IX. 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. G. R. Granlund
Dr. C. C. Jaffe
Dr. D. M. Ozonoff
Dr. O. J. Tretiak
F. X. Carroll
M. Sharon Hunnicutt
E. R. Jensen
Mary J. Naus
R. A. Piankian

Graduate Students

J. Alba
W. R. Bogan
J. C. Borum
J. E. Bowie
B. E. Boyle
R. M. Boza
Becky J. Clark
P. Coueignoux
C. H. Cox
J. L. Davis
C. J. Drake
J. R. Ellis, Jr.
I. S. Englander
A. M. Fakhr
D. I. Falkenstein
S. G. Finn
T. Fukuda
D. H. Harris
D. W. Hartman
M. E. Jernigan
C. K. S. Kau
J. W. Klovstad
T. T. Kuklinski
S-P. Kuo
C. W. Lynn
H. S. Magnuski
L. C. Makowski
P. L. Miller
B. Minkow
J. A. Myers
D. O'Shaughnessy
J. E. Ostrowski
R. S. Putnam
J. G. Richardson
R. J. Shillman
D. G. Sitler
J. R. Sloan
A. A. Smith
R. D. Solomon
H-m. D. Toong
J. M. Tribolet
B-K. Tye
K. P. Wacks
J. E. Walker
J. A. Whealler, Jr.
A. W. Wiegner
H. M. Wolfson
K. J. Wong

A. LOCALIZATION OF CELLULAR STRUCTURES

National Institutes of Health (Grant 5 PO1 GM14940-07)

I. T. Young, I. L. Paskowitz

1. Introduction

The localization of biological cells and their cellular components in a scanned image system is a problem of considerable interest in automated cytology. In this report we describe a technique that we have devised based on spectral transmission measurements and Boolean picture operations for the rapid identification of white-cell nuclei and white-cell cytoplasm in Wrights-Giemsa stained preparations. While its principal use is in the area of the automated white cell differential, we believe that the technique might have application to other areas of cytologic interest such as Pap smear screening and tissue analysis. In addition, a similar technique has been reported by Stark et al.[1] for use in remote sensing of the Earth's resources.

2. Data Preparation

Blood samples are prepared by placing approximately 1 ml of either fresh or anti-coagulated blood, obtained through venipuncture, on a precleaned glass slide and

spinning it for .75 s at 6000 rpm on a Platt Blood Centrifuge (Model 102M). While this technique of slide preparation is not required to implement the algorithm, it is a standard part of our procedure, since it leads to slides with an evenly distributed monolayer of cells and a minimum of cell disruptions.[2,3] The slide is then automatically stained on an Ames Hema-Tek slide stainer, by using a modified Wrights-Giemsa polychromatic stain. After staining, the slide is placed in our television-microscope system for analysis. This system, diagrammed in Fig. IX-1, has the following components.

Leitz Ortholux microscope with multiple light sources, automated stage, microdensitometer, and multiple objectives/oculars.

High-resolution vidicon camera with 1029 lines per frame and 2:1 interlace. (Granger Associates system and RCA Vidicon Type 8507A).

Analog and digital circuitry to convert the video signal to a low rate (1 sample/horizontal line), quantized (6-bit) data stream.

Tempo-II 16-bit minicomputer assigned to control and process the video data.

The Tempo is connected by a high-speed (1 mbit/s) serial link to a larger computer facility (PDP-9) so that the cell pictures obtained can be processed by our existing software.[4] The algorithms to be described here are, in fact, implemented on the PDP-9 facility. As indicated by the sampling density curve (Fig. IX-2) and the modulation transfer function curve (Fig. IX-3), the resolution of the system is limited only by the resolution of light optics (approximately .25 μm in the visible spectrum). The 37 dB (measured) SNR of the video signal plus certain picture-averaging techniques that we have implemented provide the 6 bits of gray-level resolution.

3. Preliminary Cell Detection

In the initial stage of processing we determine whether white cells are located in the current microscope field of view (100 × 100 μm). This initial processing was devised by Bourk and has been discussed elsewhere[5,6]; only a brief description will be given here. First, the specimen is illuminated with light that is spectrally filtered by an interference filter with a center wavelength of 570 nm and a bandwidth of 10 nm. The entire (100 μm) field of view[2] is then scanned and the histogram of brightness values calculated. A typical histogram for a field containing a white cell is shown in Fig. IX-4a, and for a field without a white cell in Fig. IX-4b. The difference in the histograms at this wavelength is the presence of dark nucleic material. Thus if no appreciable number of dark points occur in the picture the current field is rejected and a new one obtained. If, however, there is dark material in the field (as represented by dark points), a clipping level is generated to separate all points of this darkness (or darker) from all other lighter points in the picture. Two special histograms, the row and column histograms, are then calculated to locate where the dark material in the picture occurred. These positional histograms are generated by looking along any given row (or column) and counting the

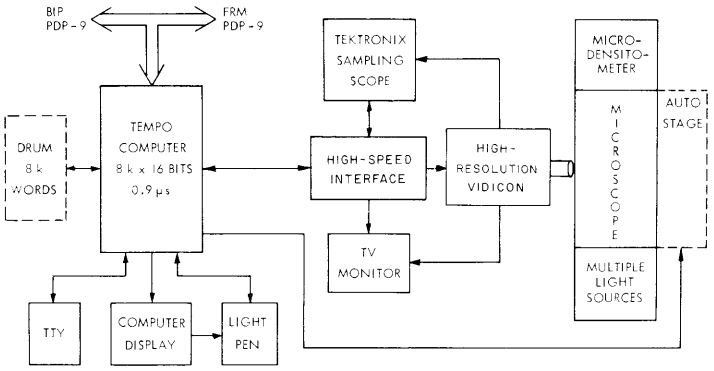

Fig. IX-1.

Television-microscope scanning system.

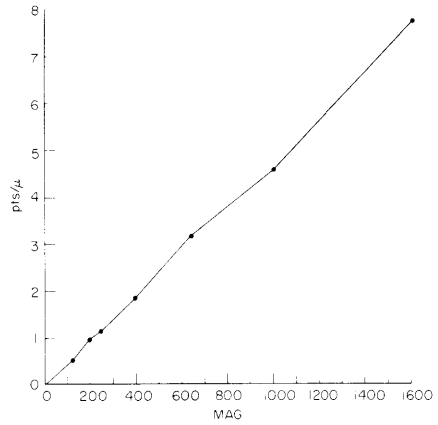

Fig. IX-2. Sample density curve of the television-microscope scanner.

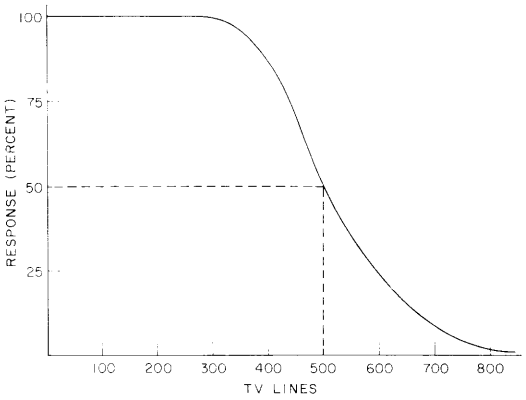

Fig. IX-3. Modulation transfer function of the television-microscope system.

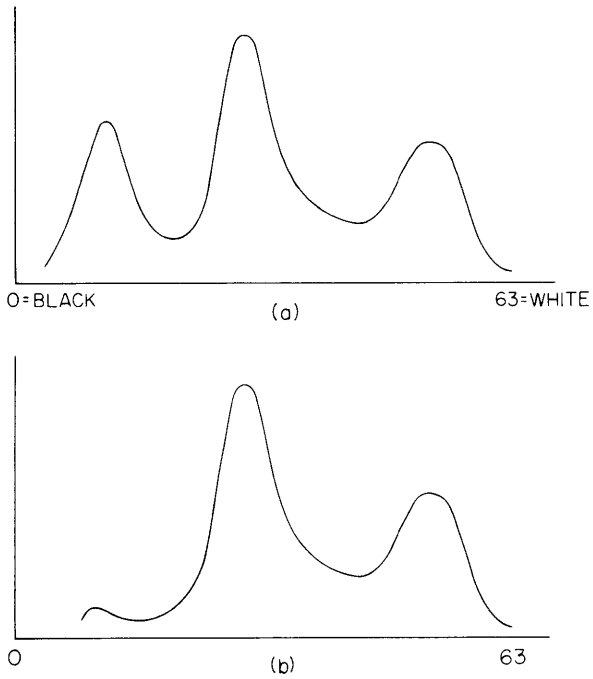

Fig. IX-4.

(a) Histogram of field with white cell.
(b) Histogram of field without white cell.

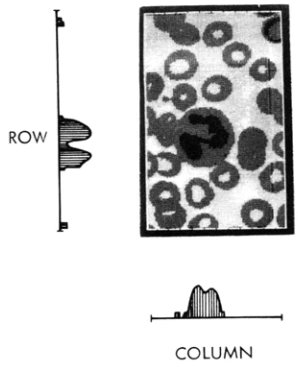

Fig. IX-5. Row column histogram of cell picture.

Fig. IX-6. Typical field of blood smear: located cell.

number of points that are fewer than the clipping level. As seen in Fig. IX-5, it is a simple matter to determine the position of the nucleic material, given these two curves.

The coordinates of the nucleus are specified by a rectangular box (X_{min}, X_{max}, Y_{min}, Y_{max}). We are not only interested in studying the nucleus (and its substructures); we are interested in the cytoplasm and its substructures, too. Therefore we need an algorithm to locate all points in the picture that correspond to cytoplasm, as well as those points that correspond to nucleus. The Wrights-Giemsa staining characteristics of the nucleus cytoplasm and red cells when examined under light of various spectral contacts enable us to secure this information.

4. Algorithm for Component Location

The first stage in processing is to redefine the rectangular box that contains the nucleus:

$$X_{min} \rightarrow X_{min} - \left(\frac{X_{max} - X_{min}}{2}\right) = \frac{3}{2} X_{min} - \frac{1}{2} X_{max}$$

$$X_{max} \rightarrow X_{max} + \left(\frac{X_{max} - X_{min}}{2}\right) = \frac{3}{2} X_{max} - \frac{1}{2} X_{min}$$

$$Y_{min} \rightarrow Y_{min} - \left(\frac{Y_{max} - Y_{min}}{2}\right) = \frac{3}{2} Y_{min} - \frac{1}{2} Y_{max}$$

$$Y_{max} \rightarrow Y_{max} + \left(\frac{Y_{max} - Y_{min}}{2}\right) = \frac{3}{2} Y_{max} - \frac{1}{2} Y_{min}.$$

This effectively doubles the size of the box in both directions while the center of the box is kept in the same location. We have found in practice that doubling the "field" size is sufficient to guarantee the inclusion of all cytoplasmic material for human white blood cells. An example of a typical field is shown in Fig. IX-6.

Next we record the picture and brightness histogram of all points in the new rectangular box for three different spectral illuminations:

1. 570 nm – the "yellow" picture Y(x, y)
2. 530 nm – the "green" picture G(x, y)
3. 420 nm – the "blue" picture B(x, y).

At the sampling density at which this work is done (3 pels/μm) the storage requirements for a typical 25 × 25 μm picture and its histograms are 5827 words of storage in the PDP-9 computer where pels (picture elements) are packed three to the 18-bit computer word.

The characteristics of the three color pictures that are essential to the algorithm are listed in Table IX-1. Our technique translates this linguistic description into a set of Boolean operators on binary pictures.

Table IX-1. Average intensities of different cell structures at different wavelengths.

Illumination	Nucleus	Cytoplasm	Red Cells	Background
570 nm	dark	gray	gray	light
530 nm	dark	dark	dark	light
420 nm	light	light	dark	light

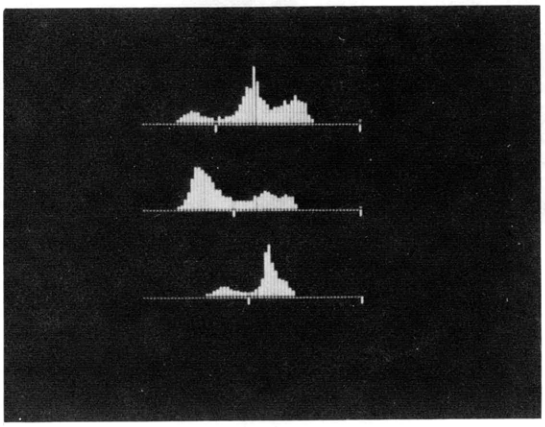

Fig. IX-7. Typical set of histograms with clipping levels.

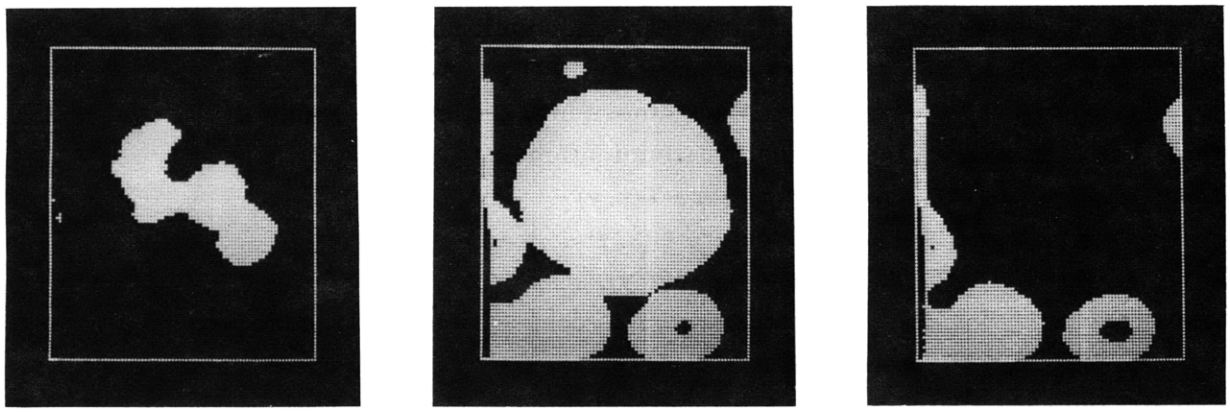

Fig. IX-8. Binary color pictures.

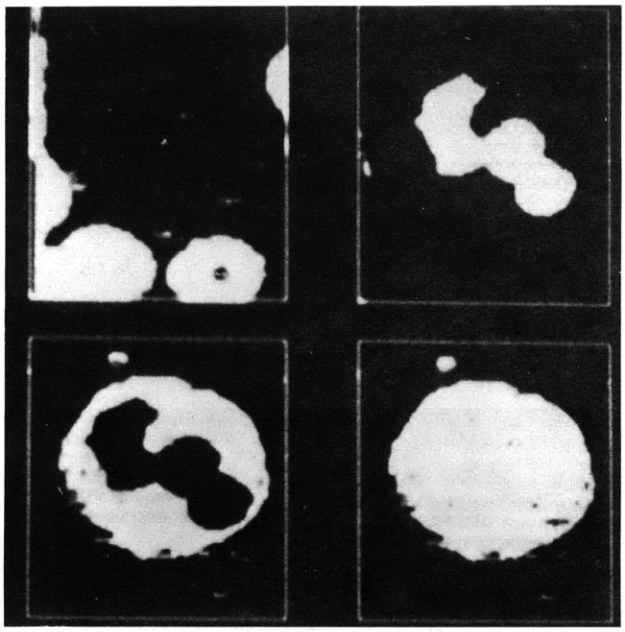

Fig. IX-9. Four cells.

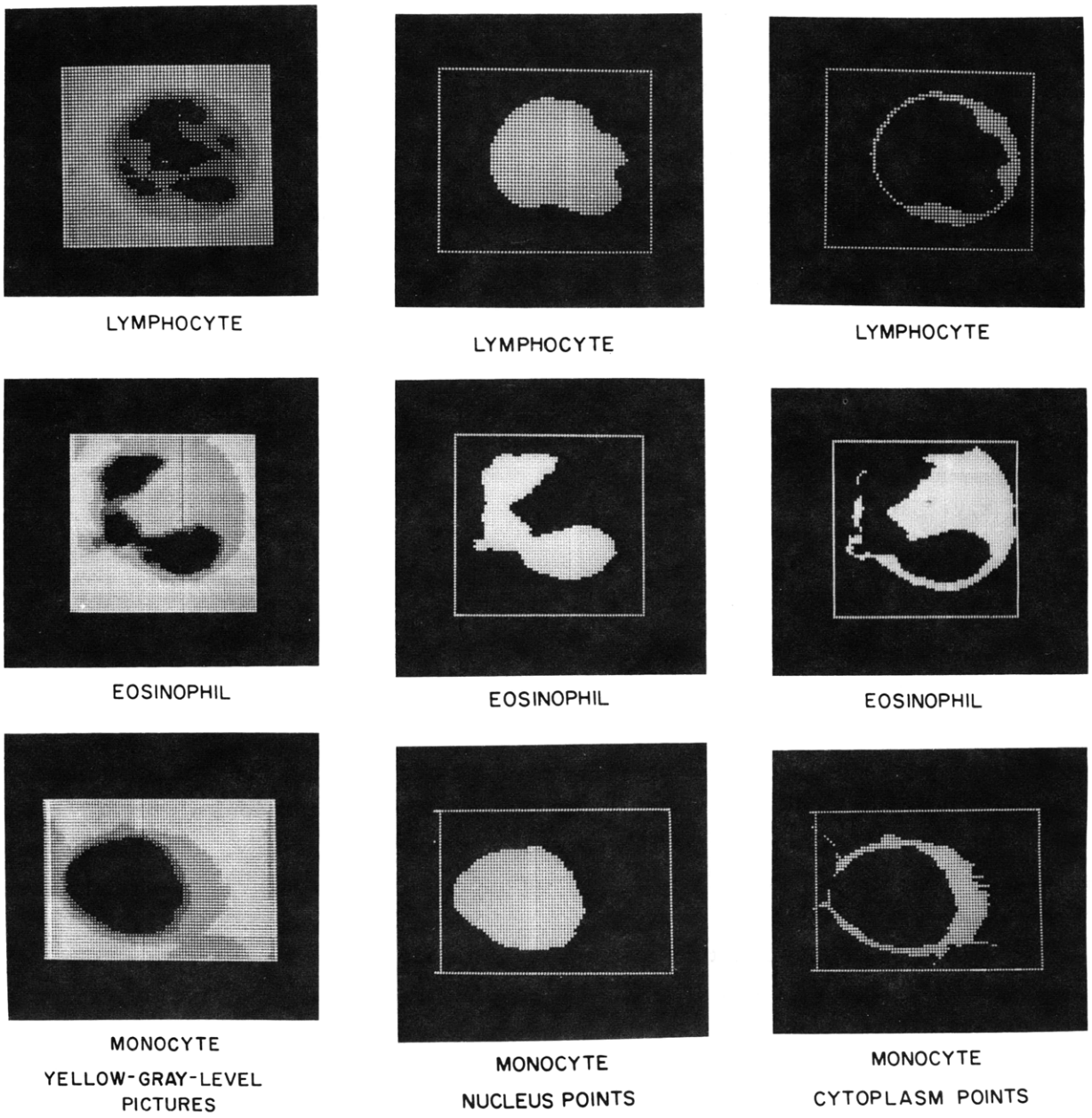

Fig. IX-10. Cells with diverse staining characteristics.

To first translate each 6-bit gray-level color picture into a binary color picture, a clipping level for that picture is extracted from the appropriate histogram. A typical set of histograms for one cell is shown in Fig. IX-7, together with their attendant clipping levels. Thus, for example, in the yellow picture we set

$$Y(x,y) = 1 \quad \text{if} \quad Y(x,y) < Y\text{-CLIP}$$

$$Y(x,y) = 0 \quad \text{if} \quad Y(x,y) > Y\text{-CLIP}$$

and similarly for the green and blue pictures. Three resulting binary color pictures are shown in Fig. IX-8. Finally the three color pictures are combined by the following operators to generate cytoplasm points C(x, y), nucleus points N(x, y), red-cell (erythrocyte) points E(x, y), and total white-cell points W(x, y):

$$N(x,y) = Y(x,y)$$

$$C(x,y) = G(x,y)\,\overline{Y}(x,y)\,\overline{B}(x,y)$$

$$E(x,y) = B(x,y)$$

$$W(x,y) = G(x,y)\,\overline{B}(x,y) + Y(x,y)$$

or

$$W(x,y) = G(x,y)\,\overline{B}(x,y) \qquad \text{(see Table IX-1).}$$

Since the data are packed 3 pels per 18-bit computer word and the logical operators operate on a bit-by-bit basis, it is convenient to define logical "1" to have octal value 77. This permits us to process 3 pels simultaneously, and considerably speeds up the computation. The final set of pictures produced by the algorithm is shown in Fig. IX-9.

To observe how this algorithm operates on cells whose staining characteristics are diverse, we have a set of cells (Fig. IX-10) with (a) the original gray-level (yellow) picture, (b) the nucleus points, and (c) the cytoplasm points.

References

1. H. Stark, R. Barker, and D. Lee, "Some New Techniques for Processing Remotely Obtained Images by Self-Generated Spectral Masks," Appl. Opt. 11, 2540-2550 (1972).
2. M. Ingram and F. Minter, "Semiautomatic Preparation of Coverglass Blood Smears Using a Centrifugal Device," Am. J. Clin. Pathol. 51, 214 (1969).
3. I. T. Young, "The Measurement of Cell Adhesiveness in White Blood Cells," Proc. 24th ACEMB, Las Vegas, Nevada, October 31-November 4, 1971.

4. J. E. Green and O. J. Tretiak, "Modular Picture Processing Package (MP^3)," Quarterly Progress Report No. 94, Research Laboratory of Electronics, M. I. T., July 15, 1969, pp. 261-279.

5. T. Bourk, "Automated Characterization of Leukocyte Nucleus Morphology," S. M. Thesis, Department of Electrical Engineering, M. I. T., 1970.

6. I. T. Young, "The Classification of White Blood Cells," IEEE Trans., Vol. BME-19, No. 4, pp. 291-298, July 1972.