XVI. COGNITIVE INFORMATION PROCESSING*

Academic and Research Staff

Prof. S. J. Mason
Prof. M. Eden
Prof. F. F. Lee
Prof. W. F. Schreiber
Prof. D. E. Troxel
Prof. W. L. Henke
Prof. T. S. Huang
Prof. K. Ramalingasarma
Dr. W. L. Black
Dr. K. R. Ingham
Dr. P. A. Kolers
Dr. O. J. Tretiak
F. X. Carroll
C. L. Fontaine
J. E. Green
E. R. Jensen
J. Laurino
R. L. Rees
Sandra A. Sommers
Y. D. Willems

Graduate Students

G. B. Anderson
T. P. Barnwell III
B. A. Blesser
A. Gabrielian
R. E. Greenwood
E. G. Guttmann
R. V. Harris III
D. W. Hartman
A. B. Hayes
A. N. Kramer
W-H. Lee
C-N. Lu
J. I. Makhoul
G. P. Marston III
O. R. Mitchell, Jr.
S. L. Pendergast
D. R. Pepperberg
G. F. Pfister
R. S. Pindyck
D. S. Prerau
J. E. Richards, Jr.
G. M. Robbins
K. B. Seifert
C. L. Seitz
D. Sheena
D. R. Spencer
W. W. Stallings
P. L. Stamm
R. M. Strong
G. A. Walpert
J. W. Woods
I. T. Young

A. PHOTOCHROMIC SPATIAL FILTERING

Photochromic materials change color spontaneously and reversibly when irradiated with light of appropriate wavelength. Such materials may be used for storage of pictorial information in the same way as ordinary photographic films are used. Like photographic films, photochromics offer the advantages of high-density information storage, high readout speed, and low cost. Photochromics, however, have three major advantages over photographic (silver halide) films.

1. The stored information may be viewed immediately, without the need for development.
2. The information can be erased and the material reused.
3. The potential resolution of photochromic film is better than the resolving capability of modern optics.

The "instant development" advantage of photochromic material may be used to construct a real-time optical correlator that is usable with EDP systems. A matched spatial filter may be generated on photochromic film for an arbitrary input image or object.

*This work was supported principally by the National Institutes of Health (Grants 5 PO1 GM-14940-02 and 5 PO1 GM-15006-02), and in part by the Joint Services Electronics Programs (U.S. Army, U.S. Navy, and U.S. Air Force) under Contract DA 28-043-AMC-02536(E).

The output of the system using the matched filter yields an optimum indication of the presence of the signal in noise.

Spatial filtering systems using conventional silver halide recording media as filters have been constructed before.[1-3] The object of this research was to demonstrate the feasibility of using photochromic materials. The major disadvantage of photochromics is their low sensitivity as compared with silver halide films. A Spectra Physics 125 laser with a nominal output of 50 mW at 6328 Å was used to optically bleach a matched filter into pre-darkened American Cyanamid 43-540 sheet film. With exposure times of 45-90 sec, successful reconstructions and correlation functions were generated for simple letters and words. When combined with an input device described by Stetten,[4] a real-time optical correlator might be designed which would be compatible with digital EDP systems. This report is based on work done for an S. M. thesis.[5]

S. L. Pendergast

References

1. L. J. Cutrona, E. N. Leith, C. J. Palermo, and L. J. Porcello, "Optical Data Processing and Filtering Systems," IRE Trans., Vol. IT-6, pp. 386-400, 1960.
2. A. Vander Lugt, "Signal Detection by Complex Spatial Filtering," IEEE Trans., Vol. IT-10, pp. 139, 1964.
3. V. V. Horvata, J. M. Holeman, and C. Q. Lemmond, "Holographic Technique Recognizes Fingerprints," Laser Focus, Vol. 3, No. 11, June 1964.
4. K. J. Stetten, "Real Time Photochromic Projection Display," a paper presented at 4th National Symposium, Society for Information Display, Washington, D. C., 1964.
5. S. L. Pendergast, "Applications of Photochromic Materials in Optical Data Processing," S. M. Thesis, Massachusetts Institute of Technology, 1968.

B. DETECTING PHASE INFORMATION IN OPTICAL FOURIER TRANSFORMS

When an object transparency is placed at the front focal plane of a lens and illuminated by collimated coherent light, the distribution of light at the back focal plane can be described by a properly scaled Fourier transform of the object. The Fourier transform is, in general, complex. Any measuring device or photographic plate responds to the square of the magnitude of the light and the phase information is lost. Thus if the light at a point (ξ, η) is $A(\xi, \eta)\, e^{i\phi(\xi, \eta)}$, the photographic emulsion responds to $|A(\xi, \eta)|^2$. In this report, we describe a method of measuring the phase.

If a point of reference light is introduced into the object plane at (X_o, Y_o), it is described by $B\delta(x-X_o)\, \delta(y-Y_o)$, where B is a real constant, and δ is the Dirac delta function. The combined effect on the photographic plate at (ξ, η) is

$$\left| A(\xi,\eta)\, e^{i\phi(\xi,\eta)} + B\, e^{-i\frac{2\pi}{\lambda D}(\xi X_o + \eta Y_o)} \right|^2$$

which we shall call $C^2(\xi,\eta)$.

If three separate exposures are made at the back focal plane of the lens: one of the object alone, one of the object and the reference beam acting simultaneously, and one of the reference alone, then, we have

$$\frac{C^2(\xi,\eta) - A^2(\xi,\eta) - B^2}{2A(\xi,\eta)\, B} = \cos\theta, \tag{1}$$

where

$$\theta = \left[\phi(\xi,\eta) - \frac{2\pi}{\lambda D}(\xi X_o + \eta Y_o) \right].$$

Now $C^2(\xi,\eta)$, $A^2(\xi,\eta)$, and B^2 can all be measured, and X_o and Y_o are known, therefore the value of cos θ can be determined.

Now, if we make a fourth exposure of the object with the reference point, putting a quarter-wave plate right next to the reference point, we shall record on film:

$$D^2(\xi,\eta) = \left| A(\xi,\eta)\, e^{i\phi(\xi,\eta)} + iB\, e^{-i\frac{2\pi}{\lambda D}(\xi X_o + \eta Y_o)} \right|^2.$$

And we can determine sin θ by

$$\frac{D^2 - A^2 + B^2}{2AB} = \sin\theta.$$

From cos θ and sin θ, we can determine θ, and hence $\phi(\xi,\eta)$. Some preliminary experiments[1] have been done in calculating cos θ from measured values of C^2, A^2, and B^2.

In doing this one encounters immediately the problem of measuring intensity. Densitometry is difficult and often inaccurate. If one uses 649F photographic plates, however, there is a range of exposure for which the relationship between exposure and transmittance is linear.[2] If plate I is a recording of $A^2(\xi,\eta)$, then the transmittance of I is

$$T_a(\xi,\eta) = KA^2(\xi,\eta)\, t,$$

where t is the exposure time, and K is some constant. The results are similar for the two other plates. Let T_b be the transmittance of the plate made with the reference point alone, and T_c that of the plate made with the object together with the reference

point. If all three exposures are for the same length of time, and if all three fall within the linear region, then

$$\frac{C^2(\xi,\eta) - A^2(\xi,\eta) - B^2}{2A(\xi,\eta)\,B} = \frac{T_c(\xi,\eta) - T_a(\xi,\eta) - T_b(\xi,\eta)}{2\sqrt{T_a(\xi,\eta)\,T_b(\xi,\eta)}}. \tag{2}$$

For details of this procedure, see the author's thesis.[1]

It remains, then, to measure the transmittances of the three plates. The transmittances were measured on the new CIPG scanner. This instrument is capable of producing approximately 1000 resolvable points per line, so a raster of 1024×1024 points was used.

Steps must be taken to ensure that corresponding points of the three plates are scanned. In this project, a registration plate was placed in front of the photographic plate during exposure. The registration plate consisted of two opaque dots on a clear background, so that two small shadows were cast onto the plate, resulting in two points of high transmittance on the developed plates. If a television camera tube such as a vidicon could be used to read the light intensity at the back focal plane of the lens, then one would not have a severe registration problem.

In order to ensure that spatial frequencies were low enough to be resolved by the scanner, the objects to be Fourier-transformed were small. The objects were photoreduced onto Kodak "Kodalith" film. A black point was placed near the object before photoreduction to provide a transparent point for the reference beam. During exposure, light had to be concentrated on the point with a lens and iris to bring it to the proper intensity with respect to the main beam. Various relative intensities were tried.

The three exposures resulting from the object alone, the object with the reference beam, and the reference beam alone, were made and developed together in D-19 at 20°C for 5 minutes with identical agitation, and then scanned. The magnetic tape containing the transmittances was then processed on the IBM 360 computer. The transmittance of each point was quantized to 256 levels. The registration point shadows were found, and beginning points for each plate were determined. Finally, for each point, the calculation of Eq. 2 was performed, and cos θ obtained.

Three objects were used: (i) two points, (ii) a square, and (iii) the letters "MIT". The experiments indicated that the best results were obtained near, but not at the center, of the plate. Near the edges, transmittance was high, and the background noise made the results meaningless. Near the center, the transmittances were often so low that the linearity condition necessary to obtain Eq. 2 from Eq. 1 was no longer valid. Elsewhere, however, the results were reliable. The phase of the center point could be measured by shortening the exposure time, and knowledge of the intensity at the edges could be obtained by lengthening the exposure time. By this method, then, the phase

information in a Fourier transform can be determined.

K. B. Seifert

References

1. K. B. Seifert, "Phase Information Detection in a Fourier Transform Hologram," S.M. Thesis, Department of Electrical Engineering, M.I.T., September 1968.

2. A. A. Friesem, A. Kozma, and G. F. Adams, "Recording Parameters of Spatially Modulated Coherent Wavefronts," Appl. Opt. 6, 851-856 (1967).

C. VIDEO BANDWIDTH COMPRESSION BY LINEAR TRANSFORMATION

In the redundancy reduction of pictures, making use of the high correlation of adjacent samples, it is necessary to code blocks of picture elements to obtain an efficient encoding. Rather than code blocks of data, the coding may be done independently if the data are first transformed so as to be independent. A necessary condition for independence is that the variables be uncorrelated. To find the exact nonsingular matrix that will uncorrelate a block of data is computationally difficult. Thus it was decided to investigate a class of matrices that might approximately uncorrelate data.

Transformation by Hadamard matrices has been used successfully for the redundancy

x_1	x_2	...	x_N
x_{N+1}	x_{N+2}	...	x_{2N}
⋮	.	..	
$x_{N(N-1)+1}$	...	...	x_{N^2}

Fig. XVI-1. Positions of x_i.

reduction of vocoder data.[1] Thus it was suggested[2] that this method might prove useful for the reduction of video data.

Assume that a black and white picture has been sampled on a regular square lattice so that it is a collection of random variables $\{f_{ig}\}$. Divide the picture into squares of side N. Then form a column vector $\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ . \\ \vdots \\ x_{N^2} \end{bmatrix}$, where the positions of x_i are shown in Fig. XVI-1. Let $\underline{y}$ be the N^2 dimensional random variable resulting from a

Hadamard (H) transformation on $\underline{x}$, that is, $\underline{y} = \frac{1}{N} H\underline{x}$, where H is a symmetric matrix of +1's and -1's whose rows are orthogonal. For the two-dimensional case H is taken as $H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. For four dimensions $H_4 = \begin{bmatrix} H_2 & H_2 \\ H_2 & -H_2 \end{bmatrix}$, etc. For 2^n dimensions H is taken to be built up in this fashion.Thus these matrices have the property that they may be partitioned successively in a manner analogous to the Discrete Fourier Transform. As a result only $N \log_2 N$ operations are required for an N-dimensional Hadamard Transform. (In fact this is a decomposition into square waves instead of sine waves.) For $N^2 = 64$ and a 256×256 picture, there are $(32)^2 = 1024$ squares and hence 1024 Hadamard transforms. A transform and an inverse (plus read and write time from tape) takes approximately $2\frac{1}{2}$ minutes on the IBM 360/ computer. No multiplications need be performed, just additions and subtractions. Thus the transform may be implemented by a simple interconnection of wires.

It has been determined that if $R_{\underline{x}}$, the correlation matrix of $\underline{x}$ is of the type

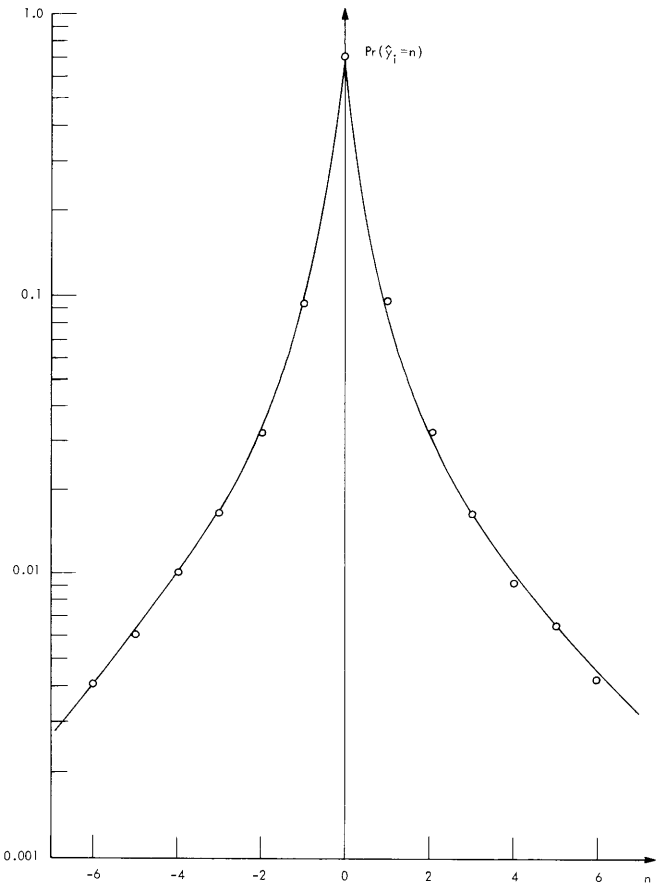


Fig. XVI-2. Experimental distribution for $i \geq 2$.

$$R_{\underline{x}} = \begin{bmatrix} 1 & a & b & c \\ a & 1 & c & b \\ b & c & 1 & a \\ c & b & a & 1 \end{bmatrix}$$

then the y_i will be uncorrelated. Owing to the high correlation of neighboring samples in pictures, as long as N does not get larger than ~4 for a 256 × 256 picture, this is approximately the correct correlation matrix (mainly relying on the high correlation of all of the neighboring elements out to distance 4). For N = 8, the approximation is not as good, but considerable redundancy reduction is obtained in addition to some welcome averaging effects.

Such a scheme of separately coding the y_i would be optimum with respect to mean-square error if the y_i were uncorrelated and jointly Gaussian and thus independent. For N = 8, however, the experimental distribution of the $\hat{y}_i$ $i \geq 2$ appears to be negative exponential $e^{-\rho|n|}$ for $|n| \leq 3$ on a 7-binit scale. A 7-binit scale on the y_i is sufficient to recover a picture looking almost as good as the original. Figure XVI-2 shows a representative experimental distribution for $i \geq 2$. $\hat{y}_1$ is just the average value over the square and hence does not have this type of distribution.

For N = 8, the cameraman picture was transformed and uniformly quantized in varying degrees of coarseness. The zeroth order (assuming the $\hat{y}_i$ are independent) entropy was calculated for each of the $\hat{y}_i$ and averaged. Figure XVI-3a shows the inverse transform of the y_i quantized on an 8-binit scale, the entropy $H \approx 2.2$ bits. Figure XVI-3b is quantized on a 6-binit scale, $H \approx 0.76$ bits. Figure XVI-3c is quantized on a 5-binit scale, $H \approx 0.47$ bits. The appearance of Fig. XVI-3b and 3c could probably be improved by the application of the pseudo-random noise technique due to Roberts. The pictures in Fig. XVI-4 were obtained by quantization of the y_i on a restricted scale. Figure XVI-4a shows the result of quantizing $\hat{y}_1$ on a 6-binit scale, and $\hat{y}_2$ through $\hat{y}_{64}$ quantized to ±1 on a 6-binit scale. Figure XVI-4b quantizes $\hat{y}_1$ on a 6-binit scale and $\hat{y}_2$ through $\hat{y}_{64}$ quantized to ±1 and 0 on a 6-binit scale. Figure XVI-4c shows the result of quantizing $\hat{y}_1$ on a 7-binit scale and $\hat{y}_2$ through $\hat{y}_{64}$ to 0, ±1, ±4 on a 7-binit scale. It is seen that the occasional excursions of the y_i $(i \geq 2)$ greater than 3 or 4 represent rather important subjective details, the edges, and thus require transmissions.

This research will continue with investigation of the effect of varying types of quantization of the y_i. The effect of a Robert's type pseudo-random noise technique will be examined. Other things that might be investigated are the variability of the probability distribution of the y_i with different pictures and the effect of using

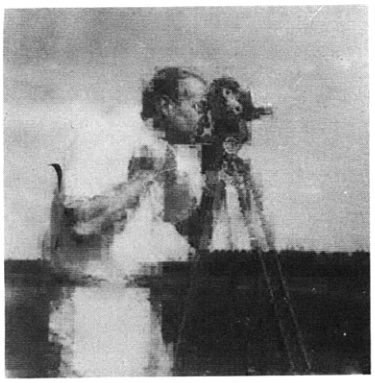
(a)

(b)

(c)

Fig. XVI-3.
(a) 8-binit processed picture.
(b) 6-binit processed picture.
(c) 5-binit processed picture.

(a)



(b)

(c)

Fig. XVI-4.

(a) Processed picture quantized to ±1 (except for y_1).

(b) Processed picture quantized to ±1.

(c) Processed picture quantized to 0, ±1, ±4.

this technique on higher resolution (512 × 512) pictures where the dependence of neighboring picture elements should be stronger.

J. W. Woods

References

1. I. Kramer and G. H. Mathews, "A Linear Coding for Transmitting a Correlated Set of Signals," IRE Trans., Vol. IT-2, pp. 41-46, September 1956.

2. T. S. Huang, "Bandwidth Compression by Linear Transformations," Autonetics Memo, May 1967 (unpublished).

D. ON A PROCEDURE FOR CODING CONNECTED FIGURES IN A PICTURE

Virtually all current research on picture processing begins with the quantization of spatial coordinates and intensity. In most cases the spatial coordinates are quantized so that each picture point is an element of a regular array. In this report we specialize to an N × N square array and to two levels of intensity, but the procedure can be modified trivially for other regular arrays and to any finite number of intensity levels.

A picture is an N × N square array of black points p^b and white points p^w. We say two points are adjacent if and only if they have an edge in common, that is, if they lie next to each other in the array (points on a diagonal are not adjacent). As is customary we call "connectedness" the equivalence relation generated by "adjacency." Hence, in particular, any picture is partitioned into connected sets. We call the black connected sets figures, $f_j(k)$, the set of all white points we call ground.

For the sake of simplicity we shall deal with the case of a single connected black set, although the method admits of extensions to more general cases. We shall describe an algorithm for representing each figure of k points, $f_j(k)$ as a binary word $w(f_j(k))$ of length ℓ, such that $\ell_w(f_j(k)) \leq k - 1 + m$, where m enumerates the white points adjacent to $f_j(k)$.

Choose some $p \in f_j(k)$, call it p_1, and some point adjacent to p_1, call it $a_1(p_1)$. If $a_1(p_1) \in f_j(k)$, define $p_2 = a_1(p_1)$ and assign 1 as the first letter of the code word; otherwise assign 0. Proceed in a clockwise sense around p_1 from $a_1(p_1)$, examining in sequence $a_2(p_1)$, $a_3(p_1)$, $a_4(p_1)$ [if they exist]. For each point a binary letter is assigned as with $a_1(p_1)$ and if $a_j(p_1) \in f_j(k)$, define $p_x = a_j(p_1)$ by an index one more than the index of the most recent element of $f_j(k)$ which has been already ordered.

If no $a_j(p_1) = p_2$, clearly k = 1 and $\{p_1\}$ is itself a figure. Otherwise there is a p_2 and the elements of $\{a_j(p_2)\}$ can now be examined sequentially, once an order among the $a_j(p_2)$ has been fixed.

We call p_1 the predecessor of p_2; $p_1 = p(p_2)$. In general $p_i = p(p_h)$, if $p_h = a_j(p_i)$ and there is no p_α, $\alpha < i$ for which $p_h = a_j(p_\alpha)$. Set $a_1(p_i) = p(p_i)$.

Obviously the figure point $p(p_i)$ has already been represented in the code word and

hence no additional letter need be added for $a_1(p_i)$. We examine $a_2(p_i)$, $a_3(p_i)$, $a_4(p_i)$, making letter assignments and indexing black points clockwise as before, except that if some $a(p_i)$ is also $a(p_\alpha)$ for some $\alpha < i$, then no letter is adjoined to the word nor is a new index assigned.

The procedure continues examining the points adjacent to each p_i in the indexed sequence, in the order of the sequence and with the conditions that $a_1(p_i) = p(p_i)$, the other indices of $a_j(p_i)$ being related to the points adjacent to p_i by the clockwise sense of rotation.

It is easy to prove (by supposing the contrary) that each $p \in f_1(k)$ has some predecessor and hence will be assigned an index. When the 1 corresponding to p_k is adjoined to the code word, the word is complete.

Note that in general, not all of the white points adjacent to $f_j(k)$ will be assigned a 0 in $w(f_j(k))$. Furthermore, $w(f_j(k))$ is a subword of words corresponding to figures of which $f_j(k)$ is a subfigure. Note, too, that there are in general 4k distinct words for each $f_j(k)$, since we have made an arbitrary choice of p_1 and of $a_1(p_1)$.

Finally, it is easy to prove (by contradiction) that no two distinct figures $f_j(k)$, $f_i(k)$ have the same code word. We now exhibit an algorithm for enciphering an arbitrary picture uniquely.

Call the upper left point in the $N \times N$ array p_{11}. If p_{11} is black, it is the first figure point of $f_1(\lambda)$. We assign $a_1(p_{11})$ as the point to the left (even though it does not exist in the array and hence requires no assignment in the code word). We may then proceed with the algorithm described earlier. In this case the code word for $f_1(\lambda)$ terminates when the adjacent points of the last indexed black point in $f_1(\lambda)$ have been examined and no new black points are reached.

If p_{11} is white, then proceed from left to right through the first row, second row, etc., until a black point is reached, enumerating the white points; call this number α_1. Clearly α_1 provides the coordinates of $f_1(\lambda)$. Furthermore, designate the last white point in this string as $a_1(p_1)$ (except if p_1 is a leftmost point in the row in which case as indicated above, assign $a_1(p_1)$ to the left of p_1, even though $a_1(p_1)$ does not exist). Thus we have $\alpha_1, f_1(\lambda)$ which is unique. Continue from the last point of α_1, disregarding any points in f_1, until a black point is reached such that $p^b \notin f_1$; call the enumeration α_2. Obviously this establishes $p_1 \in f_2(\lambda)$ and the coordinates of p_1. Proceed with representing f_2 as we did f_1. The process will terminate when α_j includes the point $p_{N,N}$.

It should be noted that we have constructed a code word for a picture that is made up of alternating sequences α_1, w_1, $\alpha_2 w_2 \cdots \alpha_m w_m \alpha_n$ in which the α's measure run length and hence are most expeditiously considered as words of uniform length. If, for example, $N \times N = 2^{18}$, the α's can be listed in 18-bit memory addresses. In general if there are m figures, their coordinates are established by $2(m+1)\log_2 N$. The maximal length of $w(f_j(k))$, $k < N$, is $3k + 2$ (This is for the case of a figure lying entirely

on one row or one column). The word length is smallest for dense, convex "blobs" and is approximately $k + 4k^{1/2}$. The code for figures is efficient in that a relatively high proportion of all possible sequences of length $\approx 2k$ containing exactly k 1's will represent figures. This is so because bounds on the number of figures with k elements are known; $\frac{3.20^k}{4k} \leq \text{Card}\{f_j(k)\} \leq \frac{6.75^k}{4k}$ (1, 2).

The code should be particularly useful when the ratio of white to black points is high. On the other hand, it will be longer than N^2 when the average size of a figure is close to 1 and hence would be much poorer than no coding at all.

Experiments are planned to explore the efficiency of this code with a variety of real pictorial objects.

M. Eden

References

1. M. Eden, "A Two-Dimensional Growth Process," Proc. Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. IV (Berkeley, California, 1961), pp. 223, 239.

2. D. A. Klarner, "Cell Growth Problems," Can. J. Math., Vol. 19, pp. 851-863, 1967.

E. IMAGE FIELD PARTITIONING

1. Introduction

There are probably two different types of images in a computer. The image may be an array of density measurements, each a multidigit number. This is the raw data for a pattern-recognition program. After some nonlinear operations, the image is thresholded, thereby producing an array of black or white points (1 or 0). The latter image is subjected to shape analysis.

2. Connectivity and Nearest-Neighbor Relations

One of the basic operations in pattern recognition is the identification of connected areas. When one is counting objects, this is a terminal operation, in a more complex shape-discrimination task, it is only an intermediate operation.

Many organizations for this algorithm are possible. One may choose a set of names (i.e., numbers) for the connected sets, and store the names at the array coordinates that each region occupies.[1] We prefer to record each connected set as an interval list. For each connected region we form a set of coordinate triplets (y_1, x_{1a}, x_{1b}), $(y_2, x_{2a}, x_{2b}) \ldots (y_n, x_{na}, x_{nb})$. Each triplet signifies that points in the set lie between (y_k, x_{ka}) and (y_k, x_{kb}). Note that it is possible for a certain area to have a given value of y in more than one triplet.

One may immediately find the area of the object from this list, and may directly refer from it to points lying within the object.

Connected regions are usually defined by specifying a nearest-neighbor set for each point in the array, and to say that two black (white) points lie in a connected region if there is a chain of points from one to the other such that all points in the chain are black (white), and each succeeding pair lies in the same nearest-neighbor set. If we use the same nearest-neighbor definition for both the black and the white points we may run into an anomaly. Under the "tight" nearest-neighbor definition of Fig. XVI-5b the set of

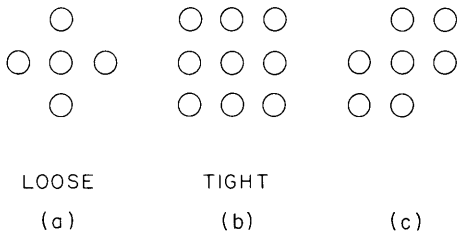


Fig. XVI-5. Nearest-neighbor definitions.

black points (1,1), (2,2), (3,3), (4,4) are connected, yet if we use the same neighborhood relation for white points, the points (2,3) and (3,2) are also connected, so that the line of black points fails to divide the white region into two pieces. One may avoid this by using the "loose" nearest-neighbor definition of Fig. XVI-5a for the white points, or by using the relation of Fig. XVI-5c for both black and white areas. There seems to be no clear a priori reason for preferring any of these nearest-neighbor criteria, but one should choose the correct pair. Thus the relations for the black and white points may have to be different.

3. Algorithms

We shall present two algorithms for connected area resolution. The first of these, the interval algorithm, operates on one line at a time. We suppose that the previous line has been processed and the intervals of black points, as well as the names of the areas to which these intervals belong, have been found and are available to us. We now locate the connected intervals of black points in the current line. These are matched against those from the preceding line.

Whenever an interval in the current line is connected to at least one interval from the preceding line, it is assigned to that name. If there is no connecting interval in the preceding line, a new name is assigned to this interval. The only difficulty that is encountered is shown in Fig. XVI-6. If an interval is connected to regions with two different numbers, it is given a name equal to that of the smaller of the two numbers, and an entry in a separate list is made signifying that the two names

(numbers) correspond to the same region. When this process is completed, the data from the previous line and the name equivalence data may be written on some bulk device.

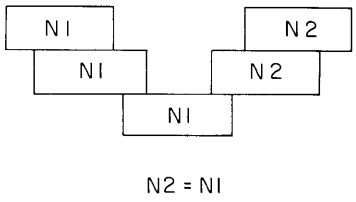


Fig. XVI-6. Interval algorithm.

After all of the lines have been processed as described, two lists will have been generated: one containing the intervals and their tentative names, and one with the name equivalence data. The latter contains a string of name-pairs. The pairs are first sorted so that the smaller name of each pair appears first in each doublet, and the set of pairs is sorted by first and second names. We now form a dictionary that specifies the actual name to be assigned to each tentative name. A pass is now made through the list of intervals and the final names are substituted.

The second algorithm uses the contour trace[2] to delineate a connected set. At any stage of this procedure we maintain an interval list of all points not yet examined and not known to contain black points. We examine these points. Whenever an interval has been shown not to have any black points, we eliminate it from this list. If a black point is found, we perform a contour trace. This gives us a set of (y, x) values, the successive points on the contour.

The interval list is constructed from these points in the following way. The array of contour points is examined to determine whether a white point is immediately adjacent and at the same y value as the contour point. This can be found from the coordinates of the following and preceding contour points, since the only time this does not happen is when the three points have the same y value, and if the x coordinates on the three points are in order. The points that fail this test are eliminated. Points at which the contour performs an up-down pivot are entered twice. The remaining points are sorted on y and on x. The sorted list is processed, attention being given to the fact that the contour can pass more than once through a given point. The output of this process is the interval list for the area. The points corresponding to this area are eliminated from the interval list used for searching, and the procedure is continued.

It is possible that black "islands" may contain white "lakes," with further "black" islands, and so on. If these are desired, one may use the interval lists of the areas as search lists, and to explore the interior. Care must be taken in outlining a white "lake," since the contouring must be done with the correct nearest-neighbor relation.

4. Conclusions and Discussion

It is difficult to make a general comparison between these algorithms. The contour algorithm requires that the whole picture be in storage, while the interval algorithm can operate on one y line of picture at a time. On the other hand, the contour algorithm automatically provides the outline, and this seems to be a valuable shape descriptor.

The interval algorithm is very similar to those described by Rosenfield[1] and by Bostrom,[3] although the form presented seems to be faster than the first and more general than the second.

Both the interval and the contour algorithms seem to be inefficient for images that contain very many very small black areas.

The procedures were coded in MAD for a 7094 computer. When applied to a micrograph of a Papanicolaou smear sampled over a 108 × 172 grid the first algorithm took 3.6 sec, and the second 7.2 sec. We can make no further claims about this, except that the object code produced by this compiler is inefficient. It is hoped that if the procedures are found to be useful they will be implemented with a special-purpose processor or with a computer that has some specially adapted instructions.

O. J. Tretiak, E. G. Guttmann

References

1. A. Rosenfield and J. L. Pfaltz, "Sequential Operations in Digital Picture Processing," J. A. C. M., Vol. 13, pp. 471-494, 1966.

2. R. S. Ledley, "High Speed Automatic Analysis of Biomedical Pictures," Science, Vol. 146, No. 3640, pp. 216-223, October 9, 1964.

3. R. C. Bostrom, H. S. Sawyer, and W. E. Tolles, "Instrumentation for Automatically Pre-screening Cytological Smears," Proc. IRE 47, 1895-1900, 1959.