed, "An Information Theoretic Measure of Image Similarity," submitted for publication.
6. J. Pearl, H. C. Andrews and W. K. Pratt, "Performance Measures for Transform Data Coding," submitted for publication.
7. W. K. Pratt, "Linear and Nonlinear Filtering in the Walsh Domain," Proc. of Conference on Application of Walsh Functions, Washington, D.C., (April, 1971).
8. W. K. Pratt, "Fast Computational Techniques for Generalized Two Dimensional Wiener Filtering," Proc. Two Dimensional Signal Processing Conference, University of Missouri-Columbia, (October, 1971).
9. W. K. Pratt, "Basic Considerations for Color Image PCM Coding," 1971 Picture Coding Symposium, Purdue University, (October, 1971).
10. W. K. Pratt, "Color Image Transform Coding - A Progress Report, 1971," 1971 Picture Coding Symposium, Purdue University, (October, 1971).
11. W. K. Pratt, "Spatial Transform Coding of Color Images," IEEE Transactions on Communications Technology, (December, 1971).
12. W. K. Pratt, "Binary Symmetric Channel Error Effects on PCM Color Image Transmission," submitted for publication.
13. W. K. Pratt, "Generalized Wiener Filtering Computation Techniques," submitted for publication.