

XVII. COGNITIVE INFORMATION PROCESSING*

Academic and Research Staff

Prof. M. Eden	Prof. K. Ramalingasarma	Dr. O. J. Tretiak
Prof. T. S. Huang	Prof. W. F. Schreiber	C. L. Fontaine
Prof. F. F. Lee	Prof. D. E. Troxel	P. H. Grassmann†
Prof. C. Levinthal	Dr. W. L. Black	E. R. Jensen
Prof. S. J. Mason	Dr. K. R. Ingham	J. E. Green
	Dr. P. A. Kolers	

Graduate Students

J. Allen	R. W. Kinsley, Jr.	R. A. Schaffzin
G. B. Anderson	A. N. Kramer	K. B. Seifert
T. P. Barnwell III	W-H. Lee	C. L. Seitz
B. A. Blesser	C-N. Lu	D. Sheena
A. J. Citron	J. I. Makhoul	D. R. Spencer
P. C. Economopoulos	G. P. Marston III	W. W. Stallings
A. Gabrielian	O. R. Mitchell, Jr.	P. L. Stamm
C. Grauling	S. L. Pendergast	R. M. Strong
R. E. Greenwood	D. R. Pepperberg	C. R. Tegnelia
E. G. Guttman	G. F. Pfister	Alice A. Walenda
R. V. Harris III	R. S. Pindyck	G. A. Walpert
D. W. Hartman	W. R. Pope, Jr.	J. W. Wilcox
H. P. Hartmann	D. S. Prerau	J. J. Wolf
A. B. Hayes	G. M. Robbins	J. W. Woods
G. R. Kalan		I. T. Young

A. CUBIC CORNERS‡

One structure that comes frequently before our eyes in this urban world is the cubic corner. This is simply any corner of a cube, of a desk or box, or in general 3 planes meeting at right angles. The simplest line drawing of a cubic corner consists of a point with 3 radiating straight lines of indefinite length. For easy reference, such a drawing will be called a 3-star.

The fact is that 3-stars tend to look like cubic corners. That is, the visual system tends to interpret them as the projections of cubic corners. This is not so under all conditions. Not only the 3-star itself, but any larger scene in which a 3-star is imbedded may influence its interpretation. Moreover, it is likely that our urban experience contributes to this propensity; the Watusi might not join the trend at all.

* This work was supported principally by the National Institutes of Health (Grants 1 PO1 GM-14940-01 and 1 PO1 GM-15006-01), 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).

† On leave from Siemens A.G., Erlangen, Germany.

‡ This work was supported in part by Project MAC, an M. I. T. research program sponsored by the Advanced Research Projects Agency, Department of Defense, under Office of Naval Research Contract Nonr-4102-(01).

(XVII. COGNITIVE INFORMATION PROCESSING)

In Fig. XVII-1 there are some 3-stars that look like cubic corners. In Fig. XVII-2 there are some that do not.

Fig. XVII-1.

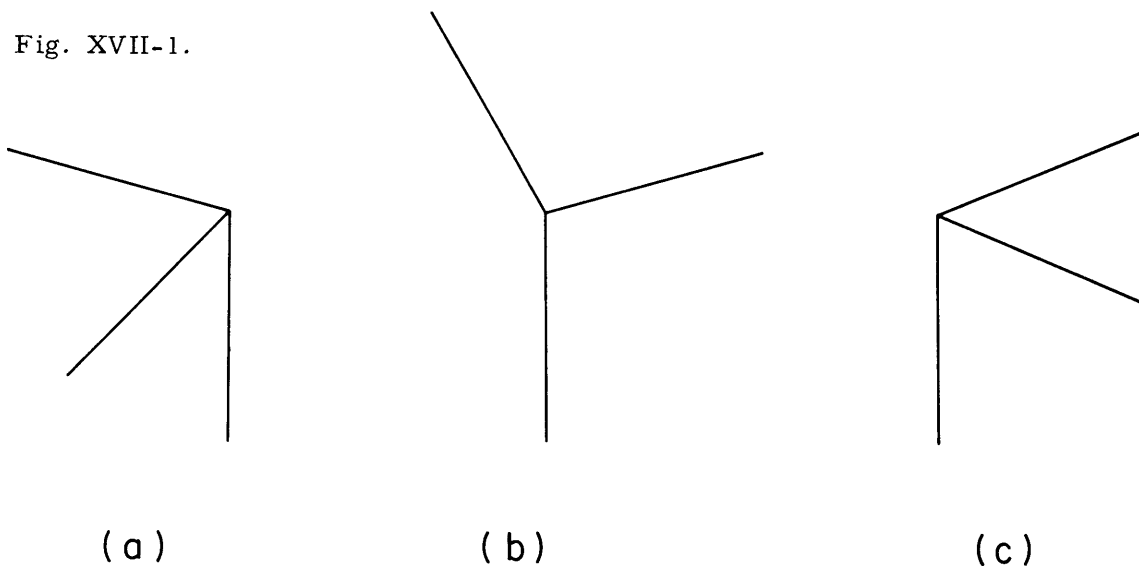


Fig. XVII-2.

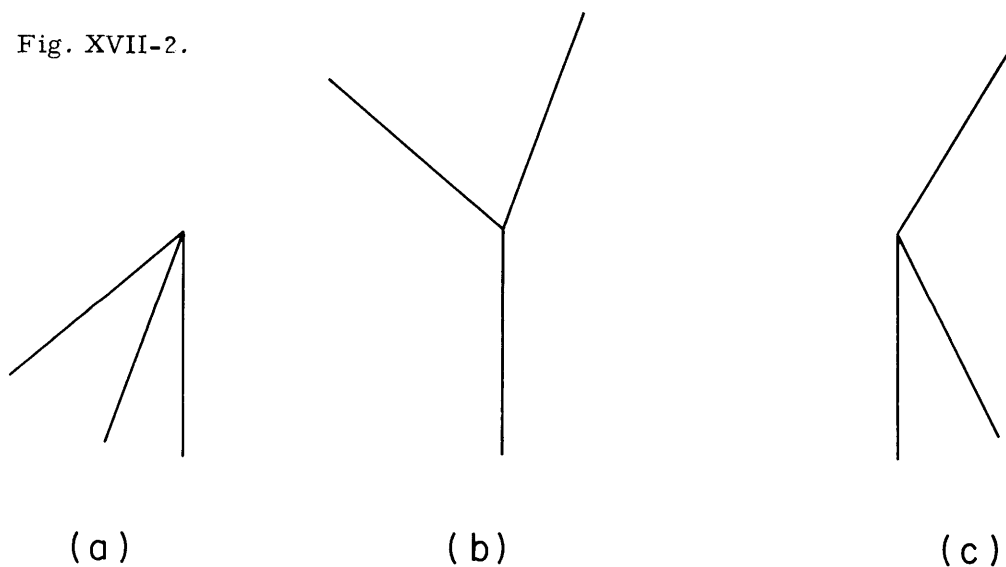
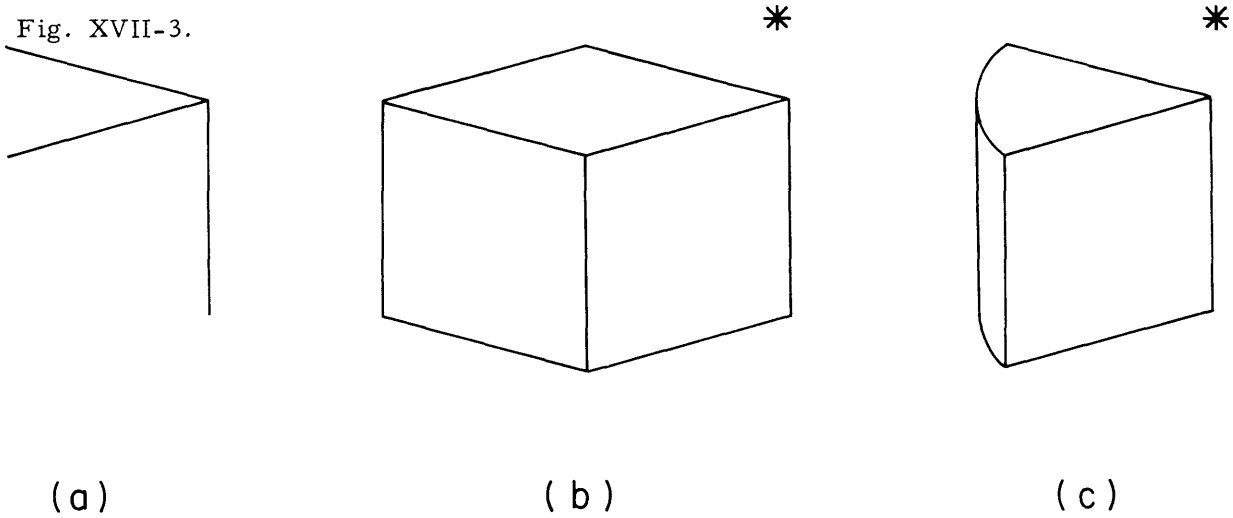


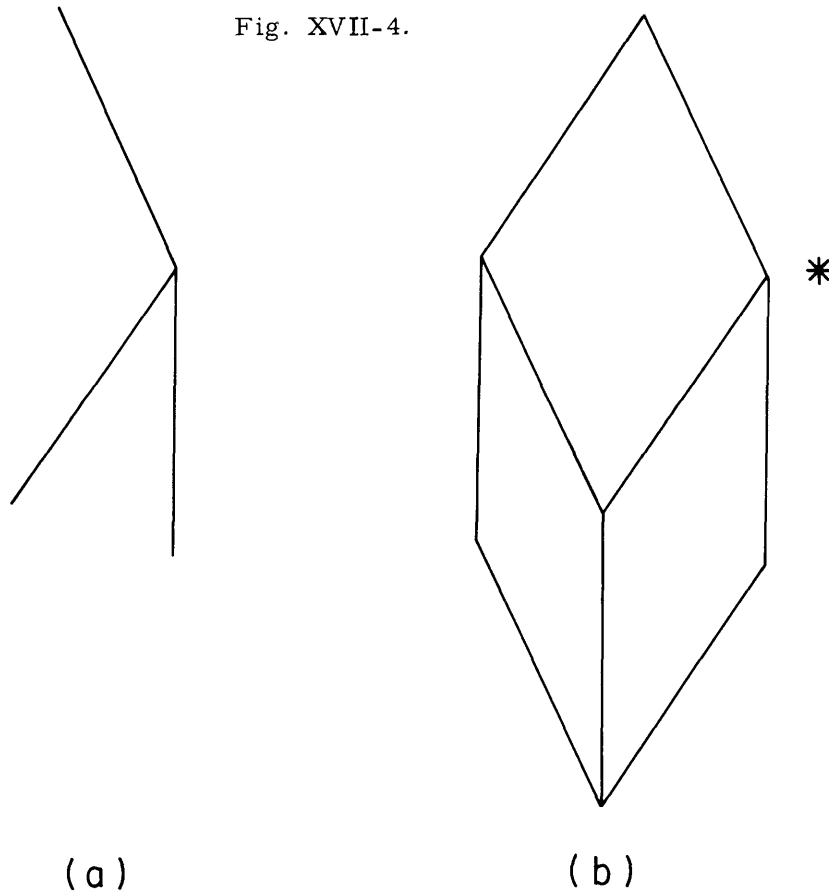
Figure XVII-3 exemplifies how context can influence interpretation. The 3-star of Fig. XVII-3a appears in Fig. XVII-3b and XVII-3c in places indicated by stars (*). Figure XVII-3a may look like a cubic corner to the reader, as it does to the author. At any rate, Fig. XVII-3b contains the 3-star in a context that certainly makes it appear cubic. Figure XVII-3c, on the other hand, contains the 3-star in such a way that it looks like a wedge of not too wide an angle. The circular arc on the left of Fig. XVII-3c suggests a wedge of cheese.

Fig. XVII-3.



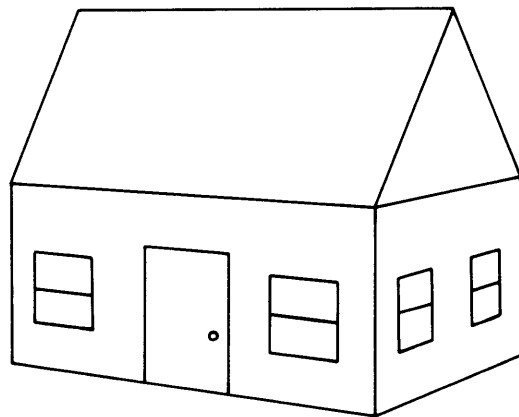
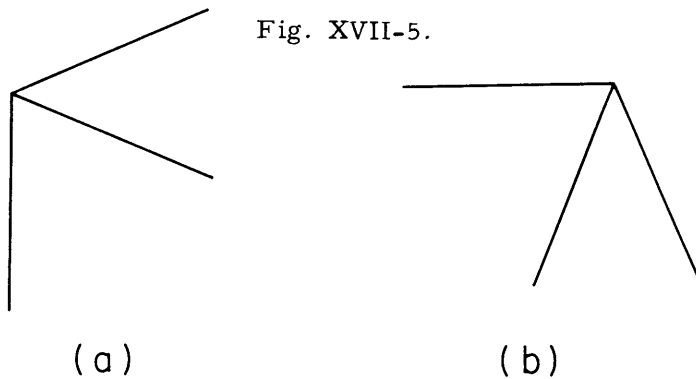
One might suspect that the interpretation of a 3-star is just a matter of context and has nothing to do with the 3-star itself. In Fig. XVII-3b, a drawing looking like a rectangular solid was constructed about the 3-star by putting the sides of the 3-star into parallelograms and completing the figure. Try this procedure with the 3-star in Fig. XVII-4a; the results are hardly analogous. Figure XVII-4b does not look rectangular at all.

Fig. XVII-4.



(XVII. COGNITIVE INFORMATION PROCESSING)

The orientation as well as the context of a 3-star can influence its interpretation. The 3-star in Fig. XVII-5a can readily be seen as a cubic corner. Yet in Fig. XVII-5b where the figure is rotated 90°, the smallest angle no longer appears to be right, but rather acute. Figure XVII-5c speculates on one of the common scenes in our experience which might lie behind such an orientation dependency.



(c)

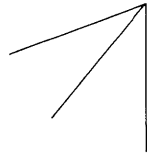
Three-stars are most likely to look like cubic corners when one of the rays points directly downward. Most cubic corners that one sees on buildings, boxes, and such things are oriented with their planes perpendicular or parallel to the ground, and consequently have one vertical edge. Probably a down-pointing ray in a 3-star invokes these experiences.

Still there is the question, How does the structure of the 3-star itself influence its interpretation? One can determine mathematically which 3-stars can be perspective projections of cubic corners. We shall illustrate this a bit later. The diagrams already presented show that the visual system will not accept that certain 3-stars look like cubic corners. The questions then are, How critical is our visual system?

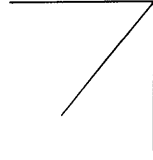
Does one reject 3-stars that could be cubic corners? Does one accept 3-stars as cubic corners when they could not possibly represent them geometrically? How well does the visual system align itself with the requirements of geometry and perspective?

Figures XVII-6 and XVII-7 contain two series of 3-stars for the judgment of the eye. All have rays pointing straight down to enhance the seeing of cubic corners. In the series of Fig. XVII-6 the 3-stars b through d and j through l are acceptable to the eye as cubic corners (if in doubt, construct parallelograms around them), and these cubic corners have two faces in view. In Fig. XVII-7 only d through g are seen as cubic corners, and these cubic corners have three faces in view.

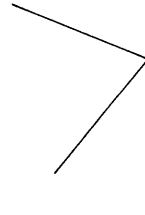
Fig. XVII-6.



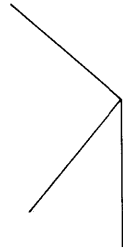
(a)



(b)



(c)



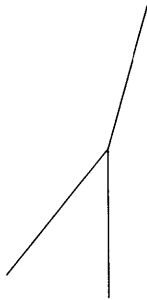
(d)



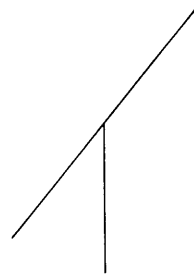
(e)



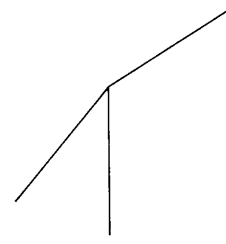
(f)



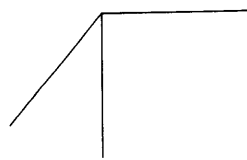
(g)



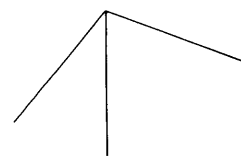
(h)



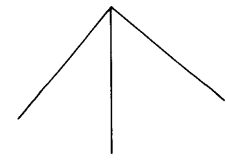
(i)



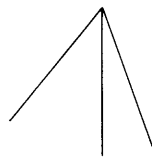
(j)



(k)

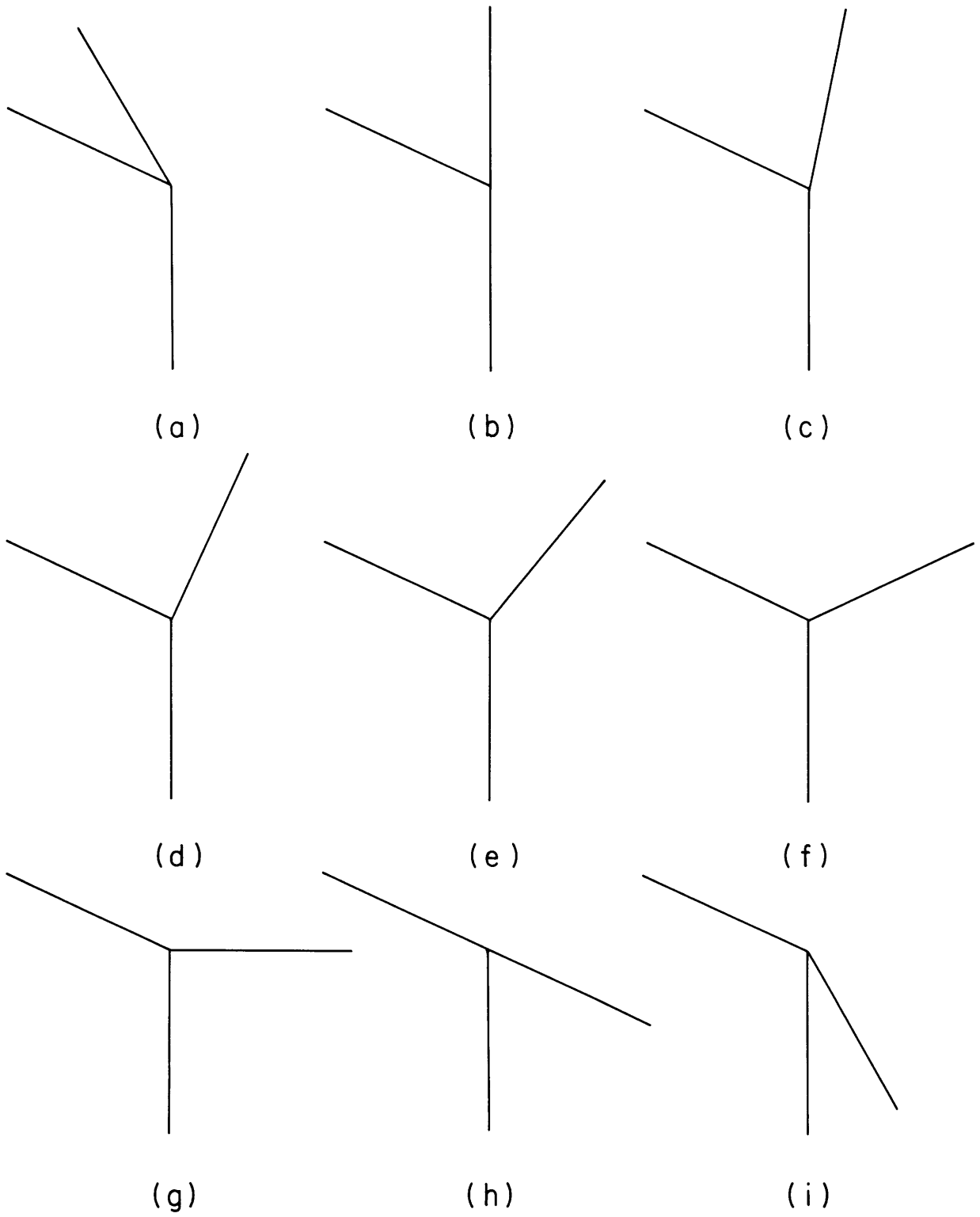


(l)



(m)

Fig. XVII-7.



The empirical generalization suggested by Fig. XVII-6 and other sketches is:

A 3-star is acceptable to the visual system as a 2-faced cubic corner if and only if it contains 2 angles less than or equal to 90° , whose sum is greater than or equal to 90° .

The generalization suggested by Fig. XVII-7 and other sketches is:

A 3-star can appear to be a 3-faced cubic corner only if all 3 angles are greater than or equal to 90° .

Now that there is some empirically derived notion of what the visual system interprets as a cubic corner, the problem is to determine which 3-stars are perspective views of cubic corners.

At first thought, it would seem necessary to calculate an arbitrary perspective projection of an arbitrarily oriented cubic corner. This is not the case for two reasons. In the first place, the eye is assumed to be focused on the vertex of the cubic corner; in the second place, only the angles between the projected rays, and not their length in projection, are of concern. Under these conditions, perspective can be ignored and the projection of a cubic corner in $X Y Z$ space will simply be its orthogonal projection on the $X Y$ plane.

Suppose that some 3-star has angles between its rays a, b, c , and also that the rays are represented by unit vectors P, Q, R in the $X Y$ plane. The question is, Are there 3 vectors in $X Y Z$ space which are mutually orthogonal and project, respectively, to P, Q, R ?

Since projection is accomplished merely by dropping the Z component, any 3 such vectors must have the form:

$$P + pZ \quad Q + qZ \quad R + rZ, \quad (1)$$

where Z is the unit vector in the Z direction, p, q, r scalars. Requiring mutual orthogonality is requiring that the dot products of these vectors in pairs be zero. Using distributivity of the dot product and remembering that P, Q, R are unit vectors yields (2), and similarly for the pairs Q, R and R, P .

$$(P+pZ) \cdot (Q+qZ) = 0 = P \cdot Q + pq = \cos a + pq = 0. \quad (2)$$

This produces three equations:

$$pq = -\cos a \quad qr = -\cos b \quad rp = -\cos c \quad (3)$$

Solving for p, q, r , we obtain

$$p = \pm \sqrt{-\frac{(\cos a)(\cos c)}{(\cos b)}} \quad q = \pm \sqrt{-\frac{(\cos a)(\cos b)}{(\cos c)}} \quad r = \pm \sqrt{-\frac{(\cos c)(\cos b)}{(\cos a)}} \quad (4)$$

(XVII. COGNITIVE INFORMATION PROCESSING)

Hence solutions exist and are given by these formulas, provided that $(\cos a)$, $(\cos b)$, $(\cos c) \neq 0$, and (so that the numbers under the radicals will be positive) either one or three of $\cos a$, $\cos b$, $\cos c$ are negative.

The solutions in (4) have singularities where $\cos a$, $\cos b$, or $\cos c = 0$; that is, where a , b , or $c = 90^\circ$ or 270° . To what do these singularities correspond in the situation under study?

If there is only one 90° or 270° angle, this represents the situation in which one ray of a cubic corner is pointed directly away from or toward the observer, and hence only 2 edges (not a 3-star) are visible.

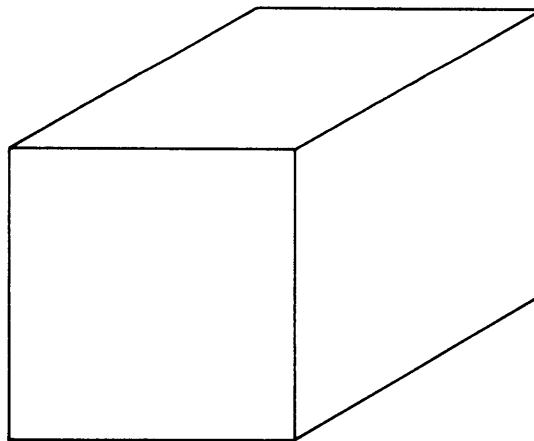
If there are two 90° angles, the solutions to (4) contain the indeterminate form $0/0$. This situation occurs when one plane of the cubic corner is oriented edgewise to the viewer. This is the only case when a 3-star that is a projection of a cubic corner can contain a right angle. Any other combination is impossible, since $a + b + c = 180^\circ$.

Aside from the singularities, What does the requirement that either one or three of $\cos a$, $\cos b$, $\cos c$ be negative represent? If, say, only $\cos c$ is negative, this implies that a is less than 90° , b is less than 90° , and $a + b$ is greater than 90° . If all three, $\cos a$, $\cos b$, and $\cos c$, are negative, this implies that a , b , c , are all greater than 90° .

The rules derived mathematically agree with the rules derived empirically, except for the 3-stars containing one right angle! To rephrase the conclusion:

Given a 3-star in favorable conditions of orientation and context and not containing one right angle, the visual system will accept that 3-star as a cubic corner if and only if it could in geometric fact be a projection of a cubic corner.

Fig. XVII-8.



A 3-star that does have one right angle cannot be the geometric projection of a cubic corner. Yet the visual system will often accept such a 3-star as a cubic corner (Fig. XVII-8).

D. N. Perkins

B. SURFACE RECONSTRUCTION FROM CONTOUR SAMPLES: SOME COMPUTER SIMULATION RESULTS

Briefly, the problem is the following: Given the constant elevation contours of some surface, how should one reconstruct the surface? This problem has been viewed as a statistical estimation problem, and the equation governing the estimate has been derived previously.¹

For simulation purposes, the two-dimensional random field was scanned, thereby producing a random process. Level-crossing times, together with the corresponding levels, were used as inputs to a reconstruction algorithm.

The computer was used to study how the reconstruction scheme would behave with such information. Only the most elementary reconstruction formula for the scanned field was implemented. The source process that was used was a black and white snapshot. The statistical processing of the snapshot was compared with one that had been quantized to the same levels and with one in which straight lines joined crossings.

1. Basic Reconstruction Formula

By assuming that the scanned field is Gaussian and by ignoring all of the inequality information except at the estimation point, the equation²

$$\hat{x}(t) = m(t) + \sigma(t) \frac{\int_A^B z \exp\left(-\frac{1}{2} z^2\right) dz}{\int_A^B \exp\left(-\frac{1}{2} z^2\right) dz}$$

may be applied. The integral in the denominator is computationally difficult to evaluate. Also, the role played by that quotient of integrals is to supply a number between A and B, that is, the center of gravity of the Gaussian density between A and B. In this case, it is justifiable to use the Cauchy density instead of the Gaussian density. If A and B are both much greater or less than zero, however, the Cauchy approximation to the center of gravity is very bad, but the estimation formula still does not suffer. The actual estimation function that was used was

$$\hat{x}(t) = \begin{cases} m(t) + \frac{1}{2} \frac{a(t) \ln [(1+B^2)/(1+A^2)]}{\tan^{-1} [(B-A)/(1+AB)]}, & AB \geq -1 \\ m(t) + \frac{1}{2} \frac{a(t) \ln [(1+B^2)/(1+A^2)]}{-\pi + \tan^{-1} [(B-A)/(1+AB)]}, & AB \leq -1, \end{cases} \quad (1)$$

where $A = (\alpha(t)-m(t))/a(t)$, $B = (\beta(t)-m(t))/a(t)$, and $a(t) = K\sigma(t)$. The constant K was

determined so that

$$\frac{1}{\sqrt{2\pi}} \int_{-2}^2 \exp\left(-\frac{1}{2}z^2\right) dz = \frac{K}{\pi} \int_{-2}^2 \frac{1}{x^2 + K^2} dx.$$

The symbols are defined as follows: \hat{x} is the estimate at point t ; the conditional mean is $m(t)$; and the standard deviation is $\sigma(t)$. Both of these are conditioned upon the waveform passing through the levels at the corresponding crossing times. The two functions $\alpha(t)$ and $\beta(t)$ bound the original waveform: $\alpha(t) \leq x(t) \leq \beta(t)$. These bounding functions arise because level crossings have been detected.

For each particular estimation point t , only the sample values at the two neighboring crossing times are used. Denote crossing times t_j , and the corresponding levels x_j . Let $R(s, t)$ be the autocorrelation function for $x(t)$. The mean of $x(t)$ restricted to pass through x_j at t_j and x_{j+1} at t_{j+1} is

$$m(t) = a_j x_j + a_{j+1} x_{j+1},$$

where the a_j , a_{j+1} satisfy

$$R(t, t_j) = a_j R(t_j, t_j) + a_{j+1} R(t_{j+1}, t_j)$$

$$R(t, t_{j+1}) = a_j R(t_{j+1}, t_j) + a_{j+1} R(t_{j+1}, t_{j+1}).$$

Also, the process is assumed to be stationary so that a normalized autocorrelation function may be usefully defined: $\rho(\tau) = R(\tau)/\sigma^2$. When σ^2 is written without an argument " t ", it will denote the variance of the stationary, unrestricted process $x(t)$. Also, for notational convenience, set $\lambda = t - t_j$, $\tau = t_{j+1} - t$, $T = t_{j+1} - t_j$. The solution for a_j and a_{j+1} is

$$a_j = [\rho(\lambda) - \rho(\tau)\rho(T)] / [1 - \rho^2(T)]$$

$$a_{j+1} = [\rho(\lambda)\rho(T) - \rho(\tau)] / [1 - \rho^2(T)].$$

The variance of $x(t)$, restricted just as the mean was, is

$$\sigma^2(t) = \sigma^2 [1 - a_j \rho(\lambda) - a_{j+1} \rho(\tau)].$$

2. Level-Crossing Processing

It is assumed that the input waveform has been sampled periodically in time at a rate fast enough so that the original may be reconstructed adequately by straight lines through the samples. The level-crossing processing has two steps: the crossing times and

corresponding levels are extracted from the input waveform by assuming that straight lines join the samples, and the input waveform is reconstructed by applying the interpolation equation.² The reconstructed waveform is represented at the original periodic sample times.

Some of the crossing times were eliminated because they included no time at which the waveform was to be reconstructed. Perhaps this would be best clarified by an

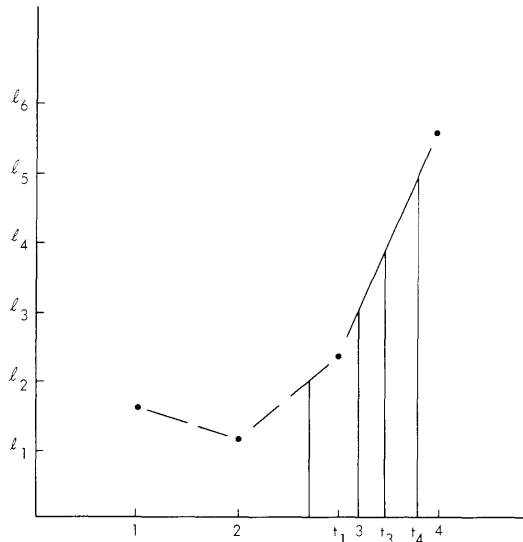


Fig. XVII-9. Detection of the level crossings of the sampled waveform.

example. In Fig. XVII-9 the input waveform is shown in its sampled form. To extract the crossing times, one may picture straight lines joining the samples. Between $t = 3$ and $t = 4$ there are 3 level crossings. In this example, the crossing times t_2 and t_3 are ignored. In general, between any two samples of the input waveform, if there are any crossings, only the rightmost crossing is retained.

To reconstruct the waveform from these crossing times, each adjacent pair is taken successively. Then, for each such pair t_j, t_{j+1} , the interpolation formula is evaluated at all points in the half-open interval $[t_j, t_{j+1})$ at which the input was sampled.

3. Processing of the Snapshot

The snapshot was quantized into 1024 brightness levels, and was sampled on a 256×256 grid. It was also necessary to know the autocorrelation function of the process. For the particular photograph used here, autocorrelation function has been estimated.³ The values for the mean, μ , and the standard deviation, σ , where $\mu = 403$ and $\sigma = 234$. The normalized autocorrelation $\rho(\tau)$ is shown in Fig. XVII-10. Also shown is a polynomial approximation to $\rho(\tau)$ that was used in the reconstruction. The approximation of $\rho(\tau)$ must be done with some care. For example, the formula for $\sigma^2(t)$ gives a negative result if

(XVII. COGNITIVE INFORMATION PROCESSING)

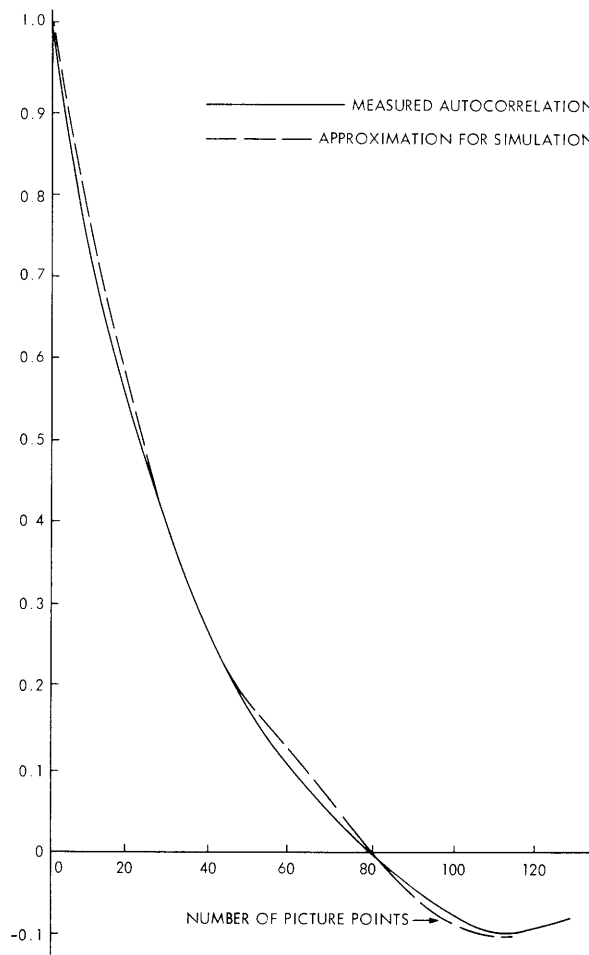


Fig. XVII-10. Normalized autocorrelation functions.

$\rho(\tau) = 1 - s\tau^2$, no matter what value of s is used and no matter how small the crossing interval, T , may be.²

The results of reconstructing the photograph from crossing-time information may be seen in Figs. XVII-11 and XVII-12. In either of these figures (a) is the original, (b) is a quantized version, (c) is a reconstruction using straight lines between crossings, and (d) is the statistical reconstruction. The difference in processing between Figs. XVII-11 and XVII-12 is that 8 equispaced levels are used in Fig. XVII-11, and 16 equispaced levels in Fig. XVII-12. The value of e shown in each case is the square root of the mean-square error per picture point:

$$e^2 = \sum_{j=1}^{256} \sum_{i=1}^{256} (x(i,j) - \hat{x}(i,j))^2 / (256)^2.$$



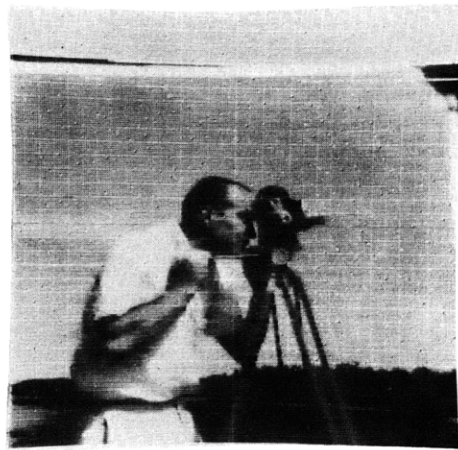
a)



b)
 $e=72.3$



c)
 $e=47.3$
 $N=11670$

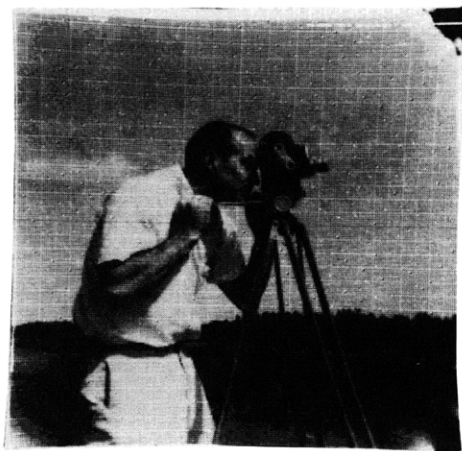


d)
 $e=34.1$
 $N=11670$

Fig. XVII-11. (a) Original.
(b) Quantized to 3 bits.
(c) Straight-line reconstruction.
(d) Statistical reconstruction.



a)



b)
 $e=37.7$

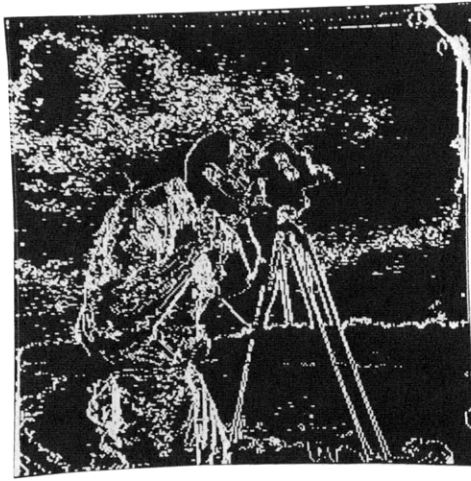


c)
 $e=22.9$
 $N=18679$

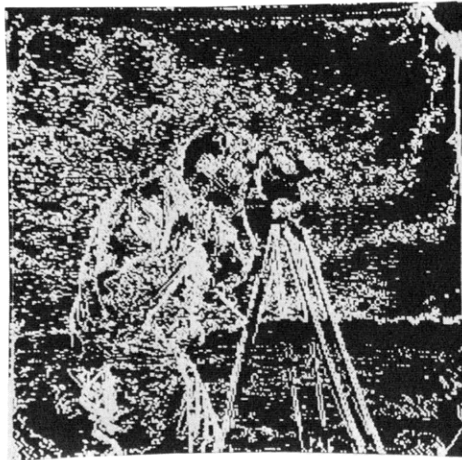


d)
 $e=16.0$
 $N=18679$

Fig. XVII-12. (a) Original.
(b) Quantized to 4 bits.
(c) Straight-line reconstruction.
(d) Statistical reconstruction.



a)



b)

Fig. XVII-13. Level-crossing points.
(a) 8 levels, equispaced in brightness.
(b) 16 levels, equispaced in brightness.

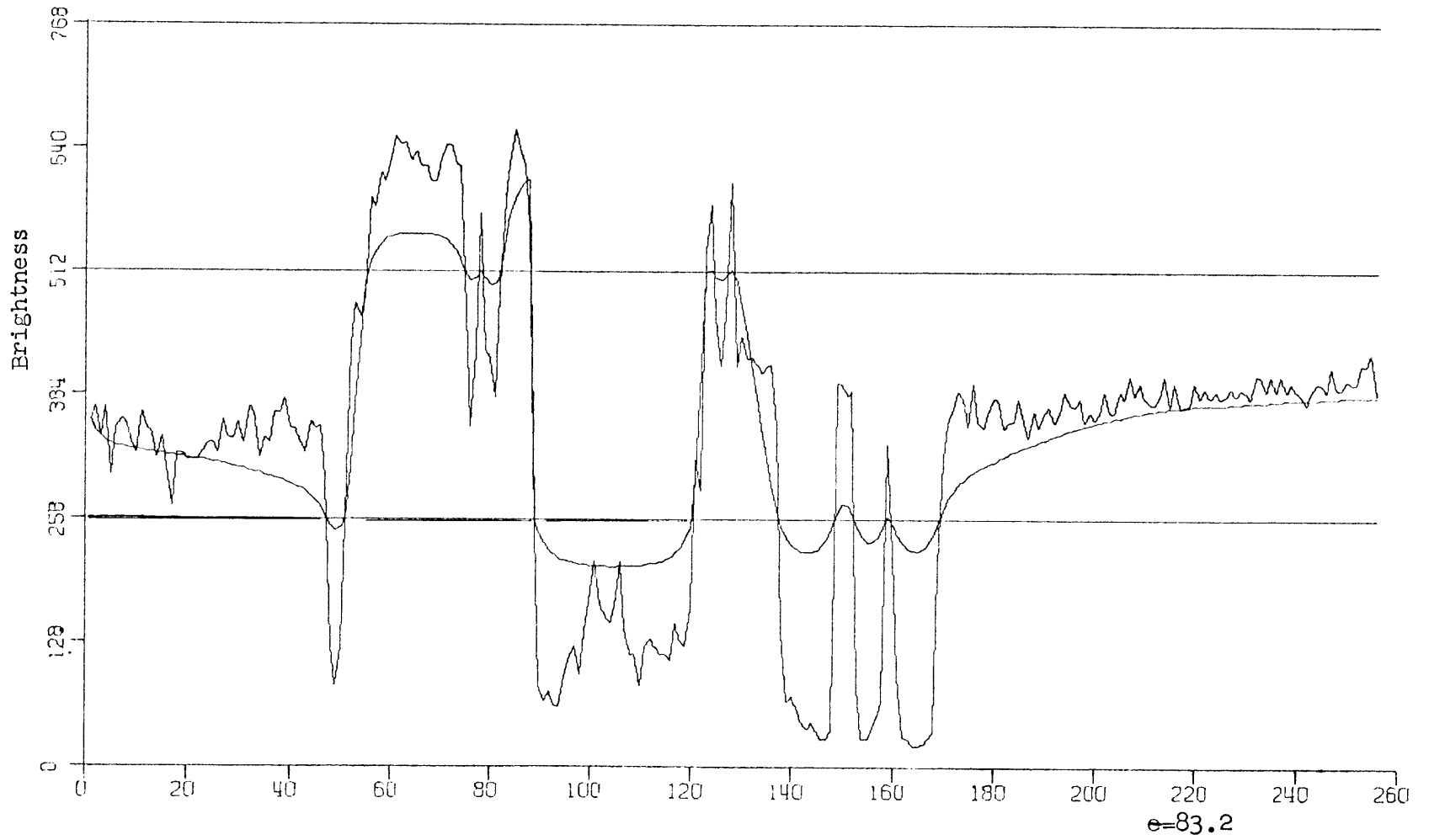


Fig. XVII-14. Reconstruction of the middle scan line - 4 levels.

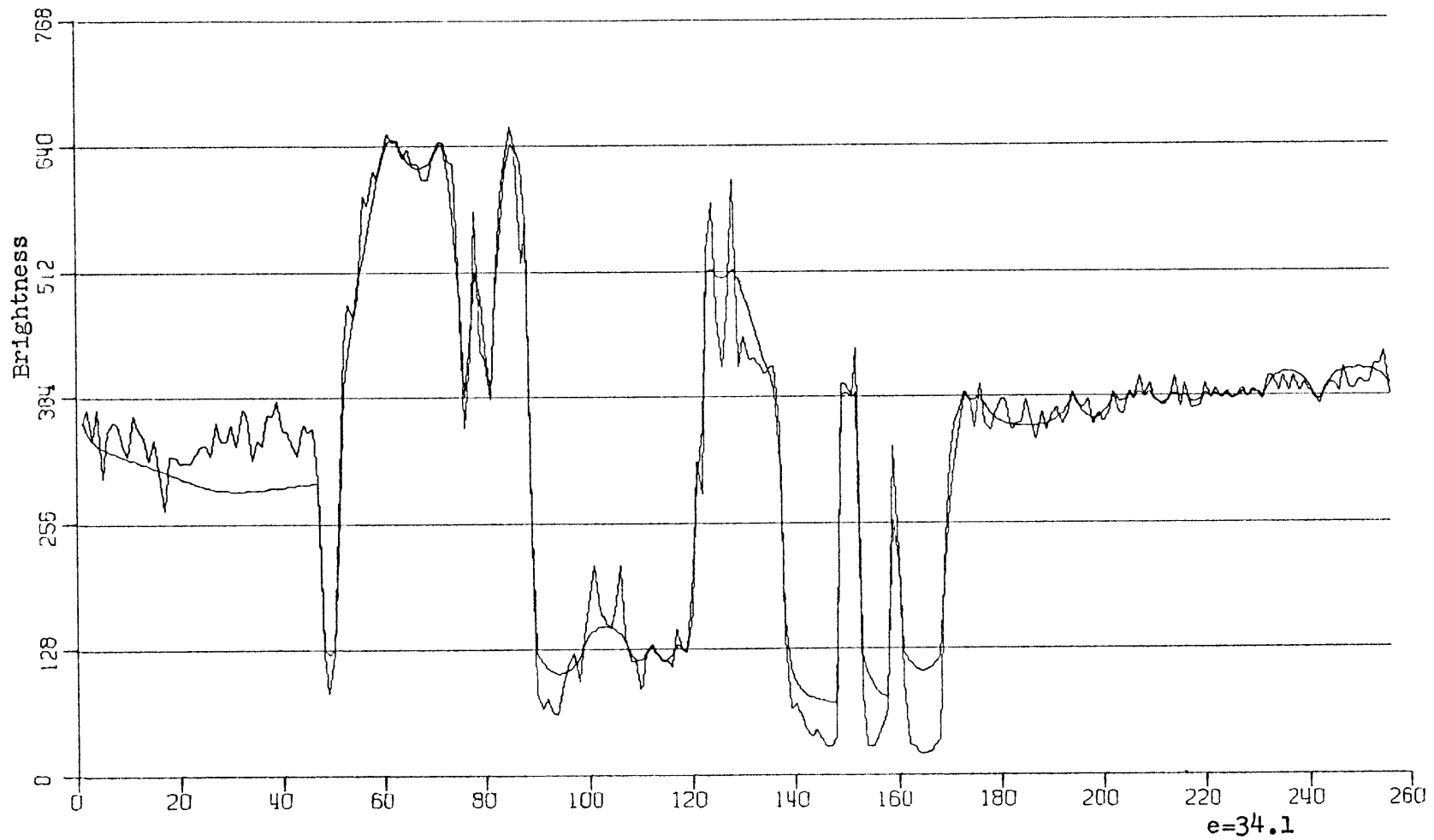


Fig. XVII-15. Reconstruction of the middle scan line - 8 levels.

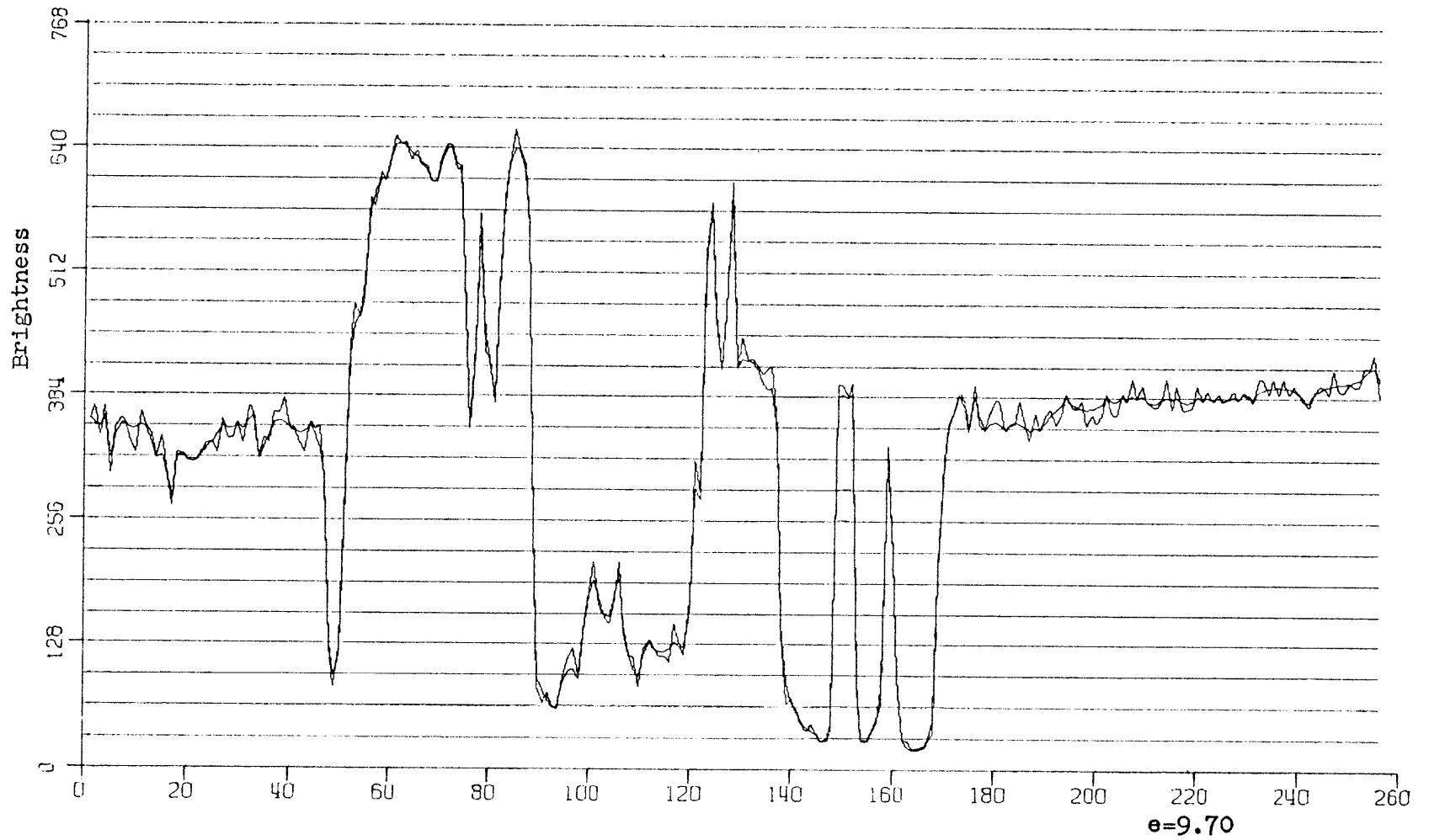


Fig. XVII-16. Reconstruction of the middle scan line - 32 levels.

Also, the value of N is given under (d) and (e). N is the total number of crossings detected. In Fig. XVII-11d the top eighth of the photograph is a constant. This was an input-output error with the magnetic tape, and is unrelated to any reconstruction procedure.

In Fig. XVII-13, the points at which the crossing times were detected are shown. Part (a) shows the crossings points for 8 levels, and (b) shows them for 16 levels. There was a total of $65,536 = (256)^2$ grid points. With 8 levels 11,670 crossings were detected, and with 16 levels 18,679 crossings were detected.

Besides having the smallest error, the statistically processed photograph has lessened the effect of the spurious contours and abrupt changes in brightness, which is so apparent in the quantized photograph. The fact that the photograph was processed in a horizontal direction is not apparent in Fig. XVII-12d, but is noticeable in Fig. XVII-11d. A disturbing feature of this statistical reconstruction is that the snapshots appear to have less contrast, but this is because the highest the reconstructed waveform may reach is approximately midway between the second highest and highest levels. Similarly, it never goes lower than the middle of the lowest and next to lowest levels.

In the straight-line interpolation scheme the same crossing times were used as those in the statistical scheme.

The straight-line scheme performs very poorly. It is apparent that the processing was done horizontally. Also, there are abrupt, artificial changes in brightness. Clearly, one could improve upon this simple interpolation scheme. The point in trying it was just to compare this simple interpolation scheme with the statistical one. Another point brought out by the straight-line interpolation is that the mean-square error is not a complete measure for goodness of fit. Even though the error was comparable to the error for the statistical interpolation, the appearance of the statistically reconstructed picture is of higher quality than the picture reconstructed by straight-line interpolation.

Shown in Figs. XVII-14, XVII-15, and XVII-16 are graphs of the statistical reconstruction of the middle scan line of the photograph. The number of levels are shown on each graph, and the original input is superimposed on the reconstructed version. They may be distinguished because the original has an apparent low-magnitude, high-frequency portion on the extreme right.

One aspect of this statistical reconstruction is demonstrated by these graphs. Very fast large-magnitude jumps in the process are always found, but the low-magnitude higher frequencies, in general, are missed.

G. M. Robbins

References

1. G. M. Robbins, "Surface Reconstruction from Contour Samples," Quarterly Progress Report No. 87, Research Laboratory of Electronics, M. I. T., October 15, 1967, pp. 160-163.

(XVII. COGNITIVE INFORMATION PROCESSING)

2. G. M. Robbins, "Surface Reconstruction from Contour Samples," S.M. Thesis, M.I.T., February 1968; see especially Equation (1), Sec. 4.1.
3. T. S. Huang, "Pictorial Noise," Sc.D. Thesis, M.I.T., September 1963.

C. FIELD SHADING AND NOISE IN THE SCAD SCANNER

1. Approximation Procedure

It is hoped that a scanner will give a constant signal when scanning a uniform field. This is not likely to happen; the signal will be biased by the deterministic nonuniformities in a scanner (shading), and perturbed by random noise.

In a flying-spot scanner such as SCAD¹ the deterministic distortion of the signal can be characterized fairly simply. The measured signal is a linear function of the transmittance, so that nonuniformities of the light transmission path can be compensated by measuring the signal produced by the scanner from a clear field, and by scaling the signal from a transparency by the reciprocal of the clear-field characteristic.

Two problems arise when one wishes to estimate the shading or the deterministic clear-field variation. The signal from the scanner is corrupted by noise, so that, to make such an estimate, one must perform the calculation in terms of a stochastic model. The analysis might be done by repeatedly scanning a clear field, and by finding histograms of values at each point in the field. Because of practical limitations, this was not feasible; hence, our analysis was done from single recordings of the signal. The second difficulty arises when one wishes to use the data for correcting the signal. Unless the data on the shading are summarized by a few parameters, it may not be feasible to perform the correction.

We decided to model the clear-field variation by a linear composition of several functions that are easy to compute and are likely to fit the shading pattern well. If $f(I, J)$ is the clear-field signal recorded over the raster $1 \leq I \leq IM$, $1 \leq J \leq JM$, and $f_A(I, J)$ is a function of the form, then

$$f_A(I, J, a_1 \dots a_n) = \sum_{i=1}^u a_i f_i(I, J).$$

The coefficients $a_1 \dots a_n$ are chosen to minimize the difference:

$$e(a_1 \dots a_n) = \sum_{I=1}^{IM} \sum_{J=1}^{JM} (f(I, J) - f_A(I, J, a_1 \dots a_n))^2.$$

We define the approximation error as

$$e_a = \left(\frac{\text{Min}_{a_1 \dots a_n} e(a_1 \dots a_n)}{IM * JM} \right)^{1/2}.$$

The difference between f and f_A may be thought of as being composed of two terms:

(XVII. COGNITIVE INFORMATION PROCESSING)

random noise, and shading that the f_A function could not fit. We take the expression

$$e_d = \left(\frac{1}{2 \cdot IM(JM-1)} \sum_{I=1}^{IM} \sum_{J=1}^{JM-1} (f(I, J) - f(I, J+1))^2 \right)^{1/2}$$

as a measure of the rms white noise in the signal. The quantity

$$e_u = (e_a^2 - e_d^2)^{1/2}$$

is an estimate of the variation in the signal that is neither white nor fitted by f_A .

2. Physical Sources of the Variation

The optical system used in SCAD is shown in Fig. XVII-17. The feedback signal modulates the grid drive to the cathode-ray tube in an attempt to keep the signal from the feedback PMT constant. The beam splitter that sends a signal to the feedback PMT is placed behind the imaging lens, so that the feedback system collects light over the

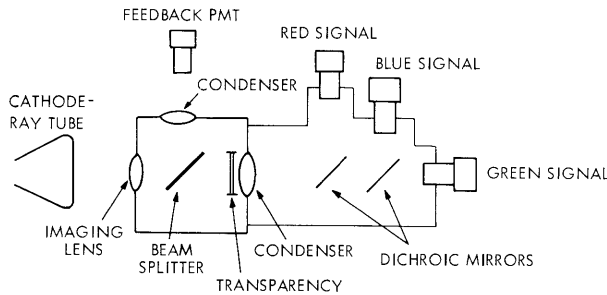


Fig. XVII-17. SCAD scanner optics (top view).

same solid angle as the transparency sensing path. One surface of this mirror is metallized to reflect approximately 30% of the light, and the other surface has an antireflection coating. The light transmitted by the transparency is split into color components by interference filters.

The principal sources of signal shading are the imperfections in the feedback system, and differences between the feedback and signal optical paths. The reflectance from a mirror depends on the incidence angle, and this angle varies when the scanner explores the transparency. The angle of incidence on the beam-splitting mirror varies over $45^\circ \pm 13^\circ$, and the angle of incidence on the dichroic mirrors is $45^\circ \pm 6^\circ$. The cone of rays from the lens, when the lens is at maximum aperture, is $\pm 6^\circ$. The reflectance-wavelength characteristics of the dichroic mirror depends on the angle of incidence and, since the light from the CRT is not white, this introduces both color and intensity variation. One may also expect variations because of artifacts such as dust or smudges

on the mirrors and condenser lenses.

One expects white noise from the amplifiers and the shot effect in the photomultiplier tubes, as well as 60 Hz pickup. With the light levels and integration times encountered in this system, these effects should be very small. Phosphor noise, or light variations because of the granularity of the phosphor, are expected to be white, and although the light feedback is supposed to cancel this, this cancellation is not perfect.

3. Results and Discussion

Clear-field signals from the three color channels were recorded on magnetic tape from a raster, 108 samples wide and 173 samples high. The placement of this raster relative to the field of view of the scanner is shown in Fig. XVII-18. The signal was quantized to 6 bits.

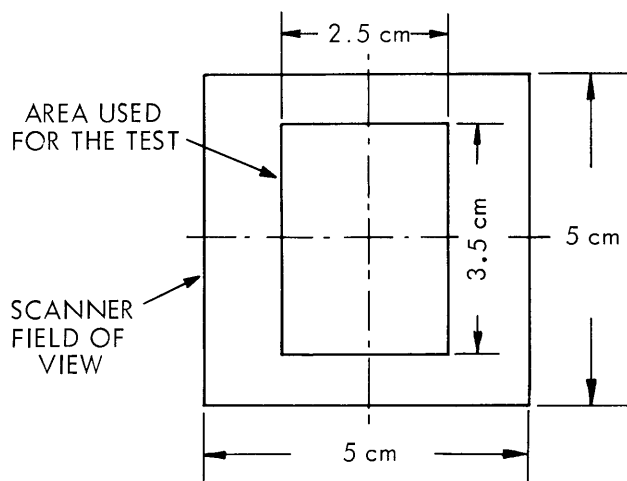


Fig. XVII-18. Area used in the field of view test.

The clear-field signals were processed according to the procedure described above. The functions used in the expansion are listed in Table XVII-1. The coefficients obtained from the analysis are given in Table XVII-2. These coefficients have been scaled so that the amplitude of the constant function, f_1 , is unity.

The variation in the clear-field signal seems to be described almost completely by the expansion in terms of the approximation functions and the white noise. The term e_u that measures the departure from this model is $\sim 1\%$, and this seems to be small enough to cause no concern under most conditions. We expect this term to be due to hum and errors in the approximation, and since there is a significant variation in e_u between the three fields and we expect the hum to be the constant, this variation is

(XVII. COGNITIVE INFORMATION PROCESSING)

Table XVII-1. Expansion functions.

$$f_1(I, J) = 1$$
$$f_2(I, J) = \cos (((I-1)/(IM-1) - .5) * \pi)$$
$$f_3(I, J) = \cos (((J-1)/(JM-1) - .5) * \pi)$$
$$f_4(I, J) = \sin (((I-1)/(IM-1) - .5) * \pi)$$
$$f_5(I, J) = \sin (((J-1)/(JM-1) - .5) * \pi)$$
$$f_6(I, J) = f_4 \cdot f_5$$

Table XVII-2. Field shading and noise coefficients.

Field	a ₁	a ₂	a ₃	a ₄	a ₅	a ₆	e _a	e _d	e _u
Red	1.00	-.0190	.0086	.1011	-.0071	.0009	.0271	.0263	.0075
Green	1.00	-.0107	.0529	.0089	.0076	.0072	.0230	.0213	.0086
Blue	1.00	-.0406	-.0230	-.0943	.0221	.0075	.0276	.0244	.0129

probably due to approximation errors.

The white noise, estimated by e_d, is much larger than is tolerable, and an effort will be made to reduce it. Its source has not been established, although it is definite that only a small part of it is caused by shot or quantization noise.

O. J. Tretiak

References

1. O. J. Tretiak, "Scanner Display (SCAD)," Quarterly Progress Report No. 83, Research Laboratory of Electronics, M. I. T., October 15, 1966, pp. 129-142.

D. BIOLOGICAL IMAGE PROCESSING – AUTOMATED
LEUKOCYTE RECOGNITION

Enormous amounts of time are expended in medical laboratories every day in the routine examination of smears prepared from human peripheral blood. The reason for making this examination is twofold: to possibly locate abnormal or pathological cells; and, more commonly, to perform a differential leukocyte (white cell) count. The very repetitive nature of this second test leads us to ask if it could be automated. That is, could we find an algorithm that could perform the differential count more efficiently and accurately than it is now being done. This is the task of automated leukocyte recognition.

1. Background

Suspended in the plasma (liquid) of mammals are 3 distinct types of cells: red corpuscles (erythrocytes), colorless corpuscles (white blood cells or leukocytes), and blood platelets. The leukocytes, which number between 5000 and 9000 cells per cubic millimeter of blood, are considered true cells, that is, they possess both a well-defined nucleus and a cytoplasm. The cells seen in dry smears are quite different from the same cells seen in living preparations. In the dry smears the normally spherical cells are flattened to large thin disks.

Leukocytes may be subdivided into two different categories: (a) nongranular leukocytes, and (b) granular leukocytes. Extensive examination of stained, dry smears has shown that these two categories may be further subdivided into the 5 types of leukocytes found in normal blood samples.

A. Nongranular Leukocytes

1. Lymphocytes
2. Monocytes

B. Granular Leukocytes

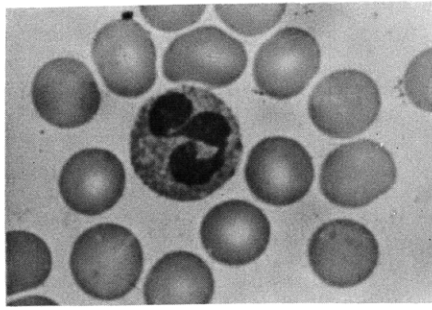
1. Eosinophils
2. Neutrophils
3. Basophils

When a blood smear, stained with Wright's stain, is examined under an oil-immersion microscope at 1000X magnification, the 5 different types of leukocytes may be differentiated according to the morphological (shape) and spectral (color) information summarized in Table XVII-3. Pictures of the different types of leukocytes are shown in Fig. XVII-19.

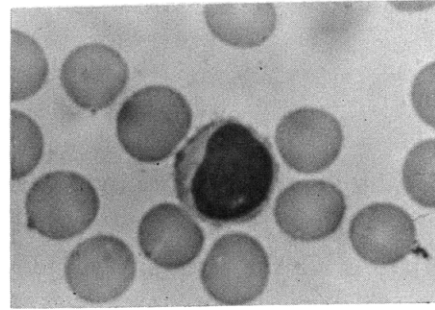
Table XVII- 3. Morphological and spectral characteristics of leukocytes.

Cell Type	Population (%)	Diameter (μ)	Cytoplasm		Nucleus			Granules
			Color	Amount	Color	Shape	Amount	Amount
Lymphocyte	25-33	7-12	Clear Light Blue	10-30%	Purple	Round	70-90%	Few
Monocyte	2-6	13-20	Grayish Blue	50%	Light Purple	Indented	50%	Few
Neutrophil	60-70	10-12	Pink to Lilac	50%	Reddish Purple	Segmented	50%	Many
Eosinophil	1-4	10-14	Red Stippling	60-70%	Dark Blue	Two-Lobed	30-40%	Many
Basophil	.25-.5	10-12	Blue	50%	Purple	Elongated	50%	Many
Red Cell	$4-6 \times 10^6$ per mm^3	4-6	Pale Orange Red	100%	None	—	0%	None

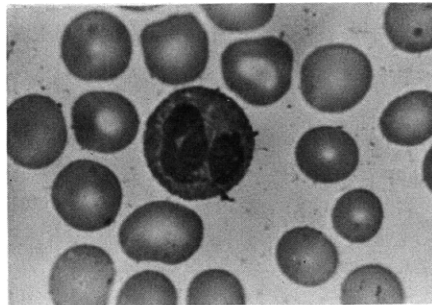
(XVII. COGNITIVE INFORMATION PROCESSING)



NEUTROPHIL

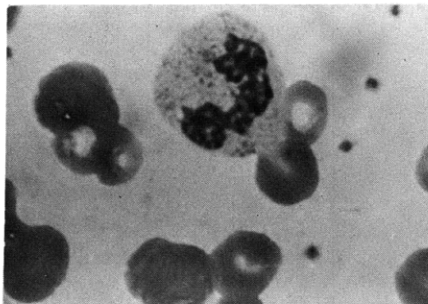


LYMPHOCYTE

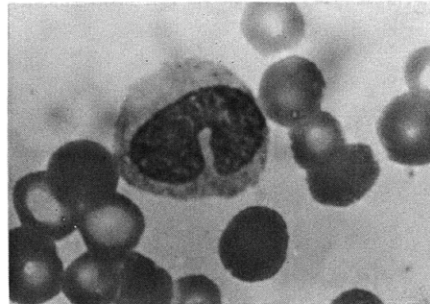


EOSINOPHIL

10 μ



BASOPHIL



MONOCYTE

Fig. XVII-19. Normal human leukocytes (1000X).

2. Optical Image Processing

Using Kodachrome II-A film, we photo-micrographed a standard blood smear stained with Wright's blood stain. The field of view was approximately $70 \mu \times 50 \mu$ and contained 50 red cells, 1 lymphocyte, and 1 neutrophil. The color transparency was scanned through SCAD, a flying-spot color scanner, and the resulting digital data were recorded on magnetic tape. The resolution of the entire system at this point was limited by the resolving power of the microscope:

(XVII. COGNITIVE INFORMATION PROCESSING)

$$R = \frac{0.61 \lambda}{n. a.}$$

For a center wavelength $\lambda_c = 0.55 \mu$, and the n. a. of the microscope objective (n. a. = 1.32) yields a resolution limit of approximately 0.25μ . For leukocytes this is roughly the size of the granules in the cytoplasm.

The three one-color pictures generated by SCAD were read into the Project Mac 7094 computer, and a program was developed to produce a multi-grey level picture on the off-line printer. By using an overprinting technique described by Mendelsohn and modified for the CTSS character set, a picture was produced that reduced the original 64 grey-level picture (corresponding to 6-bit quantization) to 17 distinct grey levels. A photograph of the "red picture" (the original color transparency viewed through a red

Table XVII-4. Grey-level overprint codes.

Output Level	Input Levels	Over-Print Symbol
1	0-2	MM
2	3-5	MM
3	6-8	MW
4	9-11	WW
5	12-15	WX
6	16-19	XX
7	20-23	XV
8	24-27	VV
9	28-31	V/
10	32-35	//
11	36-39	/,
12	40-43	W
13	44-47	X
14	48-51	V
15	52-55	/
16	56-59	,
17	60-63	

(XVII. COGNITIVE INFORMATION PROCESSING)

filter) is shown in Fig. XVII-20. A list of the grey-level overprint codes is given in Table XVII-4.

Examination of Fig. XVII-20 immediately reveals a major problem. Because of a nonuniformity in the scanner, the background illumination varies across the picture.

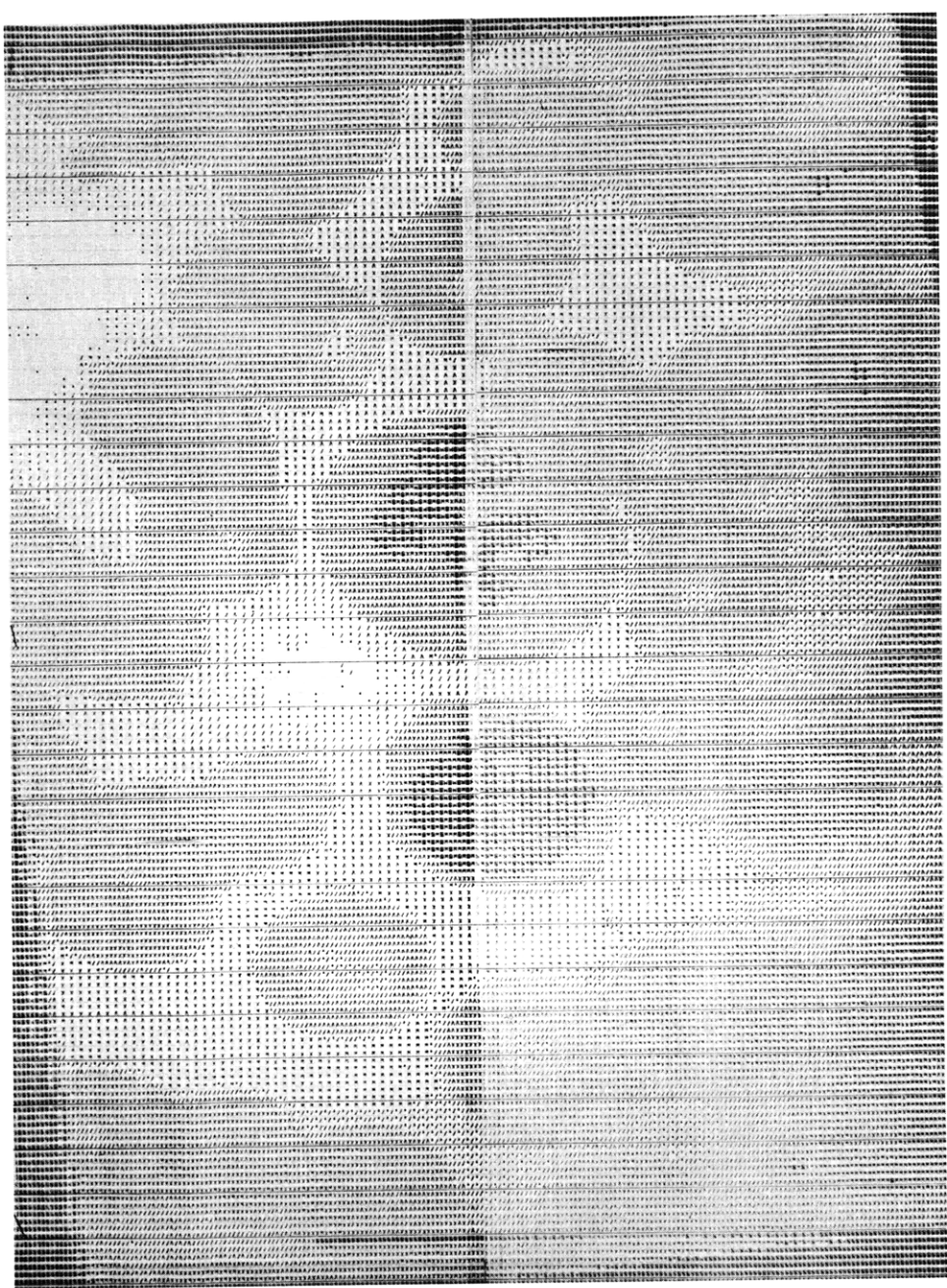


Fig. XVII-20. Original "red" picture with background.

(XVII. COGNITIVE INFORMATION PROCESSING)

The result of this is that contrast is reduced in many portions of the picture, and in some places it is almost impossible to separate the cells from background. This problem was solved by the following analysis. Let a pure black point be denoted by the number 0 and a pure white point by 63. Now, for the sake of simplicity, assume that the nonuniformity occurs in a purely horizontal direction, that is, along a scan line. If there were no distortion in the scanning illumination and we recorded a perfectly clear field, then we would get a brightness versus position curve as shown in Fig. XVII-21a. Because of the nonuniformity, however, plus our observation that in Fig. XVII-20 we are in general able to distinguish the cells from background, we may conclude that the background illumination is a slowly varying function of position. We shall call this phenomenon "tilt" after the simulated picture of the nonuniformity in Fig. XVII-21b.

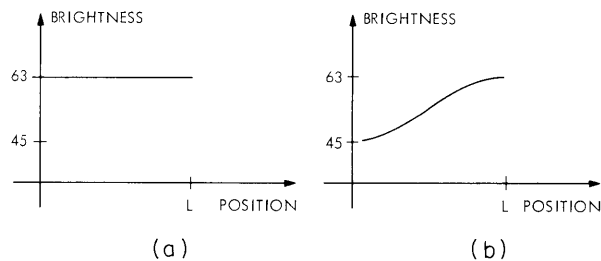


Fig. XVII-21. Effect of scanner nonuniformity.

Although the details of the derivation will not be given here, it is an easy matter to show that the effect of the two-dimensional tilt $B(x, y)$ on any picture brightness distribution $f(x, y)$ is to change the distribution to

$$f'(x, y) = B(x, y) f(x, y).$$

Since $B(x, y)$ is slowly varying as compared with $f(x, y)$, we may approximate it by the lowest frequency terms in a two-dimensional Fourier series expansion of $f'(x, y)$ when $f(x, y)$ is a constant (that is, when a clear field is recorded). Thus we have

$$B(x, y) \approx 1 + B_1 \sin(b_1 x) + B_2 \cos(b_2 x) + B_3 \sin(b_3 y) + B_4 \cos(b_4 y) + B_5 \sin(b_5 x) \sin(b_6 y).$$

By dividing the recorded picture $f'(x, y)$ by the calculated $B(x, y)$, pictures are obtained that have good contrast and uniform background illumination.

3. Pattern Recognition

Before a leukocyte can be identified it must first be located in the $70 \mu \times 50 \mu$ field of view. Since there are approximately 650 red cells for every leukocyte, the problem of overlapping leukocytes can be ignored. Thus if we consider the picture to be

represented by its 18,500 points, every point can be classified as belonging to one of three types: (a) background points; (b) red cell points; (c) leukocyte points.

In order to speed computation, the background points are immediately eliminated from further consideration by the following technique. For each of the three one-color pictures a histogram is computed of the first-order brightness distribution. A sample histogram is shown in Fig. XVII-22a. The histogram is then smoothed as shown in Fig. XVII-22b, and the relative maxima and minima are located on the smoothed histogram. By analysis of several pictures, we have found that the second minimum from the left represents a clipping level to separate background points from red and white cell points. The results of this type of clipping are shown in Fig. XVII-23.

To separate the remaining points in the picture into the two classes – red cell and leukocyte – the "classical" approach of pattern recognition was used. In this approach an appropriate feature or set of features is determined and then, by using an appropriate decision function on the resulting feature space, unknown objects may be classified by measurement of these features. The feature used in our case is the two-dimensional chromaticity per point. Briefly, chromaticity may be defined as follows. Consider the spectral distribution S_λ of a point in the color picture. By passing the light from this point through the two dichroic mirrors of the SCAD system, the three digitally recorded numbers (red, green, and blue) are given by the expressions

$$R = \int C_1(\lambda) S_\lambda d\lambda$$

$$G = \int C_2(\lambda) S_\lambda d\lambda$$

$$B = \int C_3(\lambda) S_\lambda d\lambda,$$

where the $C_K(\lambda)$ are the spectral transmission functions of the respective light paths. The chromaticity pair (r, g) per point is computed by taking the ratios

$$r = \frac{R}{R + G + B}$$

$$g = \frac{G}{R + G + B}.$$

By computing chromaticity statistics over a set of known and identified points and then using a maximum-likelihood decision function to classify unknown points, we have had reasonable success in separating red cell points from leukocyte points (Fig. XVII-24).

MAXIMUM VALUE = 833

EACH SYMBOL = 12

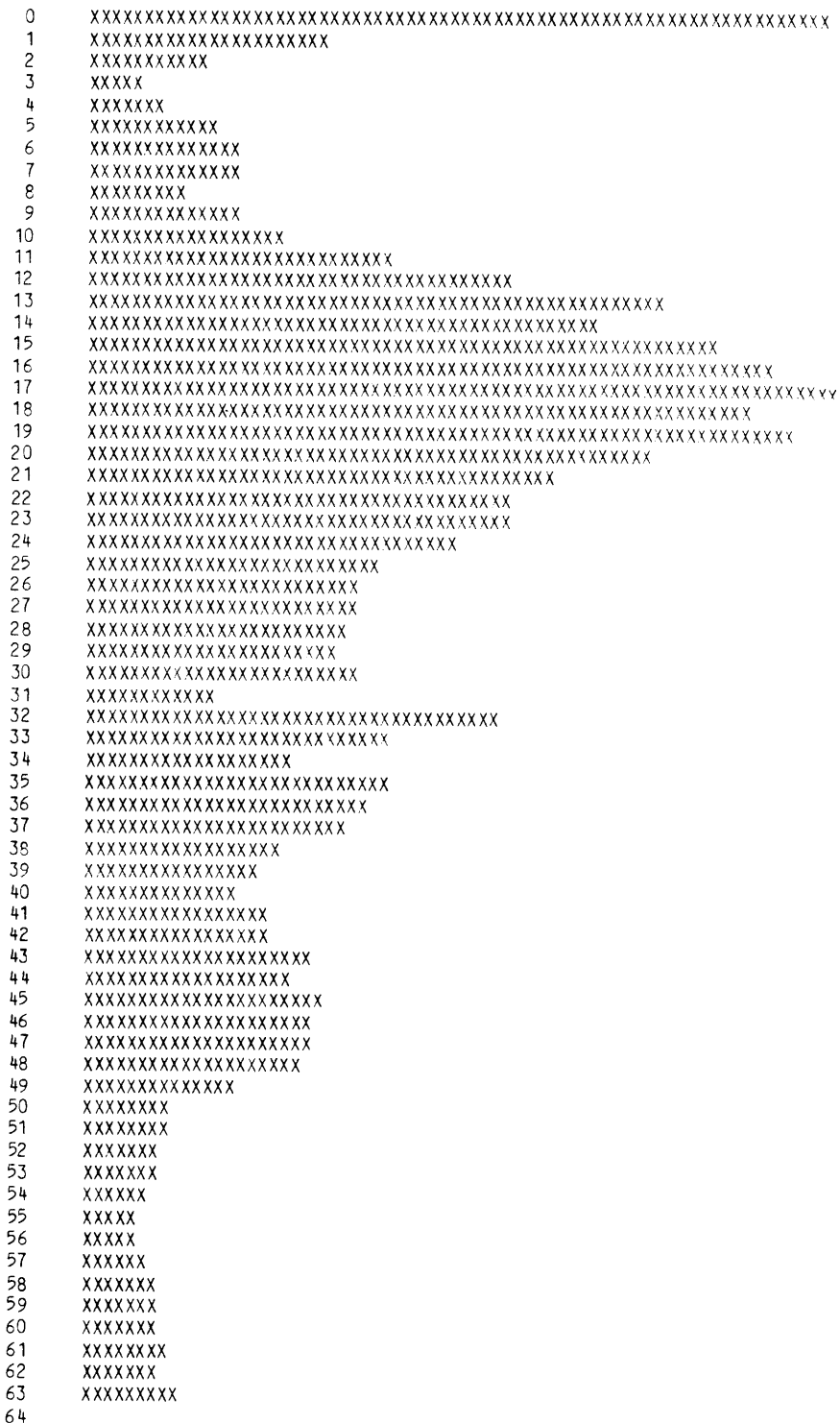


Fig. XVII-22. (a) First-order brightness distribution of the "red" picture.

MAXIMUM VALUE = 976

EACH SYMBOL = 14

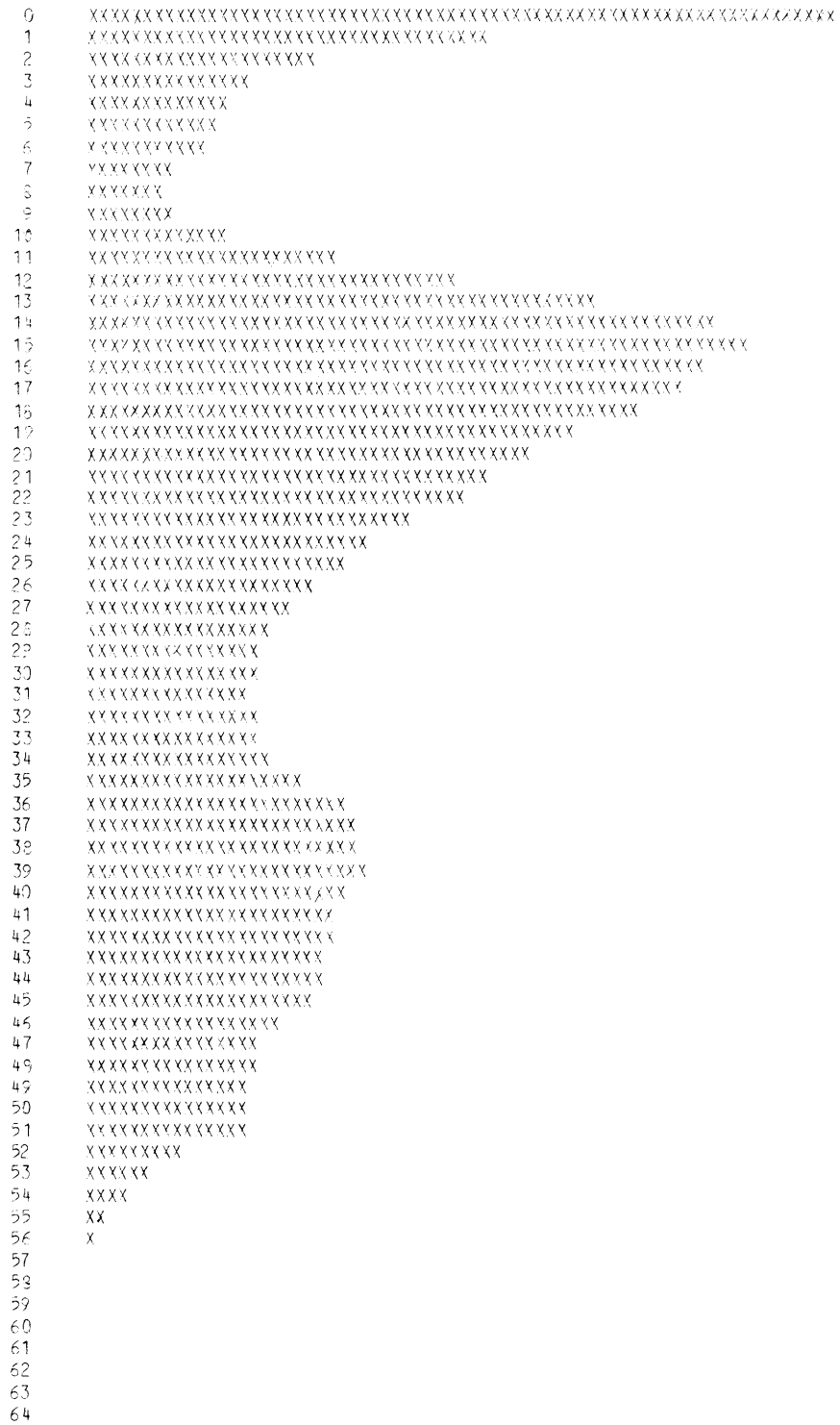


Fig. XVII-22. (b) Smoothed distribution of the "red" picture.

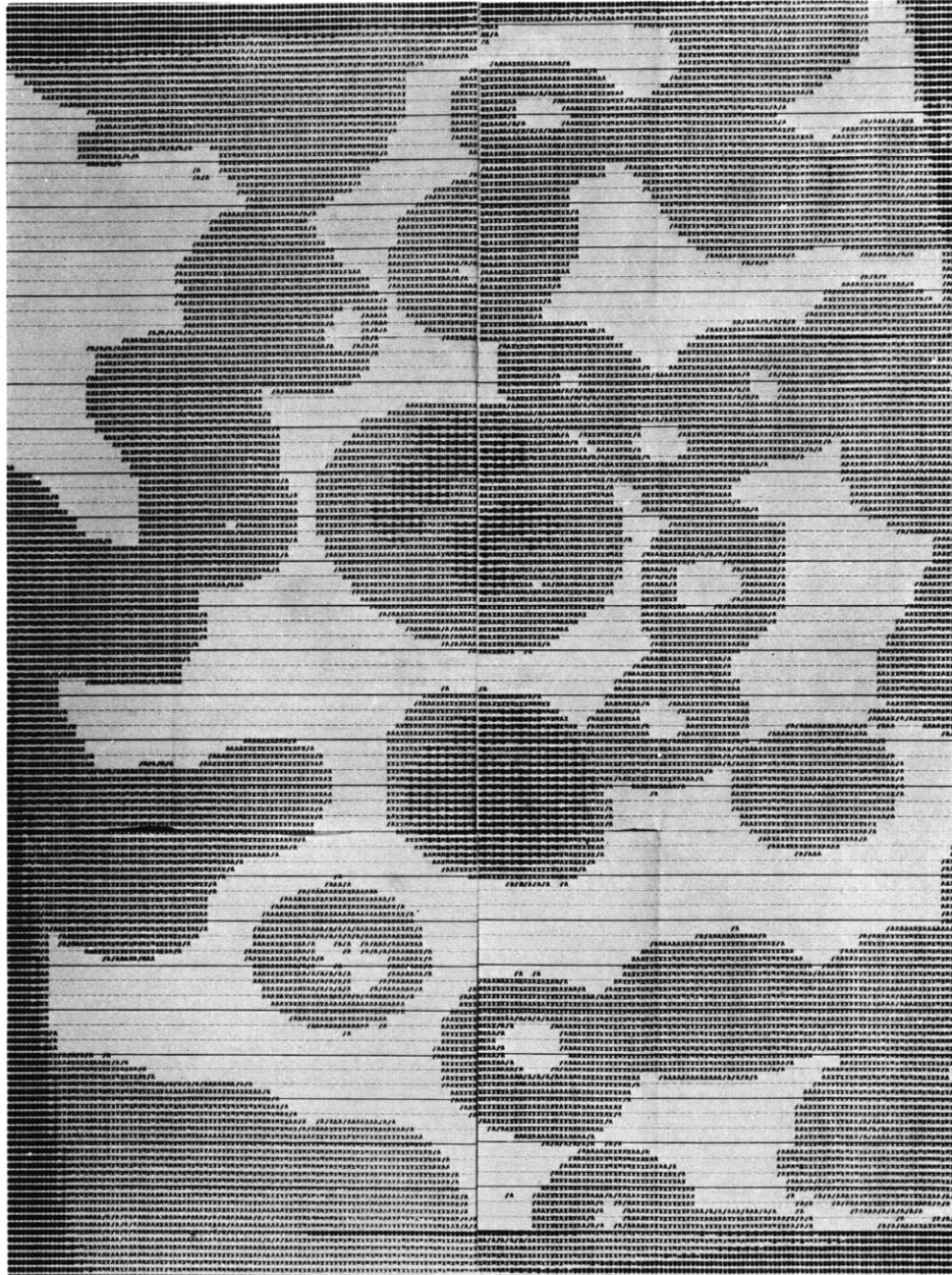


Fig. XVII-23. "Red" picture with background removed.

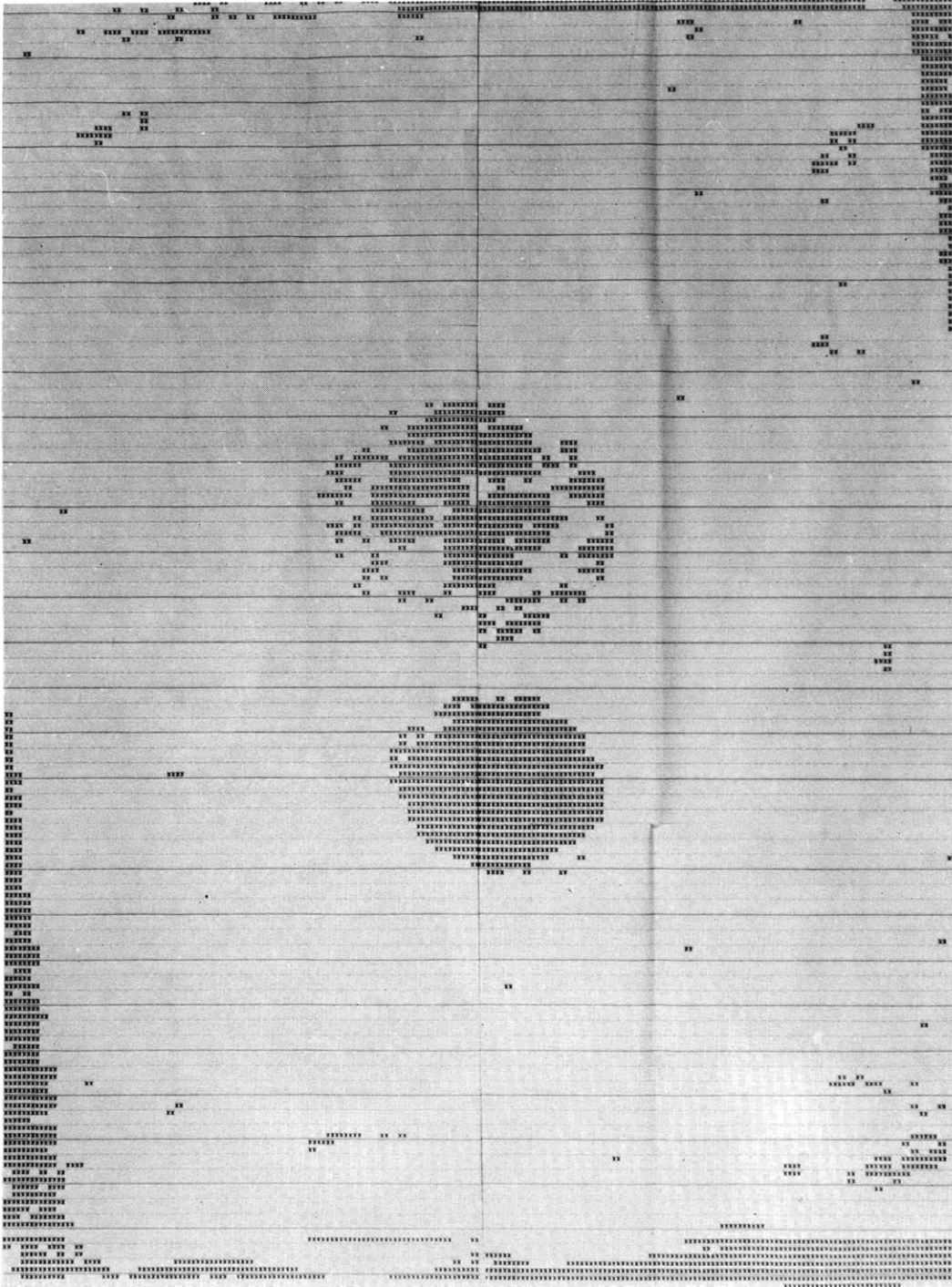


Fig. XVII-24. "Red" picture with leukocyte points identified.

(XVII. COGNITIVE INFORMATION PROCESSING)

4. Summary

We have recently recorded and are in the process of analyzing 6 more full-color blood-smear pictures to obtain refined chromaticity and shape statistics.

Using these additional pictures and the investigation of specific morphological properties of leukocytes and their nuclei, we expect to complete our task of leukocyte differentiation.

I. T. Young

References

1. I. Davidsohn and B. Wells, Clinical Diagnosis by Laboratory Methods (W. B. Saunders Company, London, 13th edition, 1965), Chaps. 1, 2, and 4.
2. M. L. Mendelsohn and B. Perry, "Picture Generation with a Standard Line Printer," Communs. ACM, Vol. 7, No. 5, pp. 311-313, May 1964.
3. M. L. Mendelsohn and J. Prewitt, "Analysis of Cell Images," Ann. N. Y. Acad. Sci. 128, pp. 1035-1053, January 1966.
4. O. J. Tretiak, "Scanner Display - SCAD," Quarterly Progress Report No. 83, Research Laboratory of Electronics, M.I.T., October 15, 1966, pp. 129-142.
5. G. S. Sebestyen, Decision Making Processes in Pattern Recognition (Macmillan Company, New York, 1962).

E. DIFFERENTIAL LEUKOCYTE ANALYSIS BY OPTICAL METHODS

We are investigating the possibility of differentiating leukocytes from the characteristics of their scattered light, rather than by examining their image. It is well known that if an object is illuminated with collimated light along the axis of the optical system, the spatial Fourier transform of the object will be seen in the back focal plane of the system. We hope to use this property in an arrangement as shown in Fig. XVII-25.

Our primary reason for considering this system is that the magnitude of the Fourier transform, that is, the scattered light in the back focal plane, does not depend (in principle) on the distance between the cell and the lens, nor on the position of the cell relative to the axis of the optical system. Therefore the cell need not be either registered

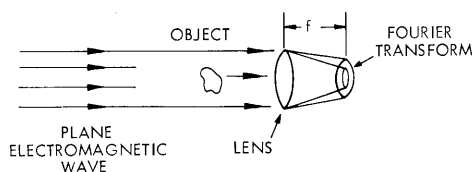


Fig. XVII-25. Fourier transform property of lenses.

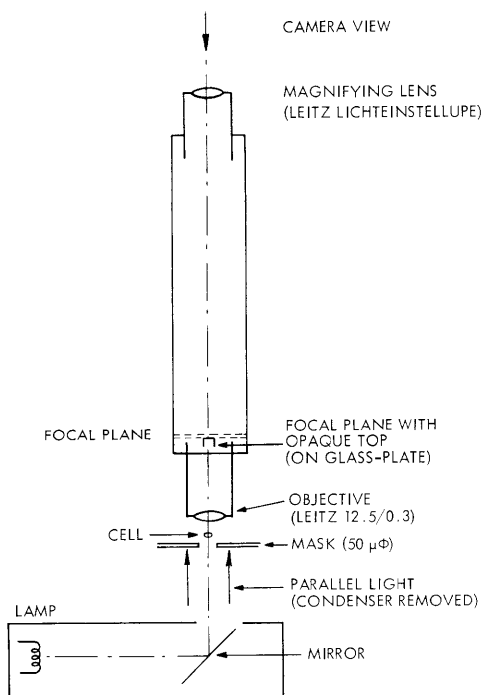


Fig. XVII-26. Microscope changed to perform Fourier transform of leukocytes.

or in focus. Thus this technique can be applied in a system in which the cells flow in a liquid past some sort of sensing element. It seems obvious that a system like this would be much easier to automate than one in which the cells must be prepared on a slide and scanned in focus.

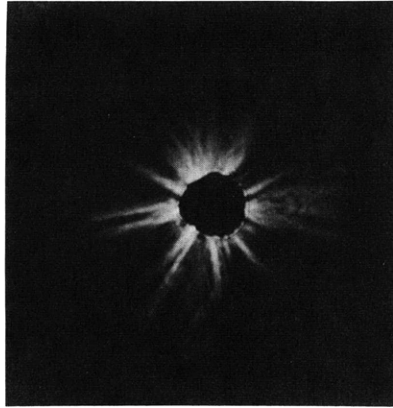
(XVII. COGNITIVE INFORMATION PROCESSING)

Another possibility is that since index of refraction variations scatter light, it may be possible, in the system under discussion, to differentiate unstained cells.

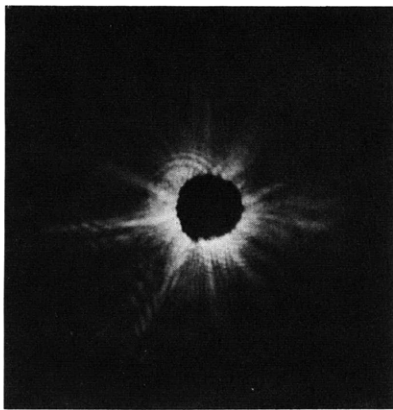
The signal that can be recorded in the system under consideration is essentially the squared amplitude of the Fourier transform. Since this is the transform of the autocorrelation function, we see that the final signal will contain much less information than the original image. The major leukocyte types differ with respect to the following attributes: size, shape or nucleus, and granules in the cytoplasm.

The size differences will affect the spatial Fourier transform at approximately 10^{+5} m^{-1} , the nucleus shape at approximately $3 \times 10^{+5} \text{ m}^{-1}$, and the nuclear granules at $3 \times 10^6 \text{ m}^{-1}$.

The system shown in Fig. XVII-26 was used for our initial experiments. Since the diameter of a single cell is only approximately 10μ , we expected that some effort would be needed to overcome noise problems, that is, the background of the Fourier transform produced by light scattering inside the lens, by diffraction by its margins, and by general reflection inside the system. The experimental arrangement consisted of a microscope changed in the following manner: (i) We removed the condenser to illuminate the picture with parallel light. (ii) We replaced the ocular by a magnifying lens, which enlarged the focal plane of the objective, in which the Fourier transform of the cells occurs. (iii) We placed an aperture of approximately 50μ diameter behind the cell to keep the illuminating beam as small as possible in order to reduce the background. (iv) We mounted a small opaque stop in the focal point of the objective to absorb the DC component of the Fourier spectrum. With this arrangement signals of single leukocytes could be detected. Figure XVII-27a shows the background of the system, and Fig. XVII-27b shows the background and the additional signal produced by two unstained leukocytes.



(a)



(b)

Fig. XVII-27.

(a) Background of the Fourier transform system.

(b) Signals of two leukocytes.

The system does not seem to produce very accurate Fourier spectra, and before trying some different optical systems we have decided to study some promising configurations by simulating them with a computer.

P. H. Grassmann, O. J. Tretiak

F. READING MACHINE FOR THE BLIND

Members of our group are working on experimental reading machine systems with the eventual aim of providing sensory aids for the blind. The present experimental system was not intended as a prototype of a practical reading aid but rather as a real-time research facility for experimental investigation of human information requirements and human learning capabilities upon which prototype designs must be based. The experimental system development has involved, and will facilitate, research on print scanning, character recognition, control by a blind reader, and auditory and tactile displays.

The reading machine system diagram is shown in Fig. XVII-28. The PDP-1 com-

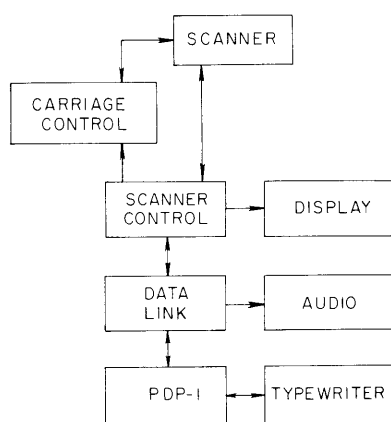


Fig. XVII-28.

Reading machine system diagram.

The scanner control is an open-loop digital servo that permits a page of text to be moved in horizontal and vertical directions at approximately 200 steps per second, where there are approximately 4 steps per letter.

The opaque scanner is shown in Fig. XVII-29. For the cathode-ray tube we used a Tektronix oscilloscope mounted on end. The dot of light from the cathode-ray tube is imaged through a single lens onto the paper, and the diffuse reflected light is collected by means of 2 photomultiplier tubes. When the scanner controller wishes to determine the blackness of a particular location on the page, it supplies the proper xy signals to the oscilloscope, unblanks the beam, and the output of the photomultiplier tubes is integrated for a fixed time. If this integrated output exceeds a threshold voltage, then the point is said to be white.

In Fig. XVII-30 the over-all sequence of operation is indicated. The initialization of the program also includes selecting the proper initial position for the carriage. The first line of text is then acquired, by using a special directive provided by the scanner

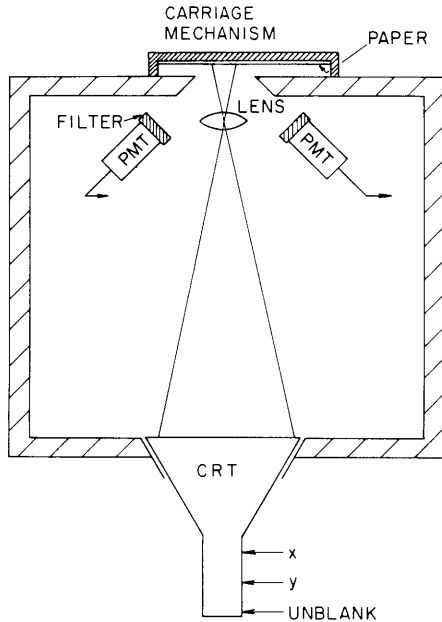


Fig. XVII-29. Opaque scanner.

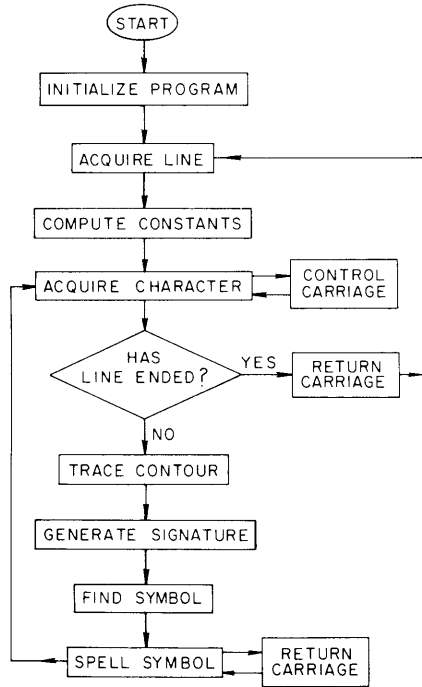


Fig. XVII-30. PRM-1 flow chart.

control. This is where the size of the type font is determined and various constants are computed to aid in the recognition of punctuation and in the rejection of miscellaneous specks. After this, the first character is acquired, with a carriage motion if it is necessary, and it is determined whether or not the line has ended. If not, the external contour of the letter is traced, again by using another special directive provided by the scanner control, and a characteristic number or signature is generated. After the signature has been generated, a table of previously encountered signatures is searched for a match. If a match is found, a symbol is retrieved and an auditory output is initiated, after which the next character is acquired and the process repeats.

The process for acquiring a line of text uses a horizontal histogram that is computed by determining the total number of black points at each y level and normalizing the curve. A simple threshold procedure is used to determine the two interior lines, and the remaining lines are computed from these two lines. As successive letters are traced, these y lines are updated, and the carriage moved vertically, if necessary, to track text that is not exactly aligned horizontally.

The next two operations are the acquisition scan and the contour trace. The vertical limits of the acquisition scan are the two exterior lines as determined from the horizontal histogram. The scan starts at the lower left, traces a vertical line, and then indexes to the right by one and repeats the vertical line until it encounters a black point. At this point the contour trace instruction is given, and if the exterior of the

black area is very small, it is considered to be a speck and ignored. If it is another size classification, it is deemed punctuation, and a special routine is used to determine which of the punctuation marks it is.

Figure XVII-31 contains a clear presentation of a character contour and also gives an indication of the resolution of the scanner.

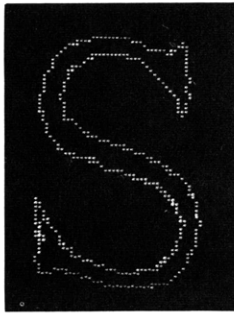


Fig. XVII-31. Capital 'S' contour.

a do not introduce extraneous extrema. The sequence of extrema can be coded with one bit per extremum if the starting point is assumed to be an x minimum and the contour trace is clockwise.

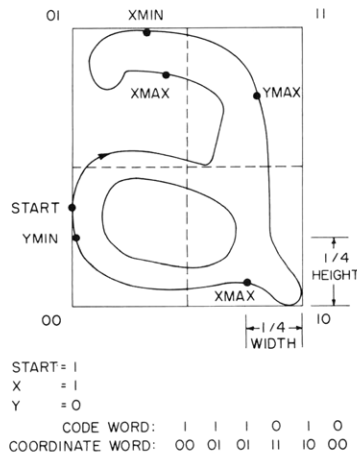


Fig. XVII-32.

Determination of extrema.

If the circumference of the black area is large enough that it is deemed a character, a signature is computed by the method indicated in Fig. XVII-32. The starting point is always at the westernmost point. Local extrema are determined for the horizontal and vertical directions if the excursions along the contour are greater than one-quarter of the letter height or width. That is, for example, an x maximum is not denoted unless one had to travel east for at least one-quarter of the letter width and then west for one-quarter of the letter width. With this scheme, minor features of the letter such as the droop in the upper left part of the letter do not introduce extraneous extrema. The sequence of extrema can be coded with one bit per extremum if the starting point is assumed to be an x minimum and the contour trace is clockwise. In addition to this sequence of extrema, the quadrant in which it was determined is stored for such extrema. These binary sequences, together with a small amount of information concerning the aspect ratio of the letter and its position relative to the y lines found by the histogram procedure, comprise the signature.

In order for characters to be recognized, the particular signature must have been encountered previously; that is, the machine must have been previously trained by a sighted operator. To enable the sighted operator to recognize the letters, a visual display is made. When the machine is unsuccessful in the search for a match for a particular signature, it can effectively ask the trainer what the character is. Given this information, it can thus add to its experience. There is a certain

amount of noise involved in the determination of black and white, and the presence of this noise actually simplifies the training procedure, as it appears that successive tracings of the same a have variations similar to tracings of different a's. Thus with

(XVII. COGNITIVE INFORMATION PROCESSING)

the carriage control disabled, one can lock on to a particular letter and build a repertoire of signatures for that particular letter.

In several accuracy tests, the results have been about the same, with an error rate of approximately one-third of one per cent. Training has been virtually completed on two fonts and others are in process. Punctuation marks are recognized and are spelled out to minimize the storage space required for the spelled-speech output. The spelled-speech output consists of letters that have been digitized and artificially shortened so that their duration is under 100 msec. This puts a limit of 120 wpm as the maximum output rate. The present machine, however, alternates the speech output with the recognition process, and the average speed is approximately 80 wpm.

S. J. Mason, F. F. Lee, D. E. Troxel