XIV. COGNITIVE INFORMATION PROCESSING[*]

Academic and Research Staff

Prof. M. Eden
Prof. J. Allen
Prof. B. A. Blesser
Prof. T. S. Huang
Prof. F. F. Lee
Prof. S. J. Mason
Prof. W. F. Schreiber

Prof. D. E. Troxel
Prof. I. T. Young
Dr. R. R. Archer
Dr. J. E. Green
Dr. K. R. Ingham
Dr. D. A. Sankowsky
Dr. O. J. Tretiak

F. X. Carroll
C. L. Fontaine
E. R. Jensen
A. J. Laurino
D. M. Ozonoff
E. Peper
Sandra A. Sommers

Graduate Students

T. P. Barnwell III
W. L. Bass
T. R. Bourk
J. E. Bowie
B. E. Boyle
R. L. Brenner
S. A. Ellias
J. R. Ellis, Jr.
A. E. Filip
R. E. Greenwood
W. B. Grossman

E. G. Guttmann
D. W. Hartman
L. Hatfield
L. P. A. Henckels
P. D. Henshaw
M. Hubelbank
J. I. Makhoul
G. P. Marston
G. G. Matison
W. McCandless

G. F. Pfister
D. S. Prerau
N. Rashid
G. M. Robbins
C. L. Seitz
R. D. Shapiro
W. W. Stallings
K. P. Wacks
G. A. Walpert
T. R. Willemain
Y. D. Willems

A. SINC^n — A FAMILY OF WINDOW FUNCTIONS

1. Introduction

Window functions have been useful in the estimation of power spectra from finite-time records. The autocorrelation function is ordinarily multiplied by a window function in order to improve the spectral estimate.[1,2] More recently, window functions have been used to reduce "leakage"[3] in the evaluation of short-time spectra as computed by using the fast Fourier transform.

Finite window functions have the property that they assume nonzero values only within a finite interval. Without loss of generality, the finite interval can be centered around the origin (t=0), so that

$$h(t) = 0, \ |t| > T, \tag{1}$$

where h(t) is the window function, and 2T is the length of the interval where h(t) can assume nonzero values. Sometimes 2T is referred to as the "window size."

Well-known window functions include the rectangular Barlett, Parzen, Hanning, and Hamming windows.[1,2] The Fourier transforms corresponding to the rectangular Barlett and Parzen windows have in common the fact that they are of the form $\left(\frac{\sin x}{x}\right)^n$, where

---

n = 1 for the rectangular window, n = 2 for the Barlett window, and n = 4 for the Parzen window. This report derives a general closed-form expression in the time domain for the family of window functions whose Fourier transforms are of the form $\left(\frac{\sin x}{x}\right)^n$, where n is any positive integer. Because the function $\frac{\sin \pi x}{\pi x}$ is sometimes referred to as sinc x, we shall call this family of window functions $SINC^n$.

2. Derivation

A window function $h_n(t)$, corresponding to the nth member of the family $SINC^n$, is defined by its transform $H_n(\omega)$ as

$$H_n(\omega) = \left(\frac{\sin a\omega}{a\omega}\right)^n,$$ (2)

where n is a positive integer, and a is a constant. $h_n(t)$ is the inverse Fourier transform of $H_n(\omega)$:

$$h_n(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} H_n(\omega)\, e^{j\omega t}\, d\omega$$

$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left(\frac{\sin a\omega}{a\omega}\right)^n e^{j\omega t}\, d\omega.$$ (3)

The problem is to compute the integral in Eq. 3 for any n.

The integral can be simplified by the help of two expansions[4]:

$$\sin^{2m} x = \frac{1}{2^{2m-1}} \left[\sum_{k=0}^{m-1} (-1)^{m+k} \binom{2m}{k} \cos (2m-2k)x + \frac{1}{2}\binom{2m}{m} \right]$$ (4)

$$\sin^{2m+1} x = \frac{1}{2^{2m}} \left[\sum_{k=0}^{m} (-1)^{m+k} \binom{2m+1}{k} \sin (2m+1-2k)x \right],$$ (5)

where

$$\binom{m}{k} = \frac{m!}{k!\,(m-k)!}, \qquad m \geqslant 0.$$

It is clear that if Eqs. 4 and 5 are properly substituted in Eq. 3, the integral must be computed separately for odd and even n.

(a) <u>n even</u>

Let $m = \frac{n}{2}$ and $x = a\omega$ in Eq. 4. Then substitute Eq. 4 in Eq. 3. The integral reduces to

$$h_n(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{1}{2^{n-1}} \left[\sum_{k=0}^{\frac{n}{2}-1} (-1)^{\frac{n}{2}+k} \binom{n}{k} \cos(an-2ak)\omega + \frac{1}{2} \binom{n}{n/2} \right] \frac{e^{j\omega t}}{(a\omega)^n} d\omega. \qquad (6)$$

Place the following substitution in Eq. 6.

$$\cos(an-2ak)\omega = \frac{1}{2} \left[e^{j(an-2ak)\omega} + e^{-j(an-2ak)\omega} \right].$$

The integral can now be evaluated by making use of the Fourier transform pair

$$\frac{e^{jb\omega}}{\omega^n} \longleftrightarrow \frac{j^n}{(n-1)!} (t+b)^{n-1} u_{-1}(t+b), \qquad (7)$$

where $u_{-1}(t)$ is the ordinary unit-step function, and b is a constant. The result of the integration can be shown to be

$$h_n(t) = \frac{1}{(2a)^n (n-1)!} \left\{ \sum_{k=0}^{\frac{n}{2}-1} (-1)^k \binom{n}{k} \left[(t+an-2ak)^{n-1} u_{-1}(t+an-2ak) \right. \right.$$

$$\left. + (t-an+2ak)^{n-1} u_{-1}(t-an+2ak) \right]$$

$$\left. + (-1)^{\frac{n}{2}} \binom{n}{n/2} t^{n-1} u_{-1}(t) \right\}. \qquad (8)$$

Equation 8 can be simplified. Since $H_n(\omega)$ is a real and even function of ω, it follows that $h_n(t)$ is a real and even function of time. Therefore it suffices to compute $h_n(t)$ for $t < 0$ and then replace every t by $(-t)$ in order to obtain $h_n(t)$ for $t > 0$.

For $t < 0$ Eq. 8 reduces to

$$h_n(t) = \frac{1}{(2a)^n (n-1)!} \sum_{k=0}^{\frac{n}{2}-1} (-1)^k \binom{n}{k} (t+an-2ak)^{n-1} u_{-1}(t+an-2ak). \qquad (9)$$

(Note that $2k < n$ always applies.) $h_n(t)$ for $t > 0$ can be obtained from Eq. 9 by replacing t by $(-t)$. Therefore, it follows that for <u>all t</u>

$$h_n(t) = \frac{1}{(2a)^n (n-1)!} \sum_{k=0}^{\frac{n}{2}-1} (-1)^k \binom{n}{k} (an-2ak-|t|)^{n-1} u_{-1}(an-2ak-|t|), \tag{10}$$

where $|t| \equiv$ absolute value of t.

(b) <u>n odd</u>

Let $m = \frac{n-1}{2}$ and $x = a\omega$ in Eq. 4. Then substitute Eq. 4 in Eq. 3. The integral reduces to

$$h_n(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{1}{2^{n-1}} \left[\sum_{k=0}^{\frac{n-1}{2}} (-1)^{\frac{n-1}{2}+k} \binom{n}{k} \sin(an-2ak)\omega \right] \frac{e^{j\omega t}}{(a\omega)^n} d\omega. \tag{11}$$

Substituting

$$\sin(an-2ak)\omega = \frac{1}{2j}\left[e^{j(an-2ak)\omega} - e^{-j(an-2ak)\omega} \right]$$

in Eq. 11 and making use of the transform pair in Eq. 7, we obtain

$$h_n(t) = \frac{1}{(2a)^n (n-1)!} \left\{ \sum_{k=0}^{\frac{n-1}{2}} (-1)^k \binom{n}{k} \left[(t+an-2ak)^{n-1} u_{-1}(t+an-2ak) \right. \right.$$

$$\left. \left. - (t-an+2ak)^{n-1} u_{-1}(t-an+2ak) \right] \right\}. \tag{12}$$

By using the fact that $h_n(t)$ is an even function of t, Eq. 12 reduces to

$$h_n(t) = \frac{1}{(2a)^n (n-1)!} \sum_{k=0}^{\frac{n-1}{2}} (-1)^k \binom{n}{k} (an-2ak-|t|)^{n-1} u_{-1}(an-2ak-|t|) \tag{13}$$

for all t.

(c) n even or odd

Equations 10 and 13 for even and odd n, respectively, are very similar. The only difference is the upper limit of the summation on k. By defining

$$\left[\frac{n-1}{2} \right] \equiv \text{integer portion of } \frac{n-1}{2},$$

both equations can be combined into one equation which is valid both for even and for odd n, and for all t:

$$h_n(t) = \frac{1}{(2a)^n (n-1)!} \sum_{k=0}^{\left[\frac{n-1}{2}\right]} (-1)^k \binom{n}{k} (an-2ak-|t|)^{n-1} u_{-1}(an-2ak-|t|). \tag{14}$$

Equation 14 completely specifies the family of window functions $SINC^n$ in the time domain.

3. Discussion

From the properties of the unit-step function it is easy to show from Eq. 14 that

$$h_n(t) = 0, \quad \text{for} \quad |t| > an. \tag{15}$$

Comparing Eq. 15 with Eq. 1, we conclude that

$$T = an. \tag{16}$$

The family of functions $h_n(t)$ is time-limited, and the window size is equal to 2an.

A property of $h_n(t)$ is that the area under the function is independent of $\underline{a}$ and $\underline{n}$ and is equal to 1. This can be easily shown.

$$\int_{-an}^{+an} h_n(t) \, dt = \int_{-\infty}^{+\infty} h_n(t) \, dt = H_n(0) = 1. \tag{17}$$

If the window is of constant size 2T, that is, a = T/n, then, as n increases, the window shape gets "thinner," but it also gets more peaked in order to keep the area of the window equal to 1. For this case of constant window size, $H_n(\omega)$ takes the form

$$H_n(\omega) = \left[\frac{\sin (T\omega/n)}{T\omega/n} \right]^n. \tag{18}$$

As n increases, the width of the main lobe increases proportionally with n, which

results in an increase in the window spectral bandwidth. At the same time, the sidelobe amplitude levels decrease with increasing n, and the high-frequency spectral roll-off is equal to $-6n$ dB/octave. Therefore, there is a trade-off between a finer spectral resolution and a reduction in leakage.

For any particular value of n it would be possible to obtain $h_n(t)$ by an $(n-1)$-fold convolution of the rectangular window. The advantage of Eq. 14 is that it gives $h_n(t)$ for any value of n by a simple summation. The reader may be interested in checking the validity of Eq. 14 for the well-known rectangular, Bartlett, and Parzen windows.

J. I. Makhoul

References

1. R. B. Blackman and J. W. Tukey, The Measurement of Power Spectra (Dover Publications, Inc. , New York, 1958).

2. P. I. Richards, "Computing Reliable Power Spectra," IEEE Spectrum, pp. 83-90, January 1967.

3. G. D. Bergland, "A Guided Tour of the Fast Fourier Transform," IEEE Spectrum, pp. 41-52, July 1969.

4. I. S. Gradshteyn and I. M. Ryzhik, Tables of Integrals, Sums, Sequences and Products (FIZMATGIZ, Moscow, 1963).

B. RELEVANCE OF PSYCHOLOGY TO CHARACTER RECOGNITION

The two disciplines which may be considered to dominate the field of character recognition are cognitive psychology and mathematical implementation of algorithms. The ultimate goal of character recognition is to characterize the visual pattern and to judge which of the possible legal set elements (in our case printed letters from an alphabet) is being represented. These algorithms are unfortunately very ad hoc and difficult to relate to human perception. Moreover, an algorithm is usually evaluated under restrictive conditions such that a high statistical success is used as evidence of an adequate solution. An algorithm yielding a 95% accuracy rate when letters are restricted to 12 Point Bodoni type, but which confuses a Bodoni "a" with an "Old Style" "s" cannot be considered as approaching a solution to the general character-recognition problem.

The argument that the algorithm should be psychologically related to human perception is often met with an apparently valid comment such as "Flying machines are not built on the principle of a bird's flapping wing, rather they are based on engineering and aerodynamic principles." The analogy is not good because there are no known physical attributes that give a letter its unique value; equivalence of stimuli is purely perceptual. Consequently, the test of whether two particular characters are the letter "z" must be judged by a human observer. Any ad hoc algorithm, of which there are many, that does

not use perceptually related features and attributes for making its decisions becomes burdened with many extra limitations, e.g., the visual data must be very clear, the letters must be taken from one type font, and multiple definitions are needed to over-come the effects of stimulus and system noise. The best long-range solution to the problem of character recognition will be by developing an algorithm that, although math-ematical in formulation, incorporates essential perceptual cues.

Psychological experiments generally indicate that the information contained in a letter is not uniformly distributed and is very redundant; some parts of the letter give essential cues necessary for resolving possible confusions, while other parts of the letter can be easily ignored. For example, text can be read when the bottom half of a line is deleted, but not when the top half is removed. The top half clearly has more relevant information than the bottom half. In the same way of thinking, the capital let-ter "A" can be described as having a centered closure at the top with two legs extending downward. This gives the essence of the "A", but there is the possibility of ambiguity with the letter "R". The ambiguity is clearly resolved by the presence of any discon-tinuity in the position and derivative of the rightmost leg at its intersection with the lower right side of the closure. From this quick analysis we can specify the psychological space and the boundary between these two letters so that this description will success-fully characterize any "A" or "R" that a human being would correctly recognize.

Our investigation is now being carried out in many directions.

1. Many letter features can be discovered by an introspective process that prompts one to ask, "How can I distort a letter so that it becomes another letter." Namely, find some style of letter in which a small change produces a new character. For example, a slight indentation in the capital "D" will make it be perceived as a "B".

2. In one experiment, subjects were asked to judge a series of defocused or obscured letters. The resulting confusion matrices portray a space with different groups of letters clustering together. For example, the letters "a", "c", "e", "x", and "z" seem to cluster in one group, whereas, the letters "b", "d", "p", and "l" cluster in another group. This would lead to the tentative conclusion that under defocusing the attributes compactness vs extendedness are perceived by the subjects.

3. We are attempting to gain some insight into a person's ability to discard visual data that are not relevant to the percept, and his ability to fill in deletion in order to complete his percept. This appears to be most difficult to accomplish with computing systems.

4. An algorithm that represents letters as a sequence of topological features, such as end point, intersection, and closure, is being investigated. We hope that the resulting encoding will have a one-to-one correspondence with the letter set.

5. The ultimate limitations of single-letter recognition can be seen in an experiment in which subjects were unable to recognize defocused letters, but were able to read

defocused sentences containing the same letters with the same degree of defocusing. Similarly, when subjects were presented with a sequence of letters from a sentence in rapid succession they were not able to read it. The sentence could easily be read, however, when all of the letters were presented simultaneously, even though the total exposure time was the same in both cases. These familiar examples of context show the inherent limitation of any mechanized letter-by-letter recognition algorithm.

B. A. Blesser, E. Peper

C. SOURCE RATE FOR A SIGNAL CORRUPTED BY NOISE

Consider a signal that is obtained from an imperfect transducer. We would like to know the rate that is required to code this signal within a certain fidelity. The situation
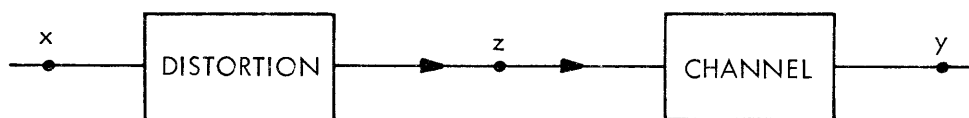

Fig. XIV-1. System under consideration.

is diagrammed in Fig. XIV-1. The true signal is x, we observe the corrupted signal, z, and the signal at the output of the channel is y.

1. Stereotype Formulation

We shall consider the problem of transmitting a single random variable. Since the channel and the transducer are disjoint, we assert that

$$P_{x,y/z}(\xi, \omega, \zeta) = P_{x/z}(\xi, \zeta)\, P_{y/z}(\omega, \zeta). \tag{1}$$

From this we obtain the following relations:

$$P_{x,y,z}(\xi, \omega, \zeta) = P_{x/z}(\xi, \zeta)\, P_{y,z}(\omega, \zeta) \tag{2}$$

$$P_{x,y}(\xi, \omega) = \int d\zeta\, P_{x/z}(\xi, \zeta)\, P_{y,z}(\omega, \zeta). \tag{3}$$

By using the standard formulation,[1] the distortion from the total system is

$$\epsilon = \iint d\xi d\omega\, P_{x,y}(\xi, \omega)\, \rho(\xi, \omega), \tag{4}$$

where $\rho(\xi, \omega)$ is the distortion weighting function. The rate is given by

$$R = \iint d\zeta d\omega \; P_{y,z}(\omega, \zeta) \log \frac{P_{y,z}(\omega, \zeta)}{P_y(\omega) \; P_z(\zeta)}. \tag{5}$$

The rate is to be minimized for a fixed value of ϵ. The problem can be cast into a form for which the solution is available. By substituting (3) in (4), we obtain

$$\epsilon = \iiint d\xi d\omega d\zeta \; P_{x/z}(\xi, \zeta) \; P_{y,z}(\omega, \zeta) \; \rho(\xi, \omega). \tag{6}$$

If we define a new function

$$\rho'(\omega, \zeta) = \int d\xi \; P_{x/z}(\xi, \zeta) \; \rho(\xi, \omega) \tag{7}$$

the new form of ϵ becomes

$$\epsilon = \iint d\omega d\zeta \; P_{y,z}(\omega, \zeta) \; \rho'(\omega, \zeta). \tag{8}$$

The solution for these equations is known,[2] and is

$$P_{z,y}(\zeta, \omega) = \mu(\zeta) \; e^{-\ell \rho'(\omega, \zeta)}. \tag{9}$$

In this equation ℓ is a parameter that determines the rate and the distortion and $\mu(\zeta)$ is a function chosen so that

$$\int P_{z,y}(\zeta, \omega) \, d\zeta = 1.$$

2. Solution

Suppose that the source has a Gaussian power density function and the corruption is caused by adding an independent Gaussian random variable.

$$P_x(\xi) = \frac{1}{\sqrt{2\pi} \; \sigma_x} \exp\left(-\frac{1}{2} \frac{\xi^2}{\sigma_x^2}\right) \tag{10}$$

$$P_{z/x}(\zeta, \xi) = \frac{1}{\sqrt{2\pi} \; \sigma_n} \exp\left[-\frac{1}{2} \frac{(\xi-\zeta)^2}{\sigma_n^2}\right]. \tag{11}$$

We use squared difference as the error weighting function

$$\rho(x,y) = (x-y)^2. \tag{12}$$

With these hypotheses, the function ρ' takes an attractive form

$$\rho'(\omega, \zeta) = \left(\zeta \frac{\sigma_x^2}{\sigma_x^2 + \sigma_n^2} - \omega \right)^2 + \frac{\sigma_x^2 \sigma_n^2}{\sigma_x^2 + \sigma_n^2}. \tag{13}$$

It follows from (9) that

$$P_{z,y}(\zeta, \omega) = c_1 \exp\left[-\ell \left(\frac{\sigma_x^2}{\sigma_x^2 + \sigma_n^2} \zeta - \omega \right)^2 \right]. \tag{14}$$

The various probability density functions that are required for finding the rate and the distortion are found by using moment-generating functions. We shall represent the Fourier transform of g by the symbol $\hat{g}$. With

$$P_{z,y}(\zeta, \omega) = f_1(a\zeta - \omega) \tag{15a}$$

$$P_z(\zeta) = f_2(\zeta) \tag{15b}$$

the moment-generating function for $P_{y,z}$ is

$$\hat{P}_{y,z}(\mu, \nu) = \frac{\hat{f}_1\left(\frac{\mu}{a}\right) \hat{f}_2(\mu + a\nu)}{\hat{f}_1\left(\frac{\mu}{a} + \nu\right)}. \tag{16}$$

In the case that is being considered

$$f_1(\delta) = \frac{1}{\sqrt{2\pi\lambda}} e^{-\delta^2/2\lambda} \tag{17a}$$

$$f_2(\delta) = \frac{1}{\sqrt{2\pi\beta}} e^{-\delta^2/2\beta} \tag{17b}$$

$$a = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_n^2} \tag{17c}$$

$$\beta = \sigma_x^2 + \sigma_n^2. \tag{17d}$$

The joint distribution $P_{z,y}$ is

$$P_{z,y}(\zeta,\omega) = \frac{1}{2\pi} \left(\lambda\left(\beta - \frac{\lambda}{a^2}\right)\right)^{-1/2} \exp - \frac{1}{2}\left[\zeta^2\left(\frac{a^2\beta - \lambda}{\lambda\beta}\right) - 2\zeta\,\omega\left(\frac{a}{\lambda}\right) + \omega^2\frac{1}{\lambda}\right]. \tag{18}$$

The rate and the error may be evaluated directly.

$$\epsilon = \frac{\sigma_x^2\sigma_n^2}{\sigma_x^2 + \sigma_n^2} + \lambda \tag{19a}$$

$$R = \frac{1}{2}\log\left[\frac{\sigma_x^4}{\left(\sigma_x^2 + \sigma_n^2\right)\lambda}\right]. \tag{19b}$$

The integrals diverge if

$$\lambda > \frac{\sigma_x^4}{\sigma_x^2 + \sigma_n^2}, \tag{20}$$

and for this critical value of λ

$$\epsilon_{max} = \sigma_x^2. \tag{21}$$

This result may be cast into a more suggestive form. We first observe that

$$\frac{\sigma_x^2\sigma_n^2}{\sigma_x^2 + \sigma_n^2}$$

is the error associated with a least mean square approximation of the corrupted signal. The distortion is always greater than this amount, but is less than or equal to σ_x^2. Distortion as a function of rate is given by

$$\epsilon = \sigma_x^2\left(1 - \frac{(1-e^{-2R})}{\left(\frac{\sigma_x^2}{\sigma_n^2} + 1\right)}\right). \tag{22}$$

For error twice the minimum value, the rate is

$$\log\left(\frac{\sigma_x}{\sigma_n}\right).$$

3. Coding Sequences of Signals

Suppose we are to transmit a sequence of Gaussian random variables $\{x_i\}$ with variances $\{\sigma_{x_i}\}$ and corrupted by added Gaussian noise with variance $\{\sigma^2\}$. We fix the total error

$$\epsilon = \sum_i \frac{\sigma_{x_i}^2 \, \sigma_{n_i}^2}{\sigma_{x_i}^2 + \sigma_{n_i}^2} + \lambda_i$$

and minimize the rate

$$R = \frac{1}{2} \sum_i \log\left(\frac{\sigma_{x_i}^4}{\sigma_{x_i}^2 + \sigma_{n_i}^2} \cdot \frac{1}{\lambda_i}\right).$$

This is very similar to the problem described by Kolmogorov.[3] The λ's are chosen as

$$\lambda_i = \inf\left(\theta, \frac{\sigma_{x_i}^4}{\sigma_{x_i}^2 + \sigma_{n_i}^2}\right).$$

The sequence is divided into two subsequences $\{x_{\alpha i}\}$ for which $\lambda_i = \theta$, and $\{x_{\beta i}\}$ for which

$$\lambda_i = \frac{\sigma_{x_i}^4}{\sigma_{x_i}^2 + \sigma_{n_i}^2}.$$

The error is

$$\epsilon = \sum_{\{\alpha_i\}} \theta + \sum_{\{\beta_i\}} \sigma_{x_i}^2,$$

and the rate is

$$R = \frac{1}{2} \sum_{\{a_i\}} \log \frac{\sigma_{x_i}^4}{\sigma_{x_i}^2 + \sigma_{n_i}^2} \cdot \frac{1}{\theta}.$$

4. Rate for a Continuous Source

For a real Gaussian stationary source with spectrum S(f) and real noise N(f), the rate may be found through a procedure similar to that used by Kolmogorov.[4] We define the associated function $\lambda(f)$ given by the relation

$$\lambda(f) = \inf \left(\theta, \frac{S^2(f)}{S(f) + N(f)} \right)$$

and define the set of frequencies for which $\lambda(f) = \theta$ as R_1.

The mean squared error

$$\epsilon = \int_{f \in R_1} df \left(\frac{S(f) N(f)}{S(f) + N(f)} + \theta \right) + \int_{f \notin R_1} df\, S(f),$$

and the rate per unit time is

$$R = \frac{1}{2} \int_{f \in R_1} df \log \left[\frac{S^2(f)}{S(f) + N(f)} \cdot \frac{1}{\theta} \right].$$

Another interesting case arises when the signal is passed through a filter whose frequency response is H(f) before N(f) is added. To reduce this to the problem solved above, we consider transmitting the spectrum S(f) with additive noise $N(f) \cdot |H(f)|^{-2}$. The result is

$$\lambda(f) = \inf \left(\theta, \frac{|H(f)|^2 S^2(f)}{|H(f)|^2 S(f) + N(f)} \right)$$

$$R_1 = \left(f;\ \theta < \frac{|H(f)|^2 S^2(f)}{|H(f)|^2 S(f) + N(f)} \right)$$

$$\epsilon = \int_{f \in R_1} df \left(\frac{|H(f)|^2 S(f) N(f)}{|H(f)|^2 S(f) + N(f)} + \theta \right) + \int_{f \notin R_1} df\, S(f)$$

$$R = \frac{1}{2} \int_{f \in R_1} df \, \log \left(\frac{S^2(f)|H(f)|^2}{S^2(f)|H(f)|^2 + N(f)} \cdot \frac{1}{\theta} \right).$$

This result may also be extended to the case in which the error is weighted with a frequency weighting function.

O. J. Tretiak

References

1. C. E. Shannon and W. Weaver, The Mathematical Theory of Communication (University of Illinois Press, Urbana, Ill., 1964), p. 111.

2. Ibid., p. 114.

3. A. N. Kolmogorov, "On the Shannon Theory of Information Transmission in the Case of Continuous Signals," IRE Trans., Vol. IT-2, No. 4, December 1956, p. 106.

4. Ibid., p. 108.

D. MODEL FOR THERAPEUTIC DIAGNOSIS OF DISEASE

1. Diagnostic Process

The steps taken by a physician to effect the cure of a patient who comes to him with a complaint vary from physician to physician. Varying degrees of intuition or experienced guessing are combined with principles of diagnosis that the physician learned in medical school to arrive at a diagnosis and prescribe a course of treatment. The ability of the physician to look at a patient and say he "looks sick" cannot be duplicated by a machine. A computer can, however, perform a statistical diagnosis if it is programmed in a logical manner to do so, and thus act as a consultant to the physician.

Concentrating on the systematic approach to diagnosis, we find that information about the state of a patient comes from four sources: (i) his medical history, (ii) a physical examination, (iii) laboratory tests, and (iv) careful observation of the course of his illness.

There is a tendency on the part of the physician to collect more information than is necessary to arrive at a probable diagnosis, sometimes with little efficiency in the selection of laboratory tests. If a more systematic approach to test selection could be made, the cost of making the diagnosis could be reduced. Also, the physician is limited in his ability to accurately consider more than a few symptoms and possible diseases, and as a result may be led to an erroneous diagnosis if the disease is uncommon and difficult to diagnose. Moreover, the physician cannot easily assign approximate probabilities or confidences to his final diagnosis. Of itself, this is not of great importance; however, in the face of alternative treatments with risk to the patient, the

physician must once again make an experienced guess, and the chances of its being correct or optimal might be improved by a quantitatively more systematic approach.

2. Modeling the Diagnostic Process

A consideration of the relative likelihood of diseases can be made through the use of Bayes' theorem. This probabilistic approach seems appropriate when we consider that medical knowledge itself is usually nondeterministic. We can determine the relative probabilities of disease complexes D_i, given the observed symptom complexes S_k, by Bayes' theorem.

$$P(D_i/S_k) = \frac{P(S_k/D_i)\, P(D_i)}{\sum\limits_i P(S_k/D_i)\, P(D_i)} . \tag{1}$$

The second term in the numerator is the probability that the patient has disease complex D_i, regardless of any symptom that he might have (i. e. , the a priori probability of D_i). This term varies from location to location, season to season, and is sensitive to the occurrence of epidemics. The denominator is just the scale factor of the unconditional probability of symptom complex S_k.

Bayes' formula also provides a simple numerical means of performing the logical elimination of disease complexes. If disease D_i and symptom S_k are logically inconsistent, then $P(S_k/D_i) = 0$, which from (1) yields $P(D_i/S_k) = 0$.

With the assumption of the availability of the necessary probabilistic data, Bayes' formula is the key to most sequential differential diagnostic procedures. As knowledge about the state of a patient increases through medical history taking, physical examination, laboratory testing, observation of the course of the illness and the reaction to therapy, the probability of each disease complex can be adjusted accordingly. If no disease is indicated with sufficiently high probability, then, if possible, further information should be gathered. On the other hand, the high probability of one disease may signal the time to initiate treatment, and by using the probabilities the risk involved can be ascertained should the chosen therapy have contraindications for other diseases with lesser likelihoods. Should the risk be too great or the diagnosis still uncertain, the current disease probabilities help provide an indication of what information should be sought (e. g. , what test to take) to move efficiently and with least expense toward a more conclusive diagnosis.

Previous studies have tended to consider diagnosis in a context somewhat divorced from the realities of the medical health care profession. Few have touched on the end purpose of diagnosis — to resolve uncertainty about the true state of the patient so that appropriate treatment can be chosen. The implication in considering the treatment — and not just a static diagnosis of the patient — is that the goal of the process has shifted

from that of ascertaining the probabilities of diseases to that of effecting the patient's return to health.

Diagnosis-treatment is always subject to certain real-life constraints: testing costs, treatment costs, risk to the patient, and so forth. Although the issue of cost in money might be dismissed as insignificant because most of the population subscribes to some form of medical insurance, the costs are nevertheless real if we consider not cost to the patient but the cost of inefficient use of the increasingly scarce resources of the physician's time, facilities for treatment, and so on. It is this problem that I wish to consider. The rest of this report presents a model (admittedly only an approximation) of the therapeutic diagnosis of a patient who has undergone a transition to a state of illness. I shall be especially concerned with exploring and making explicit those difficulties or alternatives that may exist in the implementation of the procedure.

3. Model of Therapeutic Diagnosis

We can usually characterize a patient who visits his physician or is admitted to a hospital for emergency treatment as someone who has undergone a transition from a state of health (state 1) to a state of illness (states 2, 3, ..., n), where the gravity of the state increases with its number. In reality these states form a continuum of patient conditions, but in order to gain the simplifications of a discrete model I shall consider 5 states of health: healthy, moderately ill, very ill, critically ill, and dead. The illustration below shows the possible transitions that a patient can have after an initial transition to the very ill state.
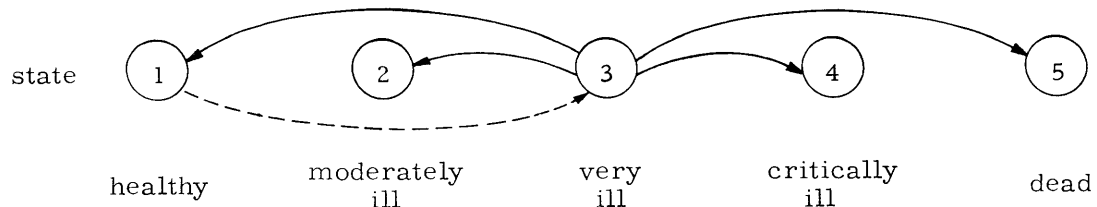
The criteria for classifying a patient into one of these 5 states might be as follows.

State 1   Healthy or cured.

State 2   Probably would recover from his illness in time without treatment.

State 3   Condition is as likely to improve as to worsen if not treated; or patient will be permanently handicapped.

State 4   Probably would die in time without treatment.

State 5   Dead.

Should it be the case that these rough classifications presuppose a tentative diagnosis in order to make a classification, then we might use some other criterion such as the degree of patient discomfort.

Assume for the time being that the patient is in the ill state i (i ⩾ 2). What should the physician's course of action be? If P(ij/A) is the probability of the patient's going from state i to state j as a result of the physician's action A, then he might want to choose an A in order to

    (i)   Maximize P(il)   (optimistic).

    (ii) Minimize P(ij), where j ⩾ i; that is, minimize the possibility of the patient's condition worsening   (pessimistic).

    (iii) Maximize P(il) − f(cost of A)   (more realistic).

    (iv) Optimize some more complex decision function.

The decision criterion used in this model will be developed for an area of medical specialization with the following set of properties.

$\{D_m\}$ $m = 1, \ldots, d$ the set of all disease classifications

$\{S_n\}$ $n = 1, \ldots, s$ the set of all symptom complexes

$\{T_q\}$ $q = 1, \ldots, t$ the set of all relevant diagnostic tests

$\{R_x\}$ $x = 1, \ldots, r$ the set of all treatments useful in treating the above-mentioned set of diseases.

Let us assume that the patient has visited the doctor who has elicited, by history taking, from a computer questionnaire, or by physical examination, some initial information about the patient. This information can be put in the form of a symptom complex S_n. From this information the doctor can formulate a probabilistic diagnosis, by using Bayes' theorem

$$P(D_m/S_n) = \frac{P(S_n/D_m)\,P(D_m)}{\sum_m P(S_n/D_m)\,P(D_m)}, \tag{1}$$

where

$P(D_m)$ = a priori probability of the occurrence of disease D_m
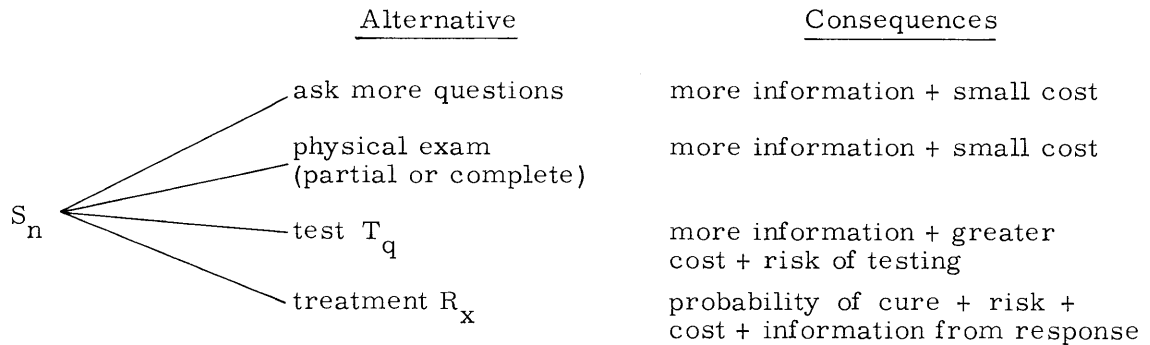
$P(S_n/D_m)$ = probability that a random patient with disease D_m will exhibit symptom complex S_n.

If the probability of one disease stands out from the rest (e.g., Condition: sore elbow — Diagnosis: sprained elbow with probability .8) the physician may wish to proceed

directly to therapy (e. g. , immobilize the elbow). Depending on the risks involved and the available alternatives, however, he may wish to proceed further before treatment. An x-ray photograph (which involves expense) may show a small fracture of the ulna. This gives a new symptom complex and alters the probabilistic diagnosis (in this case we might find P(fracture) = . 99). Although this is a very simple case history, the question that is being raised is when should the gathering of information stop and treatment begin.

Schematically we may characterize the alternatives by a decision node.

Alternative	Consequences
ask more questions	more information + small cost
physical exam (partial or complete)	more information + small cost
test T_q	more information + greater cost + risk of testing
treatment R_x	probability of cure + risk + cost + information from response

The node is labeled S_n.

To simplify the notation, we assume that the set $\{T_q\}$ includes the medical history and physical examination, and that the set $\{R_x\}$ includes the null treatment R_0 (no further treatment), as well as treatments in the general sense of the word (e. g. , wait N days). More than one test may be in order before treatment, or an unsuccessful treatment may be followed by a sequence of tests and a new treatment. This gives rise to a decision tree as a model of the diagnosis-treatment process (Fig. XIV-2). The branch that is
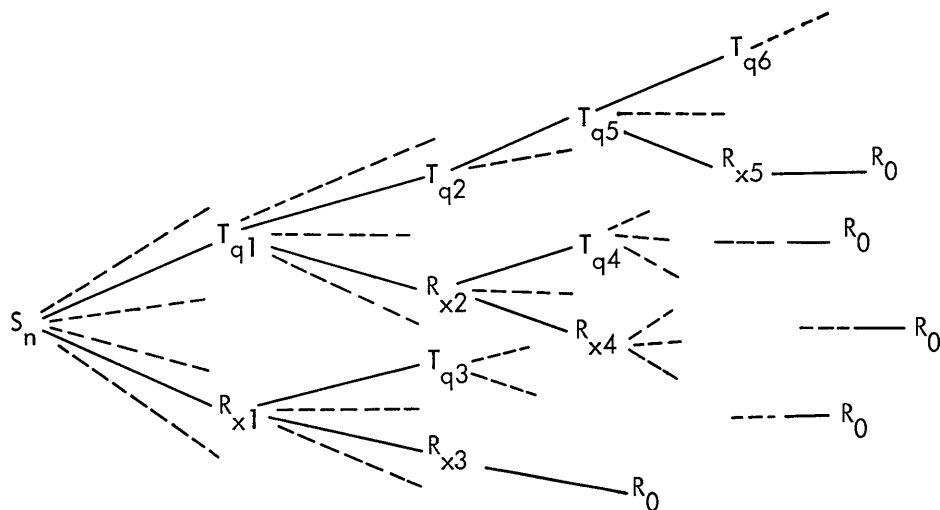

Fig. XIV-2. Decision structure for diagnosis-treatment.

followed until it eventually reaches a terminal node R_o is dependent upon the outcomes of the decisions made at all previous nodes, and probabilistically upon the true state of the patient. We would like to rank order all terminal points according to the decision criteria, and try to choose at each decision node that branch which leads to the best terminal node or nodes. Any such look-ahead procedure, however, creates the problem of computational complexity, because of a very large number of terminal points. Thus, for practical reasons, we are forced to limit look-ahead and may wish to limit consideration of only a few of the more promising branches at any given node.

A decision about how to limit the evaluation of the possibilities implied by the decision tree might be made by considering the question, Should the physician continue testing and thereby obtain a more probable diagnosis, or should he forego further testing that is costly and likely to be unnecessary and initiate treatment? In terms of the model, how does the expected outcome, given that he administers the optimal treatment R_x without further testing, compare with the expected outcome of performing the most relevant test T_q and then administering the optimal treatment based on the probabilistic outcome of that test. Thus a decision based on this limited-depth look-ahead procedure is made by evaluating an inequality of the form

$$E(T_q \text{ followed by } R_x) \gtrless E(R_x \text{ without testing}). \tag{2}$$

The expected outcome will be a function of utility changes for the possible state transitions, monetary costs of treatment and tests, and risk involved.

Before stating the decision function we define the parameters:

u_{ij} = utility of the transition from state i to state j

CR_x = dollar cost of treatment R_x

CT_q = dollar cost of performing test T_q,

and the probabilities:

$P(T_q, k)$ = probability that test T_q has outcome $k = \sum\limits_m P(T_q, k|D_m) P(D_m)$

$P(ij/R_x, D_m)$ = probability that a patient with disease D_m will go from state i to state j if treated with therapy R_x

$P(ij/T_q, D_m)$ = probability that a patient with disease D_m will go from state i to state j if administered test T_q

R^{k_q} = given the observed state of the patient before administering test T_q, R^{k_q} is the optimum treatment if the test has outcome k.

The expected utility of a treatment R_x can be written

$$E_u(R_x) = \sum_i \sum_j P(ij/R_x) \, u_{ij}$$

$$= \sum_i \sum_j \left[\sum_m P(ij/R_x, D_m) \, P(D_m) \right] u_{ij}. \tag{3}$$

The expected risk involved in administering test T_q can be written

$$E_r(T_q) = \sum_i \sum_j P(ij/T_q) \, u_{ij}$$

$$= \sum_i \sum_j \left[\sum_m P(ij/T_q, D_m) \, p(D_m) \right] u_{ij}, \tag{4}$$

where $P(D_m)$ is the probability of disease D_m computed from Bayes' formula.
The inequality

$$\max_q E\left(T_q \text{ followed by } R^{k}{}_q\right) \mathrel{\mathop{\gtreqless}\limits_{\substack{\text{test} \\ \text{treat}}}} \max_x E(R_x \text{ without further tests}) \tag{5}$$

can be evaluated as follows.

$$E(R_x \text{ without further tests}) = E_u(R_x) - CR_x$$

$$= \sum_i \sum_j \left[\sum_m P(ij/R_x, D_m) \, P(D_m) \right] u_{ij} - CR_x \tag{6}$$

$$E\left(T_q \text{ followed by } R^{k}{}_q\right) = \sum_k P(T_q, k)\left[E_u\left(R^{k}{}_q\right) - CR^{k}{}_q \right] + E_r(T_q) - CT_q$$

$$= \sum_k P(T_q, k)\left[E_u\left(R^{k}{}_q\right) - CR^{k}{}_q \right]$$

$$+ \sum_i \sum_j \left[\sum_m P(ij/T_q, D_m) \, P(D_m) \right] u_{ij} - CT_q. \tag{7}$$

Equation 6 is evaluated for R_x $x = 1, 2, \ldots, r$ and will have a maximum for some $x = x^*$. Thus, R_{x^*} is the optimal treatment if there is to be no further testing. Equation 7 is evaluated for T_q $q = 1, 2, \ldots, t$. Each test T_q has a number of possible outcomes.

If S_n is the symptom complex before the results of T_q are known, then for every outcome k of T_q we obtain a new symptom complex $S_{n(k)}$. For each outcome k we use $S_{n(k)}$ and Bayes' formula (1) to compute a new a posteriori distribution for $P(D_m)$. For each outcome k of T_q this new distribution is used to find that R_x which maximizes Eq. 6. That treatment is labeled R^{k_q}, and the maximum value of (6) for outcome k of T_q is, therefore, $E_u(R^{k_q}) - CR^{k_q}$. This value is then used in evaluating Eq. 7 for each test T_q. For some $q = q^*$, Eq. 7 has a maximum, yielding T_{q^*} as the best test to administer if we decide on further testing before treatment.

If Eq. 6 evaluated for R_{x^*} is less than Eq. 7 evaluated for T_{q^*}, then the model dictates further testing and suggests giving T_{q^*} as the next test. The results of T_{q^*} yields new information upon which the decision to treat or continue testing can again be made. If Eq. 6 evaluated for R_{x^*} is greater than Eq. 7 evaluated for T_{q^*}, then the model dictates initiating treatment R_{x^*}. If the patient is cured the process is terminated. If the patient is not cured, the additional information gained by his response to R_{x^*} is lumped with the previous data S_n and used to decide whether to test or try a new treatment.

I have tried to describe briefly a model of the diagnosis-treatment procedure. There are many obstacles to implementing such a model, even on a small scale. Obviously, estimates must be obtained for several one- and two-dimensional parameters. It is my intention to pursue this model of the diagnosis-treatment process further, in the hope of simulating actual decisions so that its accuracy and implications can be properly determined. In spite of the problem of data collection, parameter estimation, and practical application, the advantages of a well-defined decision procedure might be more efficient and accurate care in the face of an increasing scarcity of medical care resources, some implications for multiphasic screening, and, particularly, a visible and rational context in which to formulate decisions with regard to diagnosis and treatment of disease.

B. E. Boyle

References

1. N. E. Betaque, Jr., "Utility Theory Applied to Medical Diagnosis and Treatment," S. M. Thesis, Department of Electrical Engineering, M. I. T., June 1969.

2. A. S. Ginsberg, et al., "An Application of Decision Theory to a Medical Diagnosis-Treatment Problem," P3786, Rand Corporation, Santa Monica, California, February 1968.

3. G. A. Gorry, "A System for Computer-Aided Diagnosis," Project MAC Technical Report MAC-44 (Thesis), Cambridge, Massachusetts, 1967.

4. R. S. Ledley, "Practical Problems in the Use of Computers in Medical Diagnosis," Proc. IEEE 57, 1900-1918 (1969).

5. D. R. Luce and H. Raiffa, Games and Decisions: Introduction and Critical Survey (John Wiley and Sons, Inc., New York, 1968).

6. H. R. Warner, A. F. Toronto, and L. G. Veasy, "Experience with Bayes' Theorem for Computer Diagnosis of Congenital Heart Disease," Ann. N. Y. Acad. Sci. 115, 558-567 (1964).

E. LEUKOCYTE LOCATING PROCEDURE

Several members of our group are working on the development of a practical auto-mated leukocyte recognition system. The purpose of such a system is to perform a dif-ferential leukocyte (white blood cell) count on a smear prepared from human peripheral blood. An integral part of the system is the ability to rapidly determine whether leuko-cytes are present in a view and if so where. Working in Fortran IV on the PDP-9 com-puter, we have demonstrated the feasibility of a very simple locating procedure.

1. Background

When a prepared blood smear slide is examined under an oil-immersion microscope at 1000X magnification the appearance is similar to that of Fig. XIV-3, except for the loss of color. Many erythrocytes (red blood cells) can be seen surrounding and touching the single, roughly circular leukocyte distinguished by its optically dense nucleus. Leuko-cytes, 7-20 μ in diameter, are normally outnumbered by the erythrocytes 650:1. Several of the types of leukocytes have segmented or polymorphic nuclei as shown in Fig. XIV-3. A third type of cell found in mammalian blood is the blood platelet which is usually smaller than erythrocytes and of an optical density similar to leukocyte nuclei. For further details see Davidsohn and Wells,[1] Young,[2] or Young.[3]
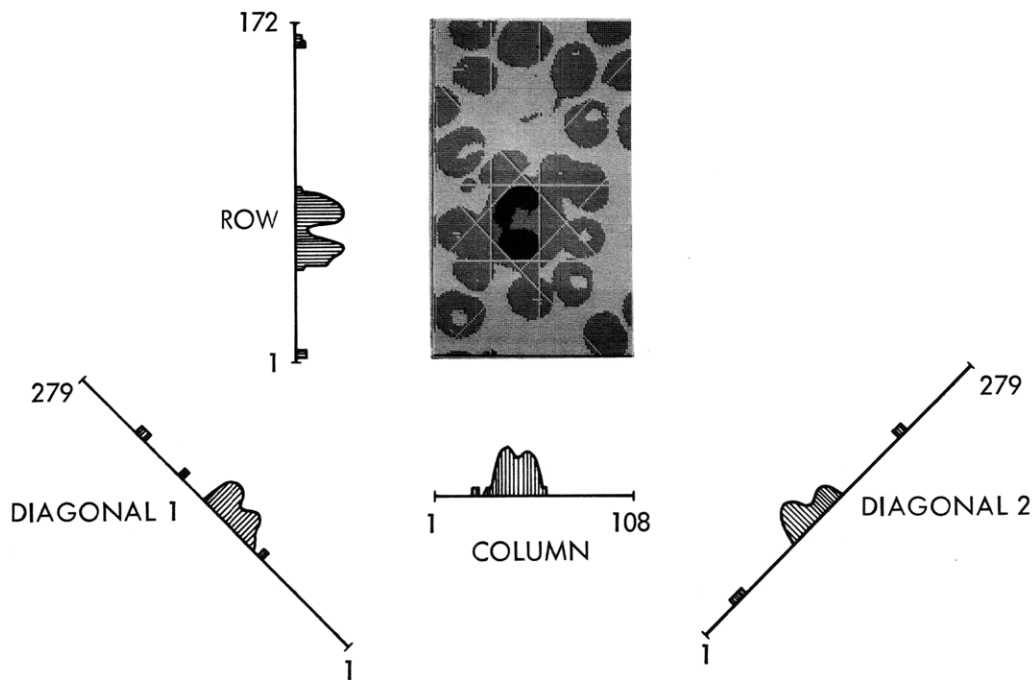

Fig. XIV-3. Typical picture showing its four histograms and the nucleus limits determined from the histograms.

2.   Pictorial Data

The pictorial data used in this study were prepared by Professor Young for his doctoral research.[2-4] A field of view approximately 70 μ × 50 μ containing one leukocyte and the surrounding erythrocytes was photographed. The resulting transparencies were scanned by the on-line flying-spot color scanner SCAD and the output stored on magnetic tape in the form of quantized blue, green, and red optical transmittances at each of 18,576 points (172 row × 108 column matrix) over the picture.

For this study only the green record was used.

3.   Image Features

As can be seen from Fig. XIV-3, the nucleus of the leukocyte is the most distinguishing characteristic, and it alone is used to locate the leukocyte. In order to distinguish between nucleic and non-nucleic material in the picture, a technique is used similar to that by which Young separated background from red and white cell points.
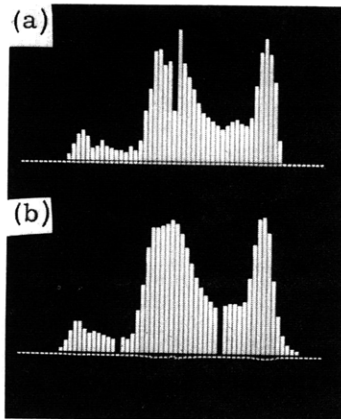

Fig. XIV-4.

(a) Brightness value histogram (raw data).
(b) Smoothed histogram with the nucleic-cytoplasm and cytoplasm-background thresholds shown.

A histogram or probability density function of brightness values is computed from the entire picture (Fig. XIV-4). Three distinct peaks are seen: the first corresponds to a distribution of nucleus brightness values, the second to leukocyte and erythrocyte cytoplasm, and the third to background points. Thus the minimum between the first and second peaks represents a threshold level by which to classify all points of the picture as either nucleic or non-nucleic.

The cytoplasm of the leukocyte provides a buffer zone of non-nucleic points around the nucleus, thereby eliminating the common problem of touching objects.

4.   Location Algorithm

Determining bounds on the nucleus extent is accomplished by the use of spatial histograms (henceforth referred to as histograms, as distinct from the brightness value

histogram). The digital two-dimensional picture is transformed to four one-dimensional
histograms by a simple threshold-summing procedure defined by the following equations:

$$h_r(r) = \sum_{r=1}^{ROWS} T(b_{cr}) \qquad 1 \le r \le ROWS$$

$$h_c(c) = \sum_{c=1}^{COLS} T(b_{cr}) \qquad 1 \le c \le COLS$$

$$h_m(m) = \sum_{c=1}^{COLS} T(b_{c,\,m+c-COLS}) \qquad 1 \le m \le (COLS+ROWS-1)$$

$$h_p(p) = \sum_{c=1}^{COLS} T(b_{c,\,p-c+1}) \qquad 1 \le p \le (COLS+ROWS-1)$$

where

h_r = row histogram

h_c = column histogram

h_m = diagonal 1 histogram

h_p = diagonal 2 histogram

$T(.)$ is a threshold function defined by

$T(X) = 1$ $X \le$ brightness value threshold

$T(X) = 0$ otherwise

$\{b_{cr}\}$ is the array of brightness values (i. e., the picture),

b_{cr} is undefined for c, r not in the range (1, COLS), (1, ROWS).

Although two histograms are sufficient for location, four are, at present, being used
for reasons that will be mentioned. The spatial orientation of these histograms is rep-
resented in Fig. XIV-3. The transformation above requires only one scan of the picture
to examine the brightness value at each point and, when a $b_{cr} \le$ threshold is encoun-
tered, to increment the appropriate element in each histogram by one. Figure XIV-5
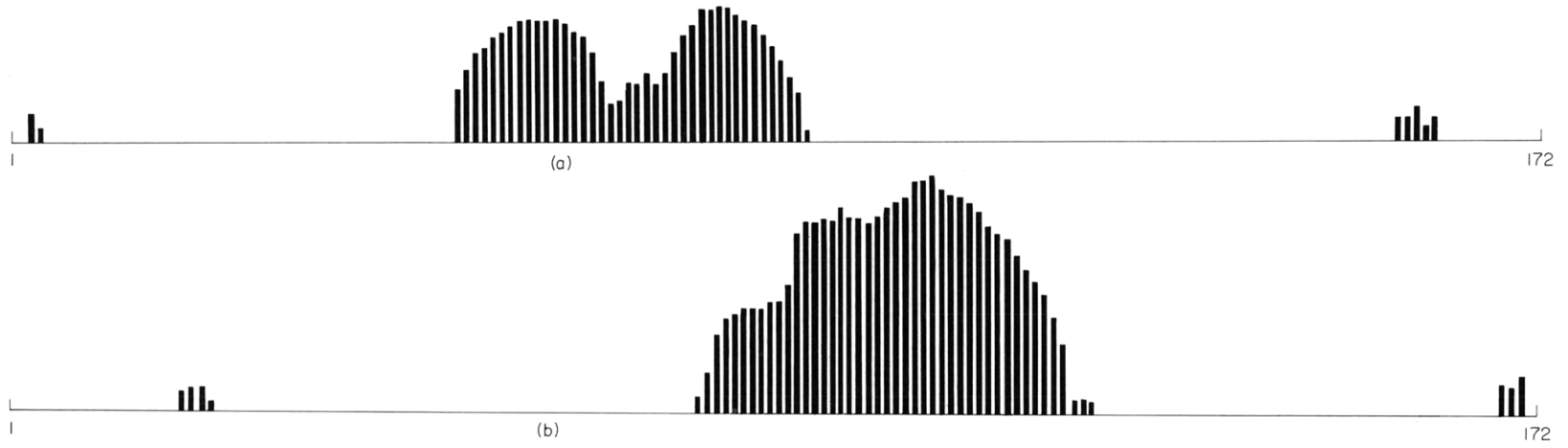shows two typical histograms.

Fig. XIV-5. Typical spatial histograms: (a) from a picture containing a leukocyte with a polymorphic nucleus; (b) from a picture containing a leukocyte with an almost circular nucleus.

All of the 25 pictures tested gave exceptionally noise-free histograms with, in general, no more than two or three extraneous peaks whose areas were much smaller than that of the nucleus.

The nucleus limits are determined from the histograms by a two-step procedure. First, the histograms are simplified by characterizing the information in the form of disjoint peaks. Each peak is represented in terms of four numbers: the two end points, the area, and the maximum value. Noise reduction takes the form of (i) eliminating any peaks not wider than 2 units, and (ii) trimming the ends of the peaks inward to sharpen boundaries and eliminate trailing edges. The second step is to examine these "simplified histograms," which are actually just lists of peaks, looking for possible groupings of peaks that could constitute a nucleus. The grouping operation is governed by the following criteria:

1. The peak having the largest area is taken as the first candidate for inclusion in the nucleus estimate.

2. All remaining peaks with an area greater than 1/4 the area of the largest peak are considered.

3. To prevent grouping two nuclei together, a maximum width parameter and maximum area parameter are checked.

4. If all peaks with area greater than 1/4 of the largest cannot be considered as one nucleus, a second nucleus is postulated and the position of the largest peak that caused the violation is stored.

The nucleus limits obtained from the histograms are checked for consistency with two tests: (i) the area under each histogram between the proposed limits should be approximately equal because each is an estimate of the number of nucleus points, and (ii) the limits should define an octagon of area $\geq$ average area from test (i). All test pictures considered thus far contained only one leukocyte. Multiple cell location has not yet been attempted.

In all of the 25 test cases, the nucleus was located. The octagonal enclosure was never in error by more than 1 or 2 units inside or 3 units outside the nucleus boundary. The usual diameter of a nucleus is 20-40 units.

5. Summary

The important advantage of this method is that only the histograms need be in core to access the presence and location of any leukocytes. Then when a leukocyte is located the pertinent part of the picture can be brought into core.

With normal blood samples, a view usually contains at most one leukocyte. Under these conditions, only two histograms are necessary to do location. Four histograms are generated from each picture, however, to obtain a better estimate of the nucleus

area that is useful in leukocyte classification and to facilitate separation of leukocyte groupings.

As we have mentioned, several types of leukocytes have a polymorphic or segmented nucleus. This polymorphism can vary from a figure eight shape to four or more isolated segments. Much of the polymorphism of a particular nucleus is evidenced in the multi-lobed character of its histograms. A study is now under way to determine whether reliable estimates of nucleus polymorphism can be made from the four histograms.

<div align="right">T. R. Bourk</div>

References

1. I. Davidsohn and B. Wells, Clinical Diagnosis by Laboratory Methods (W. B. Saunders Company, London, 13th edition, 1965).

2. I. T. Young, "Automated Leukocyte Recognition," Quarterly Progress Report No. 89, Research Laboratory of Electronics, M.I.T., April 15, 1968, pp. 231-242.

3. I. T. Young, "Automated Leukocyte Recognition," Ph.D. Thesis, M.I.T., June 1969.

4. I. T. Young, "Automated Leukocyte Recognition," Proceedings of the 8th International Conference on Medical and Biological Engineering and the 22nd Annual Conference on Engineering in Medicine and Biology, Chicago, Illinois, 1969.

F. DISCRIMINATION ALGORITHM FOR THE NORMAL AND SUSPICIOUS CELLS OF THE PAPANICOLAOU SMEAR

In our work of developing a discrimination algorithm, the data with which we are concerned are negative photomicrographs taken with a Leitz-Wetzlar Leica camera using Kodak Panatomic-X black and white film through a Leitz-Wetzlar ORTHOLUX microscope with 1000X magnification.

The transparencies prepared from the photomicrographs have been scanned with the "new scanner." It is a black and white high-resolution flying-spot scanner. The scanned point density is approximately 4 points per micron. Several features have been investigated. Thus far, only three seemed to be of value: (i) integrated nuclear chromatin density, (ii) nuclear cytoplasmic ratio and (iii) nuclear size.

Integrated nuclear chromatin density seems to be the most promising of the three features. The first set of test data comprised 18 cells (see Fig. XIV-6), of which 14 were suspicious and 4 were normal cells. Discrimination based on the nuclear-size measurement alone placed 2 normal cells high up among the suspicious ones and 6 suspicious ones low down among the normal ones. The integrated nuclear chromatin density feature placed the two large normal nuclei much lower and raised the six smaller suspicious ones much higher among the suspicious cells. Actually, 100% separation was achieved.

It is possible that small very darkly stained pyknotic normal nuclei will be difficult
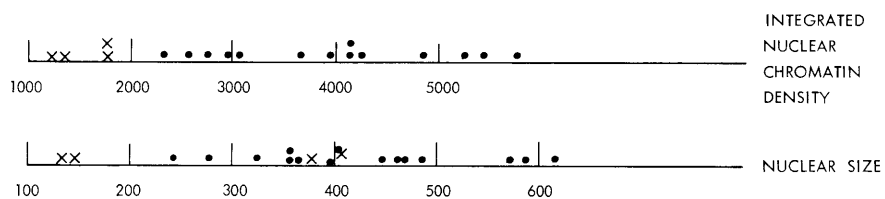
Fig. XIV-6. Test results.
X = normal cells.
● = suspicious cells.

to separate from small suspicious cells just on the basis of their integrated nuclear chromatin density measurements. In this case the nuclear and cytoplasmic measurements will have to be taken into consideration, too.

E. Guttmann