Information Processing & Management Vol. 29, No. 6, pp. 765–774, 1993
Printed in Great Britain.

0306-4573/93 $6.00 + .00

A NEW MACHINE CLASSIFICATION METHOD APPLIED TO HUMAN PERIPHERAL BLOOD LEUKOCYTES

MARK E. RORVIG
Software Technology Branch, PT4, Lyndon B. Johnson Space Center,
National Aeronautics and Space Administration, Houston, TX 77058-3696

STEVEN J. FITZPATRICK
Measurement and Evaluation Center, University of Texas at Austin, Austin, TX 78712-2219

CHARLES T. LADOULIS
Department of Pathology, Mercy General Catholic Hospital, Darby, PA 19023-0000

SANJAY VITTHAL
Reservoir Research and Engineering, Halliburton Services Company,
Duncan, OK 73536-0416

Abstract—Human beings judge images by complex mental processes, whereas computing machines extract features. By reducing scaled human judgments and machine extracted features to a common metric space and fitting them by regression, the judgments of human experts rendered on a sample of images may be imposed on an image population to provide automatic classification.

1. INTRODUCTION

Pattern classification of imagery by computational devices is usually approached in two phases. The first phase is the specification of image exemplars representing the classes by an expert as a training set, with a subsequent classification phase occurring as the joining of image features extracted from the target image population with the features similarly extracted from the specified exemplars (Duda & Hart, 1973). Various difficulties arise with these techniques in both phases. For example, in the training phase, the expert's knowledge must be properly decoded to record accurately the salient features used for exemplar classification, a process of recognized difficulty with many pitfalls (Hayes-Roth *et al.*, 1983). Additionally, in the classification phase, information from the expert must often be encoded as specific programs for identification and matching, thus restricting the applicable domain of the algorithm (Young & Fu, 1986). Even the most robust of these methods, the Fisher linear discriminant (where neither the features of the exemplar nor the domain features of the target population of images need be exactly specified) suffers from the noise introduced in exemplars when the expert makes judgments on only a few features of a multi-featured image.

The method described in this paper, however, requires neither explicit decoding of expert judgments nor domain-specific feature matching. Further, it removes from consideration the noise introduced in the Fisher method. This method, called the Two-Domain Method, introduces two unique processes in both the training and classification phases. First, expert knowledge is acquired through multidimensional scaling (Young & Hamer, 1987) of judgments of dissimilarities rendered by an expert on a sample of images from the target population. Second, general pattern features extracted from images of the target population are transformed to points in a Euclidean space. With this method, the problem of image classification is reduced from the complex one of creating machine-based validity rules to the simple matter of creating a linear mapping between two datasets derived from the human domain and the machine domain, respectively.

2. THE TWO-DOMAIN-METHOD

Consider a collection of images denoted C. Let the goal of the expert be to define pairwise dissimilarities among a sample of these images chosen by a random process. These dissimilarities judgments may be collected by presenting all possible pairs of the images in the sample, and asking the expert to place a mark on a line labeled dissimilar at one end and similar at the other. (A ruler applied to these lines establishes a matrix of dissimilarity values among the sampled images.) By processing these judgments in an n-dimensional space using conventional multidimensional scaling (MDS) techniques, a unique, real-valued ordering of these images by their dissimilarity may be produced. Let this ordering be denoted Φ. With this procedure it becomes unnecessary to know explicitly the portions, features, or aspects of the image, or even the deductive rules used by the expert, in rendering the judgments. Whatever features, aspects, or rules the expert may have attended to or employed are already implicit in the ordering, Φ.

Consider again the collection C. Let it be assumed that each image in this collection has been digitized and processed so as to extract a number of general, primitive features rendered as histograms. (In the application of this paper, six features are extracted: grey levels; edge intensity; edge slope; line length; line distance from the origin; and angle distance from the origin. No claim is made that these features are the only possible features that might be used, or even that these features are optimal. These features are used only because they are very general, convenient ones.) By converting the histograms for each image into Lorenz information measures (Chang & Yang, 1973), and calculating the Euclidean distance among all pairs of images over all feature measures, a matrix, denoted M, of primitive machine image interpretations may be produced. In this manner, the complex problem of image classification is reduced to the far simpler one of creating a linear mapping of Φ on M.

In this method, the mapping is performed by extracting from C the original machine measures matching the subset of C judged by the human expert, calculating Euclidean distances for both machine measurements and human coordinates, deriving weights, β, by multiple regression (where the Euclidean distances from the MDS solution for the human judgments are the dependent variable and the Euclidean distances among images based on machine measurements are the independent variable), and multiplying M by β. By resubmitting the predicted values to the multidimensional scaling process, the final ordering is produced, segregated into classes in an n-dimensional space. Let this last result be denoted Φ'. The complete procedure is displayed as a diagram in Fig. 1, with an example of the complete calculations used in the application below available in Appendix A.

3. AN APPLICATION OF THE TWO-DOMAIN-METHOD TO THE CLASSIFICATION OF TWO POPULATIONS OF HUMAN PERIPHERAL BLOOD LEUKOCYTES

In this article, we have chosen to apply the Two-Domain Method to a problem of discriminating two populations of microscopic images of circulating human white blood cells (leukocytes).

Specifically, the Two-Domain Method was tested for its power to discriminate two distinct patterns of human blood leukocyte distribution: an abnormal pattern associated with acute liver failure exhibiting abnormal circulating white blood cell frequency and distribution (Subject 1), and a normal pattern from a normal, healthy subject (Subject 2).

Circulating human leukocytes were separated by flotation from red blood cells by a standard flotation method, and uniform monolayer films prepared and cytochemically stained by a routine clinical laboratory automated instrument using hematoxylin and eosin dyes. The resulting slides therefore include all nucleated circulating white blood cells, predominantly neutrophils, eosinophils, lymphocytes, and monocytes, as well as platelets.

Eight representative sample fields were selected for each subject. The photographic recording was standardized using one continuous film strip of Ektachrome color reversal film rated at ASA 200. All slides were photographed at the same magnification. Effects of exposure variations and background density were tested in the Two-Domain Method

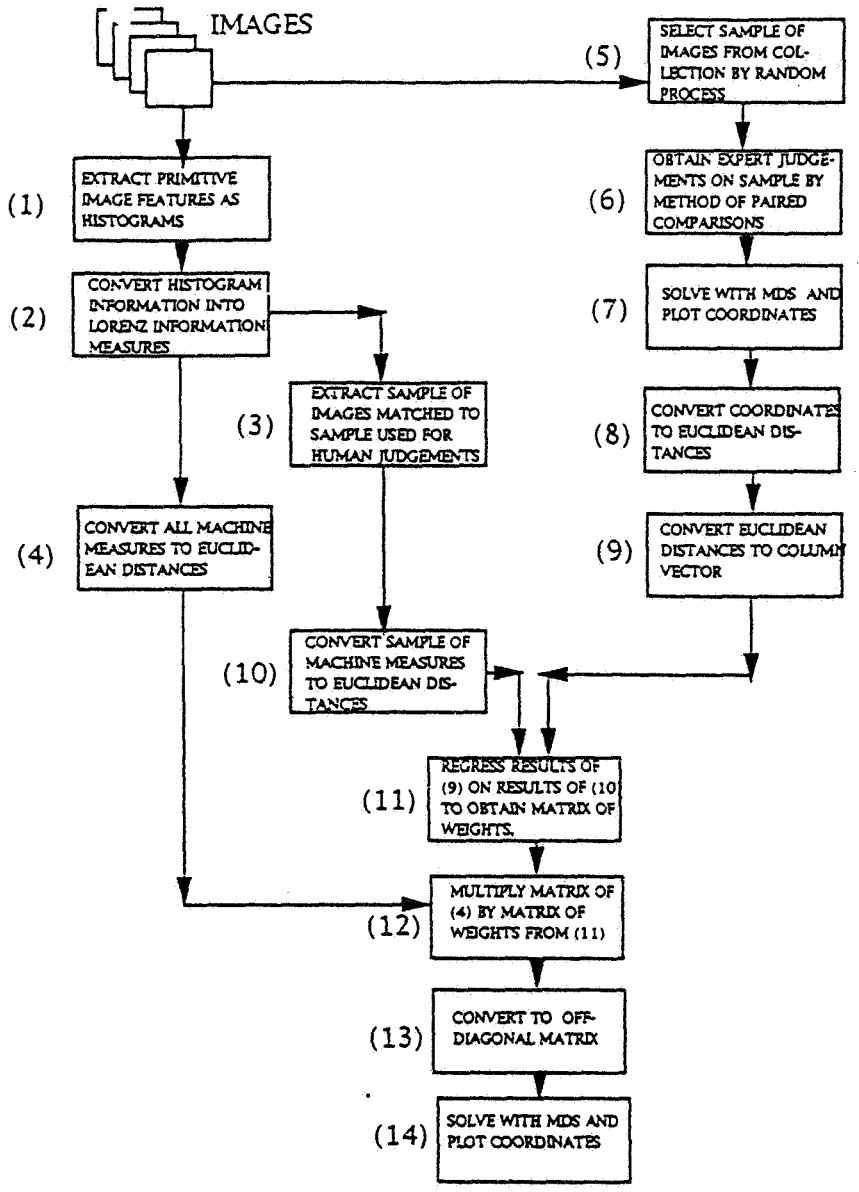


Fig. 1. The procedural steps necessary to execute the Two-Domain Method for any collection of images. Major formulas (exclusive of those used in general MDS (multidimensional scaling) procedures as applied in the application on human peripheral blood may be found in Appendix A.

by recording each image at two different exposures. Set A images (numbered 1–16) were exposed at ASA 200, and Set B images (numbered 17–32) were exposed at ASA 400. Samples used in the test thus consisted of 16 images from each subject, at two levels of exposure, on the same photographic film strip.

The difference in exposure levels substantially alters the machine measurements of these images, and is typical of problems that confound image pattern classification generally, in that "noise" artificially introduced by one element or another distort the machine classification algorithms. Reproductions of both Set A and Set B are presented following the Appendix. The purpose of this application is thus to demonstrate that the Two-Domain Method is sufficiently robust not only to properly classify Set A (by segregation in an n-dimensional space), but also to reduce or eliminate the noise introduced by the difference in Set B film exposure levels.

Expert judgments of dissimilarities were made by an experienced pathologist (C.T.L.), primarily on the basis of the segmentation of leukocyte nuclei, and lymphocyte and monocyte shape and size. Other cell types present in the images were ignored for judgment purposes. Judgments were provided in a single session on slides 1-8 of Set A according to the procedure described in Section 2, and submitted (as are all datasets discussed in this section) to the ALSCAL procedure in SAS, a common multidimensional scaling package.

In Fig. 2, Plots 1 and 2 exhibit a strong separation between the cell populations of the two subjects. The primitive machine interpretations derived from both Set A and Set B, scaled by ALSCAL, appear in Fig. 3 as Plots 3 and 4, respectively.

The images represented by datapoints in Plot 3 appear to have some natural clustering tendency along the same lines as those provided directly by human judgments, proba-

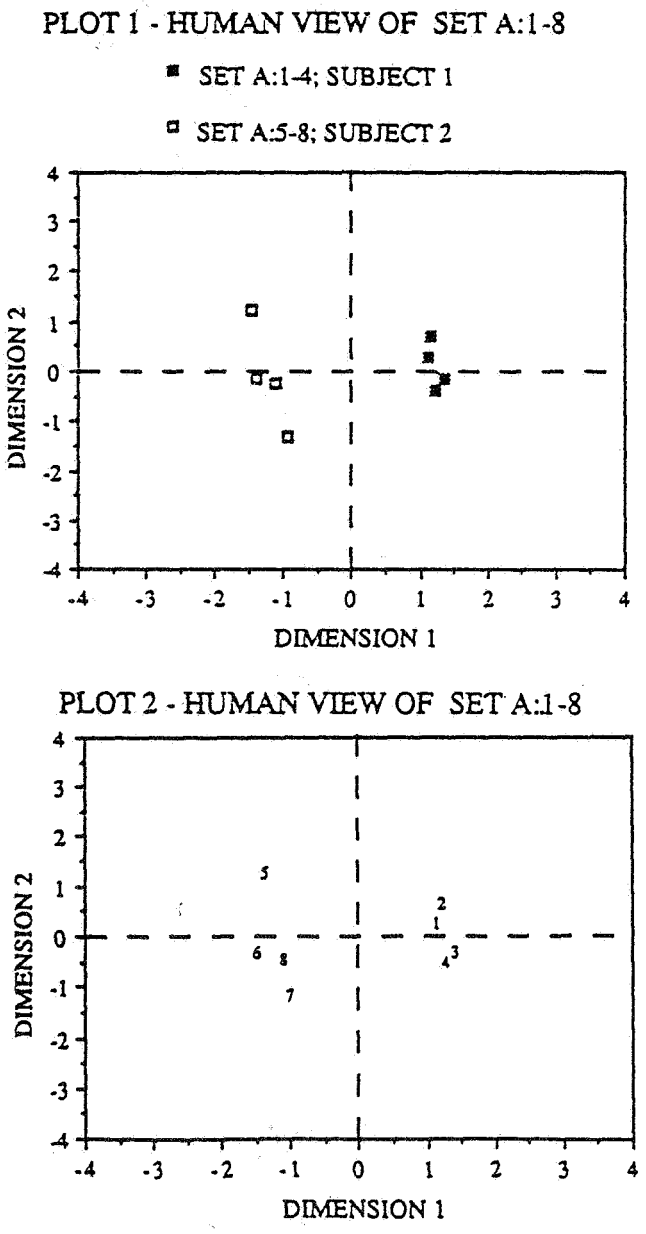


Fig. 2. MDS ALSCAL plots of the original human view of a sample of eight images of peripheral white blood cells. The human judgments were collected through the method of paired comparisons, and show a clear separation between the slides from Subject 1 and Subject 2.

2. THE TWO-DOMAIN-METHOD

Consider a collection of images denoted C. Let the goal of the expert be to define pairwise dissimilarities among a sample of these images chosen by a random process. These dissimilarities judgments may be collected by presenting all possible pairs of the images in the sample, and asking the expert to place a mark on a line labeled dissimilar at one end and similar at the other. (A ruler applied to these lines establishes a matrix of dissimilarity values among the sampled images.) By processing these judgments in an n-dimensional space using conventional multidimensional scaling (MDS) techniques, a unique, real-valued ordering of these images by their dissimilarity may be produced. Let this ordering be denoted Φ. With this procedure it becomes unnecessary to know explicitly the portions, features, or aspects of the image, or even the deductive rules used by the expert, in rendering the judgments. Whatever features, aspects, or rules the expert may have attended to or employed are already implicit in the ordering, Φ.

Consider again the collection C. Let it be assumed that each image in this collection has been digitized and processed so as to extract a number of general, primitive features rendered as histograms. (In the application of this paper, six features are extracted: grey levels; edge intensity; edge slope; line length; line distance from the origin; and angle distance from the origin. No claim is made that these features are the only possible features that might be used, or even that these features are optimal. These features are used only because they are very general, convenient ones.) By converting the histograms for each image into Lorenz information measures (Chang & Yang, 1973), and calculating the Euclidean distance among all pairs of images over all feature measures, a matrix, denoted M, of primitive machine image interpretations may be produced. In this manner, the complex problem of image classification is reduced to the far simpler one of creating a linear mapping of Φ on M.

In this method, the mapping is performed by extracting from C the original machine measures matching the subset of C judged by the human expert, calculating Euclidean distances for both machine measurements and human coordinates, deriving weights, β, by multiple regression (where the Euclidean distances from the MDS solution for the human judgments are the dependent variable and the Euclidean distances among images based on machine measurements are the independent variable), and multiplying M by β. By resubmitting the predicted values to the multidimensional scaling process, the final ordering is produced, segregated into classes in an n-dimensional space. Let this last result be denoted Φ'. The complete procedure is displayed as a diagram in Fig. 1, with an example of the complete calculations used in the application below available in Appendix A.

3. AN APPLICATION OF THE TWO-DOMAIN-METHOD TO THE CLASSIFICATION OF TWO POPULATIONS OF HUMAN PERIPHERAL BLOOD LEUKOCYTES

In this article, we have chosen to apply the Two-Domain Method to a problem of discriminating two populations of microscopic images of circulating human white blood cells (leukocytes).

Specifically, the Two-Domain Method was tested for its power to discriminate two distinct patterns of human blood leukocyte distribution: an abnormal pattern associated with acute liver failure exhibiting abnormal circulating white blood cell frequency and distribution (Subject 1), and a normal pattern from a normal, healthy subject (Subject 2).

Circulating human leukocytes were separated by flotation from red blood cells by a standard flotation method, and uniform monolayer films prepared and cytochemically stained by a routine clinical laboratory automated instrument using hematoxylin and eosin dyes. The resulting slides therefore include all nucleated circulating white blood cells, predominantly neutrophils, eosinophils, lymphocytes, and monocytes, as well as platelets.

Eight representative sample fields were selected for each subject. The photographic recording was standardized using one continuous film strip of Ektachrome color reversal film rated at ASA 200. All slides were photographed at the same magnification. Effects of exposure variations and background density were tested in the Two-Domain Method

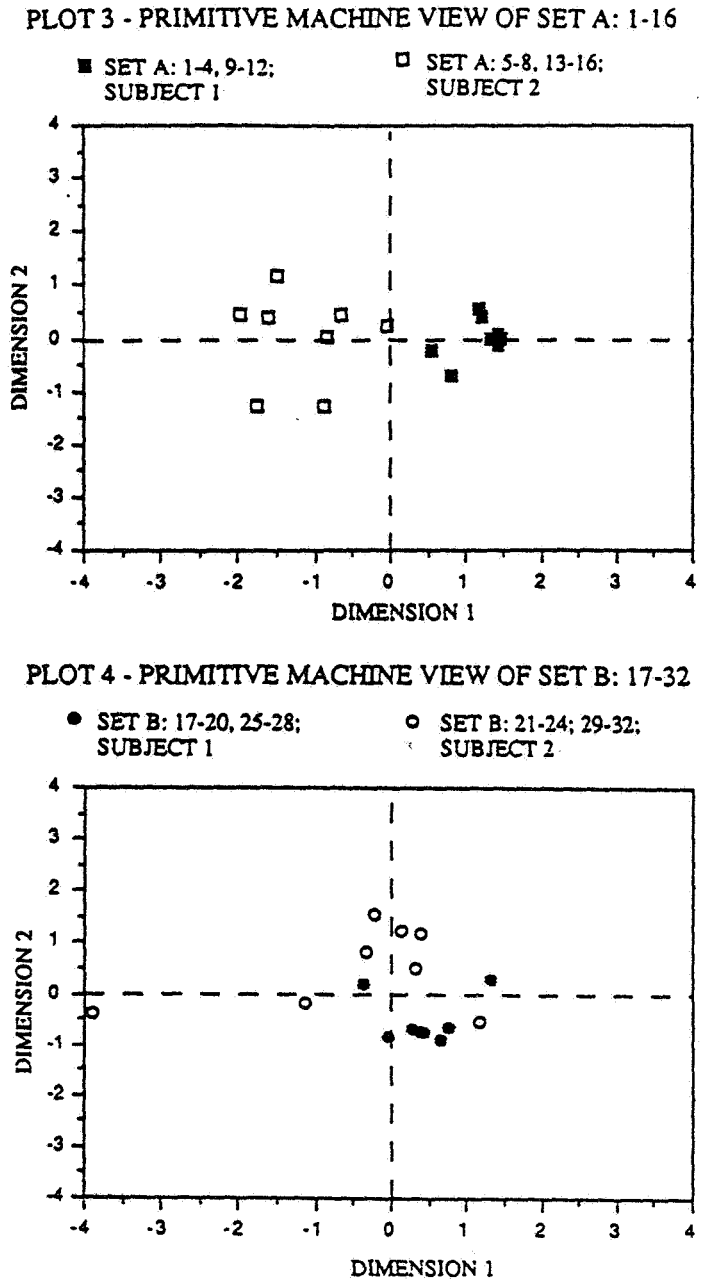


Fig. 3. MDS ALSCAL plots of the primitive machine views of Set A and Set B, including Subject 1 and Subject 2. Plot 3 of Set A (from film rated at ASA 200 and exposed at ASA 200) exhibits some natural clustering by machine features alone, whereas Plot 4 of Set B (from film rated at ASA 200 but exposed at ASA 400) exhibits little machine differentiation between the two subjects.

bly due to the increased light levels in the images produced from Subject 1 and caused by the generally lower levels of white blood cells in the sample drawn from that subject. Plot 4, on the other hand, derived from the deliberately overexposed images, reveals very little meaningful segregation.

In Fig. 4, Plot 5 reveals the strong confounding effect of the Set B data when combined with Set A and scaled together. When the sets are combined, each item acts to influence the scale value of every other item, so that the pure machine view, or interpretation, of these images becomes extremely confused. There is, for example, some segregation of

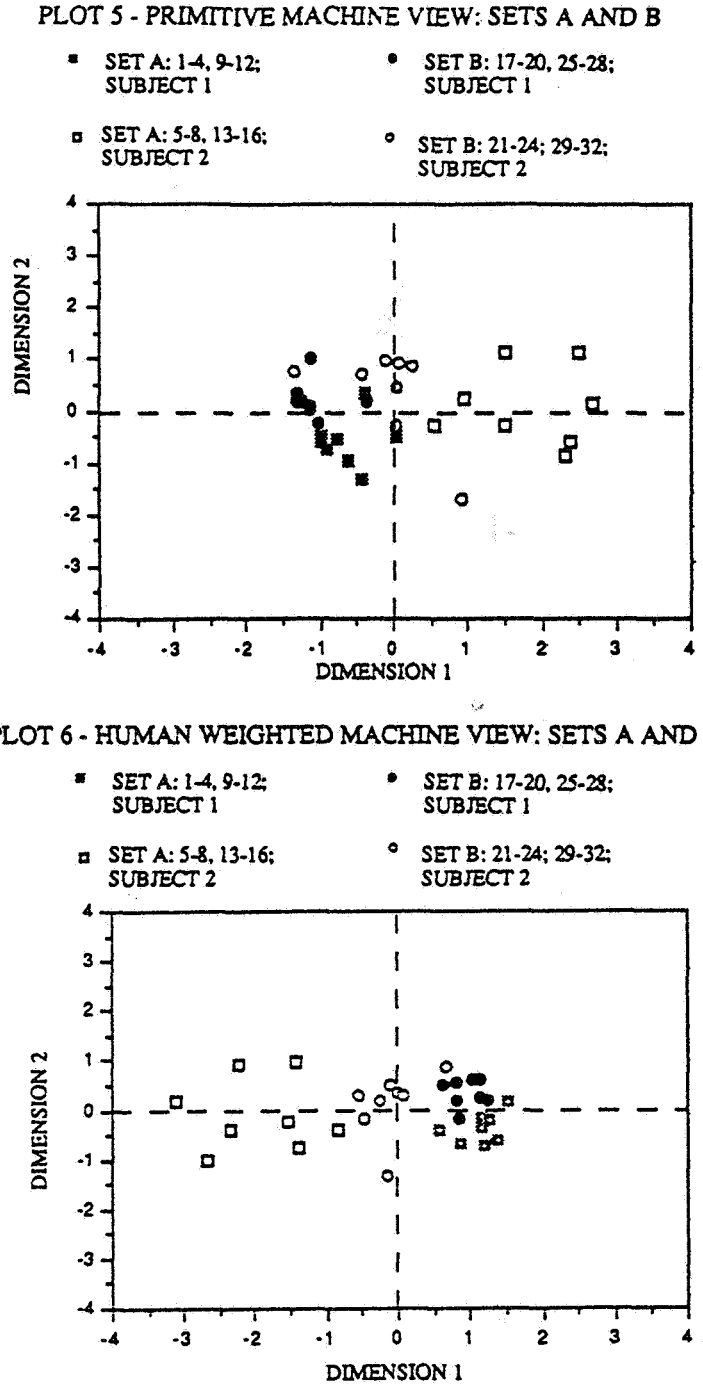


Fig. 4. MDS ALSCAL plots of Sets A and B, Subjects 1 and 2. Plot 5 exhibits distortion of the natural clustering effect displayed in Set A of Plot 3 when Set A and B are combined. Plot 6 exhibits the reordering of Subject 1 and Subject 2 classes when weighted by the human view and displayed in Plot 1. Numbered displays of these datapoints are available in Fig. 7, Plots 7 and 8, in Appendix A.

Subject 1 and Subject 2, but still much less than that appearing in the human classification of these images provided in Plot 1.

Plot 6 exhibits the effect of the Two-Domain Method on the disordered data of Plot 5. Plot 6 was produced according to the procedures of Fig. 1 with the detailed calculations provided in Appendix A. In Plot 6, Subject 1 and Subject 2 data are perfectly segregated

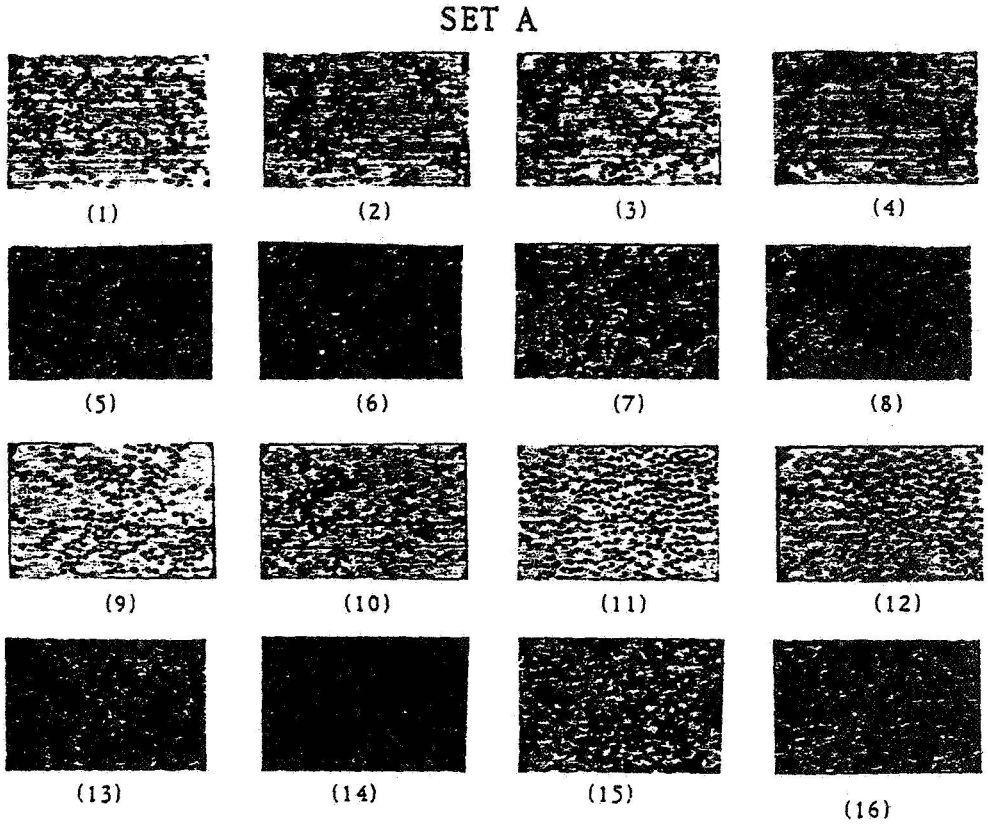


Fig. 5. Reproductions of Set A of 16 slides of peripheral blood cells used in this experiment. All slides were prepared by the same machine-assisted process. Slides 1–4 and 9–12 were extracted from a traumatized subject. Slides 5–8 and 13–16 were extracted from a normal subject. The film used to create the slides was rated ASA 200 and exposed at ASA 200.

for Set A (Fig. 5), and, with the exception of one image, also perfectly segregated for Set B (Fig. 6). Clearly, the strong, confounding effect introduced by combining Set B with Set A images is eliminated.

4. DISCUSSION

The Two-Domain Method, considered very generally, is effective simply because it reduces the intense machine activity associated with image pattern matching to the simple operations of interval scale value relations. Moreover, the scaling theory underlying the method is easily transferable to operations involving classifications among higher dimensions. Indeed, multidimensional scaling has, for some time, been more often used to record human judgments in higher dimensions for a variety of marketing applications (Green & Carmone, 1969). Finally, by using replicated multidimensional scaling methods, the opinions of multiple experts (as opposed to the single expert used in this application) may be combined in the creation of Φ.

The Two-Domain Method is also applicable to image classification systems that routinely use Bayesian methods. In this case, the operations of the Bayesian classifiers would use, as their inputs, the dissimilarity values output from multidimensional scaling matrix transforms, ignoring the plotted values (which are derived from the dissimilarity values anyway.) Along these same lines, the Two-Domain Method may facilitate neural net image classification, both by making the net more efficient due to the reduction of information that must be submitted (dissimilarities or Euclidean distances rather than vectors of pixel values) and by the increased rigor of the training set expression, which reduces noise when aspects of images are judged rather than images as wholes.

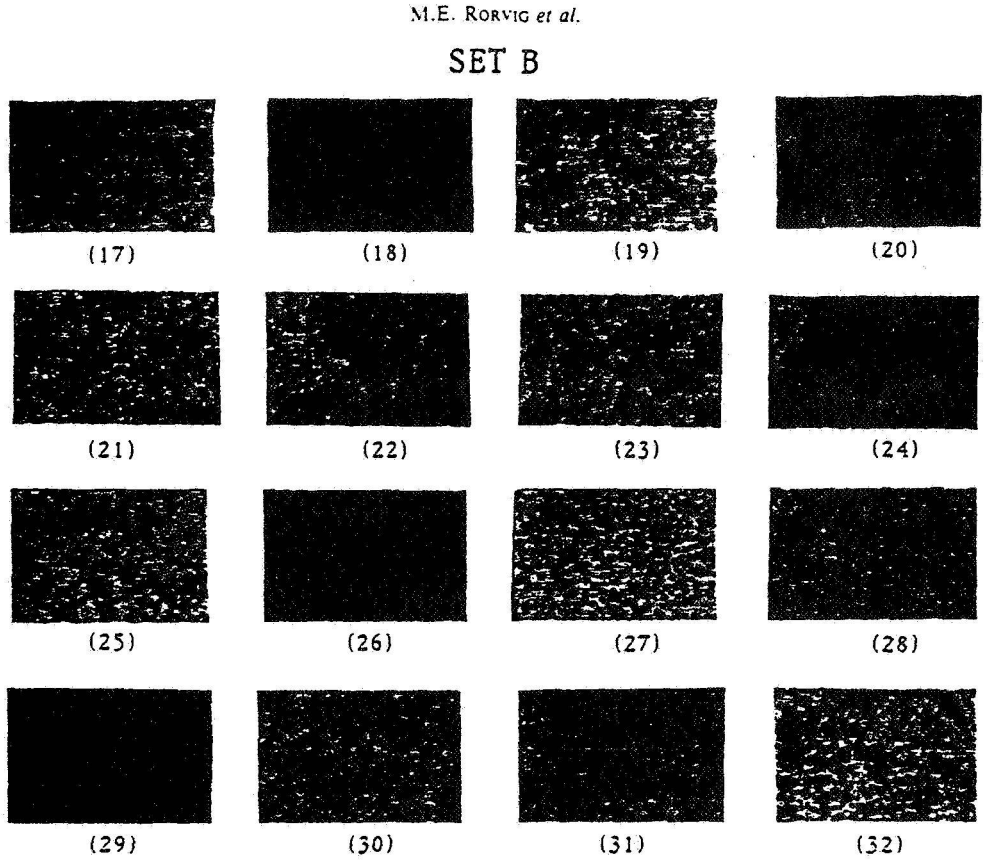


Fig. 6. Reproductions of Set B of 16 slides of peripheral blood cells used in this experiment. All slides were prepared by the same machine-assisted process. Slides 17–20 and 25–28 were extracted from a traumatized subject. Slides 21–24 and 29–32 were extracted from a normal subject. The film used to create the slides was rated ASA 200 and exposed at ASA 400.

Finally, as expressed in some earlier work (Rorvig, 1988), the Two-Domain Method may be used in the searching of large databases of images, where image representations are stored as feature components (Chang & Yang, 1983). In this application, the method would be applied to image classes iteratively, by segregating and mapping successively smaller classes of imagery. This application may be critical in locating desired sets of images that cannot be described linguistically because of either intellectual or economic constraints.

REFERENCES

Chang, S.K., & Yang, C.C. (1983). Picture information measures for similarity retrieval. *Computer Vision, Graphics, and Image Processing, 23*, 366–375.

Duda, R.O., & Hart, P.E. (1973). *Pattern classification and scene analysis.* New York: John Wiley.

Green, P.E., & Carmone, F.J. (1969). Multidimensional scaling: An introduction and comparison of nonmetric unfolding techniques. *Journal of Marketing Research, 6*, 330–341.

Hayes-Roth, F., Waterman, D.A., & Lenat, D.B. (1983). *Building expert systems.* Reading, MA: Addison-Wesley.

Rorvig, M.E. (1988). Psychometric measurement and information retrieval. In M.E. Williams (Ed.), *Annual Review of Information Science and Technology, 23*, 157–189.

Young, F.W., & Hamer, R.M. (1987). *Multidimensional scaling: History, theory, and applications.* Hillsdale, NJ: Lawrence Erlbaum.

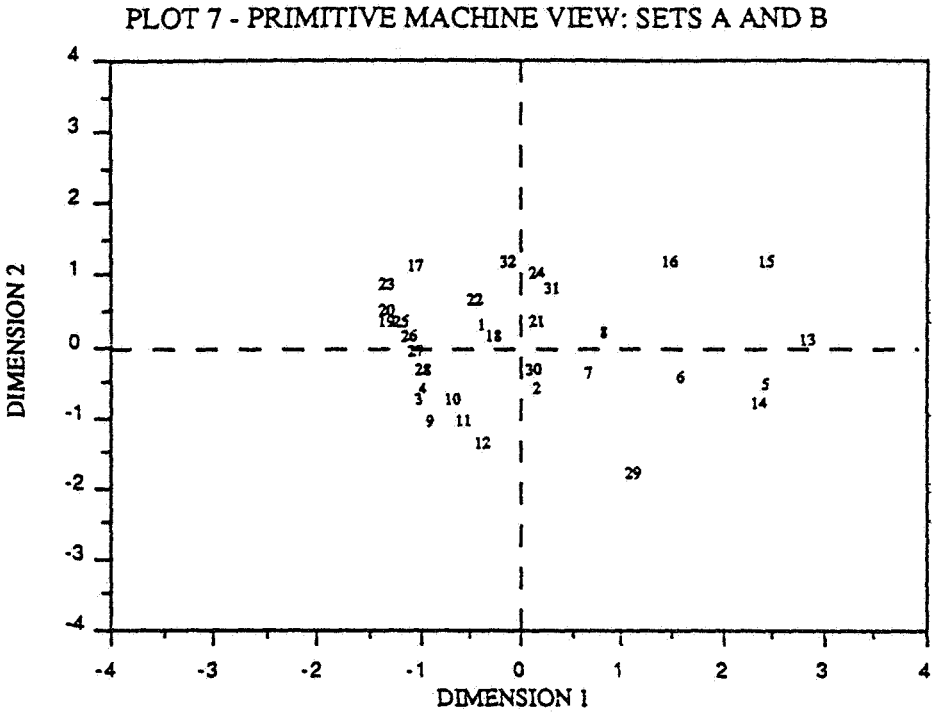
Young, T.Y., & Fu, K.-S. (1986). Part I. Pattern recognition. In *Handbook of pattern recognition and image processing* (pp. 3–167). Orlando, FL: Academic Press.

APPENDIX A

The details of Plot 6 production follow. First, the primitive machine measurements (Lorenz information measures (Chang & Yang, pp. 369–370)) for images 17–24 corresponding to the human

judgments rendered on Set A for images 1–8 were converted to six sets of squared Euclidean distances (one for each machine measurement) according to the following equation:

$$Q_k = (p_{ik} - p_{jk})^2; \quad i < j, \; k = 1,6 \tag{1}$$



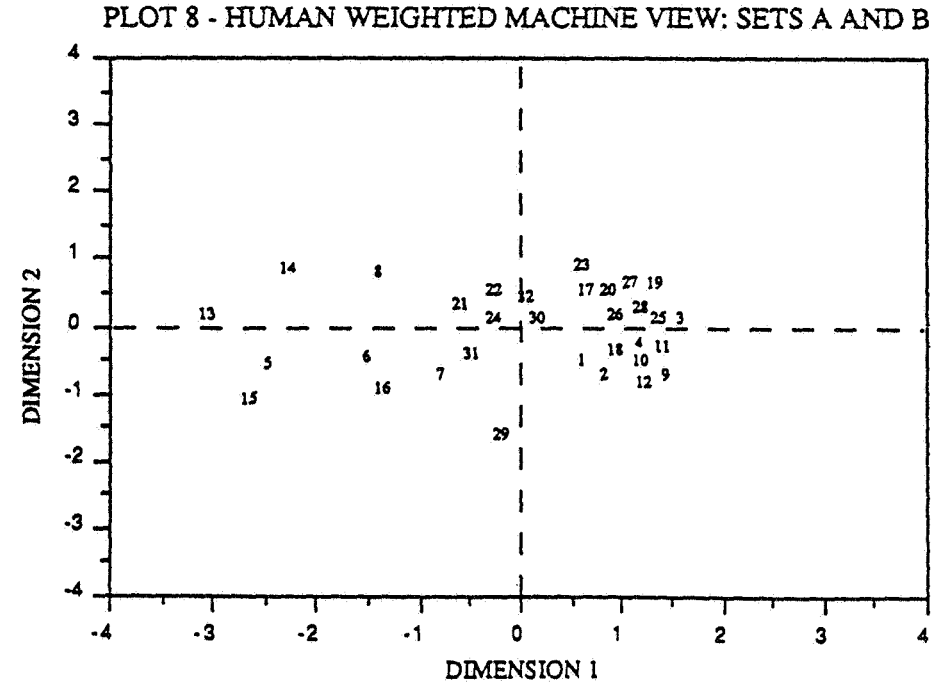


Fig. 7. MDS ALSCAL plots (in numbered display) of both primitive and human weighted views of 32 peripheral blood cell slides. The lower plot exhibits the substantial "learning" effect created by imposition of human judgments on machine interpretations.

where

Q = a matrix of 28 × 6,
Q_k = a column of matrix Q,
p = a matrix of 8 × 6,
p_{ik} = the machine measurement k for image i, and
p_{jk} = the machine measurement k for image j.

Since a column of Q contains the squared difference between all pairs of images on the corresponding machine measurements, there are $[n(n-1)]/2$ elements in each column, where n is the number of images.

Second, the squared Euclidean distances between all pairs of slides 1–8 of Set A, that is, Φ, were computed from the spatial coordinates of the MDS solution for the human judgments of Plot 1 according to eqn. 2:

$$D = \sum_{k}^{r} (x_{ik} - x_{jk})^2, \quad i < j,\ k = 1, r \tag{2}$$

where

D = the square symmetric matrix,
x_{ik} = the coordinate of image i on dimension k,
x_{jk} = the coordinate of image j on dimension k, and
r = the number of dimensions in the solution.

Third, the square symmetric matrix was converted to a column vector containing the top off-diagonal elements (for convenience also denoted D) and regressed on the matrix Q of eqn. 1 to produce the vector of weights, β. Equation 3 is the multiple regression equation in standard form and eqn. 4 is the standard least squares solution.

$$D = Q\beta' \tag{3}$$

$$\beta = (Q'Q)^{-1}Q'D \tag{4}$$

Fourth, the procedure of eqn. 1 was applied to all machine data, images 1–32, denoted M, and multiplied by the vector of weights, β, or

$$\mathbf{V} = M\beta' \tag{5}$$

where

$\mathbf{V}$ = the final vector converted to an off-diagonal matrix for submission to MDS, and
M = the 496 × 6 matrix from the procedure of eqn. 1.

$\mathbf{V}$, submitted to MDS and scaled, thus results in Φ' as displayed in Plot 6.