Digital indicators for red cell disorder

Anjan Kr. Dasgupta* and Prabir Lahiri

Department of Biochemistry and Biophysics, University of Kalyani, Kalyani 741 235, India

Digital light microscopic images of red blood cells (RBC) show a distinctive histogram. In contrast, the image histograms of red cells from patients of thalassaemia and iron deficiency anemia (IDA) show closely placed and ambiguous boundary between the object and the background (near white) peaks. The RBC image contains a centrally illuminated zone that originates from the biconcavity of the erythrocyte surface. The signature contained in the white peak of the histogram is likely to be associated with the characteristic light transmission through this zone. A learning model based on critical linear separability is presented. The model shows that when judged from biconcavity alone, thalassaemic cells are different from both normal and IDA cells. A related finding was that the fractal dimension of the inner contour of the thalassaemic RBC cells was higher (~1.28) than that of either normal (~1.19) or IDA (~1.13) cells.

RED blood cells (RBC) must maintain their specific biconcave shape for optimal functions[1]. RBC membrane skeleton determines the shape, deformability (rheologic properties) and durability of the erythrocytes[1,2]. Some important factors that correlate the shape and deformability of such cells include elastic undulation[3], viscous adaptation[4], redox load[4], and presence of excess surface area of the red cell membrane[5]. The latter depends on the presence of membrane proteins like spectrin acting on the inner interface between the membrane and the cytoplasm and on sialic acid residues acting on the outer interface between the membrane and the plasma[5]. The shape and deformability of red cells therefore serve as clinical predictors of different forms of anemia and other hematological disorders[6,7]. The deformability of red cells is measured using dynamic techniques like Laser-assisted Optical Rotational Cell Analyser (LORCA)[7] and shear stress diffractometry[8]. Digital image analysis is another important method to extract shape-related information[9,10].

Interestingly, not much attention has been paid to the characteristic image histogram of RBCs. Digital images of a normal RBC yield an image histogram that is relatively insensitive to cell numbers, the extent of microscopic illumination, image magnification or even resolution of the image processing system. A gray image in which the objects are dark and the background is white typically shows two peaks, one corresponding to the background and the other corresponding to the object. If the image consists of semi-transparent objects, the situation is more complicated. Bright field microscopic images of RBCs fall under this category. The extent of transparency depends on the concentration of light-absorbing species (primarily haemoglobin for RBCs), and the overall optical path, which is likely to be associated with the extent of the biconcavity of the erythrocyte surface. While the absorbing elements would contribute primarily to the object peak containing darker pixels, the brighter central pallor of the RBC would produce a second peak containing whiter pixels. The relative positions of the two peaks are naturally correlated with the biconcavity of the erythrocytes. The present paper is an attempt to differentiate RBC images taken from normal and pathophysiological conditions on the basis of such surface topological features.

Blood smears of normal, transfusion-dependent beta-thalassaemia and iron deficiency anemia (IDA) patients (seven from each group) were obtained in triplicate, from the Department of Haematology, Calcutta Medical College, Calcutta. For microscopic observation a 100X oil immersion objective was chosen and the smear was focused in the feathered tail of the slide where individual RBCs just touched one another. A bright field microscope equipped with a CCD imaging system (camera and frame grabber, Diastar, Reichert Scientific Instruments, USA) was used for image digitization. Other available standard images with equivalent magnification collected from web sources[11,12] were also processed for comparison. On the average a typical field of view contained 50–80 RBCs.

For the purpose of analysis the RGB images ($768 \times 512 \times 3$) were converted into their corresponding gray (256 levels) format. The image histograms were then determined[13,14]. A typical image histogram provided an average distribution of transmitted and absorbed light through an ensemble of RBCs. The bin centres adjacent to the 256 gray levels and the corresponding occurrence frequency were stored for data comparison. After thresholding, the image objects were identified on the basis of connected pixels. This method enabled us to follow the morphological parameters of each individual cell. Matlab (Mathworks) programming environment was used for this analysis. Once objects were identified, platelets and nucleated cells were eliminated using a colour separation technique. A pseudo RGB image was created from the binary image. Non-RBC objects were dumped into one of the RGB planes. The residual image served as the input object space. For each sample, an inventory was created and the following morphological parameters[14] were stored: *area, perimeter, elongation, roundness, Feret diameter* and *compactness*. The term elongation was used to indicate the ratio of the major and minor axes of the minimum ellipse that will contain the object. The Feret diameter is

*For correspondence. (e-mail: adgcal@giascl01.vsnl.net.in)

the highest length dimension that can span a given object. The quantities 4π·area/perimeter^2 and $\sqrt{(4\pi \cdot \text{area})}/(\pi \cdot \text{major axis length})$, respectively represented the roundness and compactness of the object. If the object is a perfect circle, each of the two latter quantities should be close to unity. We have used the inventory to evaluate the statistical distribution of the morphological parameters. The same could be used for picking out cells with unusual morphological character. Parameters like compactness and elongation deserve special mention as they were independent of magnification and represented inherent geometrical features of the cell.

Scaling relations (which were also insensitive) were obtained by considering log–log plots of different variables. The scaling relation between the area and perimeter expressed the fractal dimension of the contour[15–17]. If D is the fractal dimension, $D = 2/m$, where m is the slope of log (area) plotted against log (perimeter). For Euclidean surfaces, D has a value of unity. A fractal surface has a value of D greater than one. Unlike the mass fractal dimension, in which the length scale or magnification has to be artificially altered, the area–perimeter exponent can be determined using a snapshot image, assuming that the object size has some degree of variations. Such variations were indeed present and this helped us to calculate D with a reasonable degree of accuracy for both the inner and outer zones.

The inner and outer zones corresponded to regions that permitted light to pass at greater and lesser extent. Appropriate thresholding conditions were used to separate out the image matrices representing the inner and outer zones, respectively. A smaller ratio of the Feret diameters of the inner and outer zones for example, would reflect a smaller central pallor, i.e. a rounder and less concave RBC.

Our next question was the following. If we artificially mix all the RBC types in a hypothetical image plane above, can we pick up the right cell class from the knowledge we have henceforth gathered? A preliminary inspection revealed that if we plot the cells as phase points in a space spanned by the roundness and compactness axes, the different cell types would be clustered in different domains (data not shown). This permitted the introduction of a simple perceptron-based model. Perceptrons are known to have the important weakness[18] that they fail if the data are not linearly separable. We actually used this failure criterion to train the model. Our method involved a Monte Carlo simulation of the input neurons[19,20]. Neurons represented by a pair of morphological parameters, were randomly chosen within MEAN–SIGMA*σ to MEAN + SIGMA*σ across the mean vector M, σ representing the standard deviation vector and SIGMA being a chosen scalar multiplier. We attempted to train the network using hard limit neurons. The target classes (normal cell, thalassaemic cells and IDA cells) were assigned the class values [0, 0], [0, 1] and [1, 0], respectively. Assuming that initially the data set was linearly separable, a critical failure of the perceptron convergence occurred as we increased SIGMA from lower to higher values. The particular pair of parameters, which resisted a collapse of the perceptron for the highest values of SIGMA, was considered to be of the highest efficacy for discrimination.

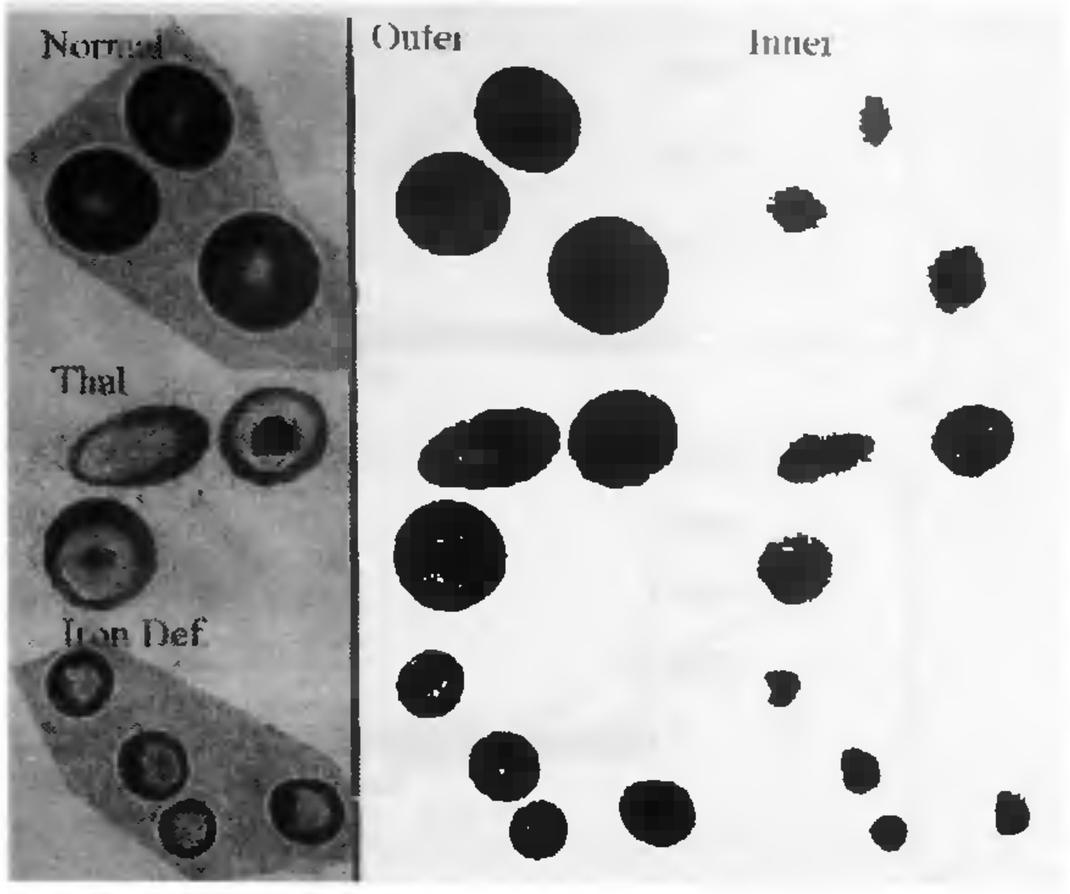

Figure 1. Types of image segmentation used for different classes of red blood cells. The digital image (1000×) is subjected to a pixel connectivity search, with white objects in dark background and dark objects in white background, respectively. The percolating pixels are identified as objects in each case. The two object classes correspond to the outer contour and the inner illuminated zone of the erythrocyte membrane, respectively.

Figure 1 illustrates the type of image segmentation used for different classes of blood cells. The percolating pixels were identified in each case. The two object classes corresponded to the outer contour and the inner illuminated zone of the erythrocyte membrane, respectively. The fractional decrease in area in going from the outer contour to the inner illuminated zone was most pronounced for normal cells, indicating less light transmission through the centre of these cells compared to the others.

Figure 2 shows the histograms observed in the three respective systems. The ratio of the object to the background peak was the highest for the normal cells when compared to the thalassaemic or IDA cells. Furthermore, the normal cell had a sharper object–background separation compared to the other two cell types.

Figure 3 shows the distribution of the morphological parameters. For normal cells, the area, perimeter and the Feret dimension of the inner and outer zones were distinctly separated. On the contrary, the inner and outer area histograms for thalassaemic cells showed a high

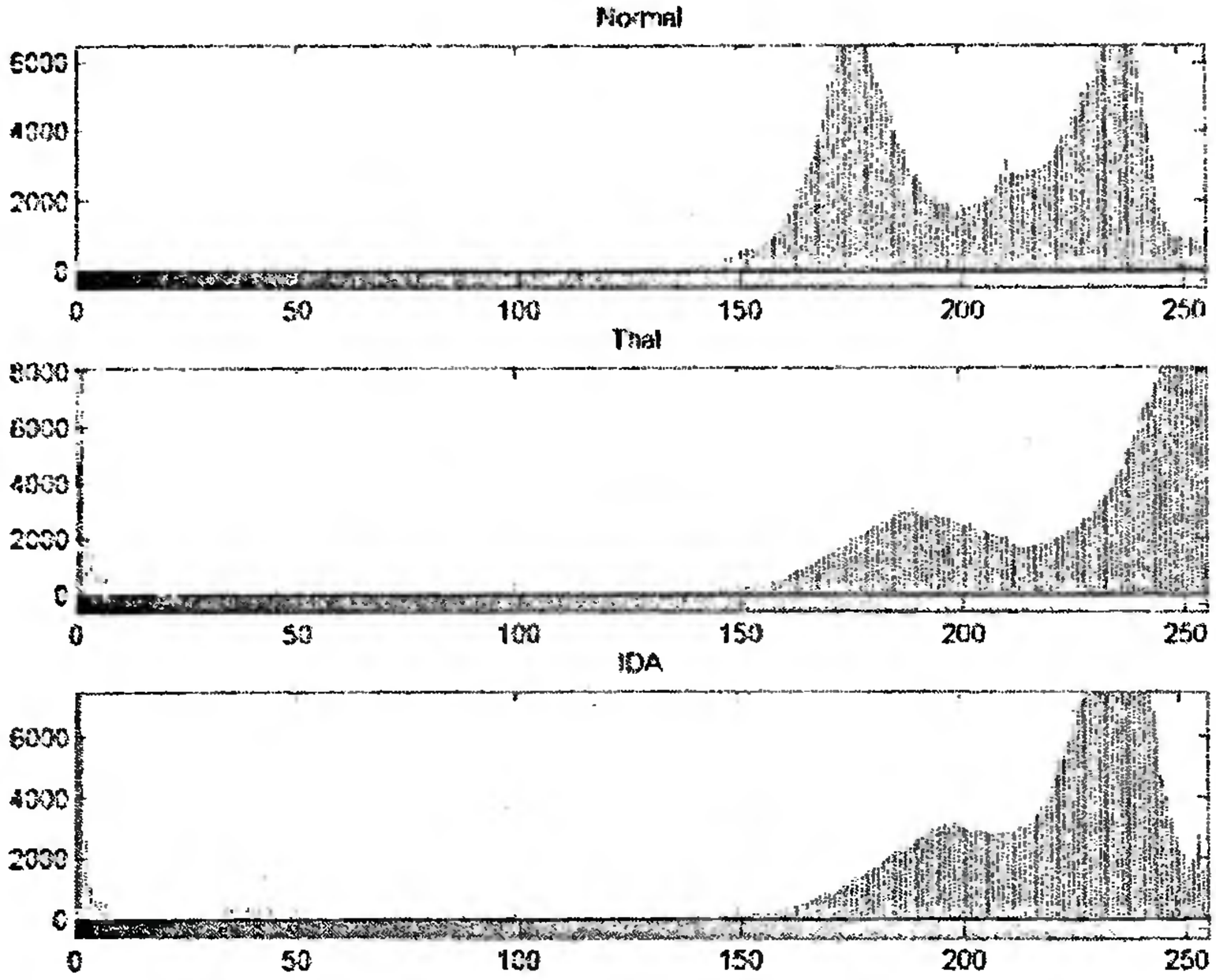

Figure 2. Image histograms of RBC from normal, thalassaemia and IDA patients. Histograms are evaluated using the 8 bit gray image with 256 gray levels. The horizontal panel below each histogram shows the shades of gray associated with different regions of the histogram.

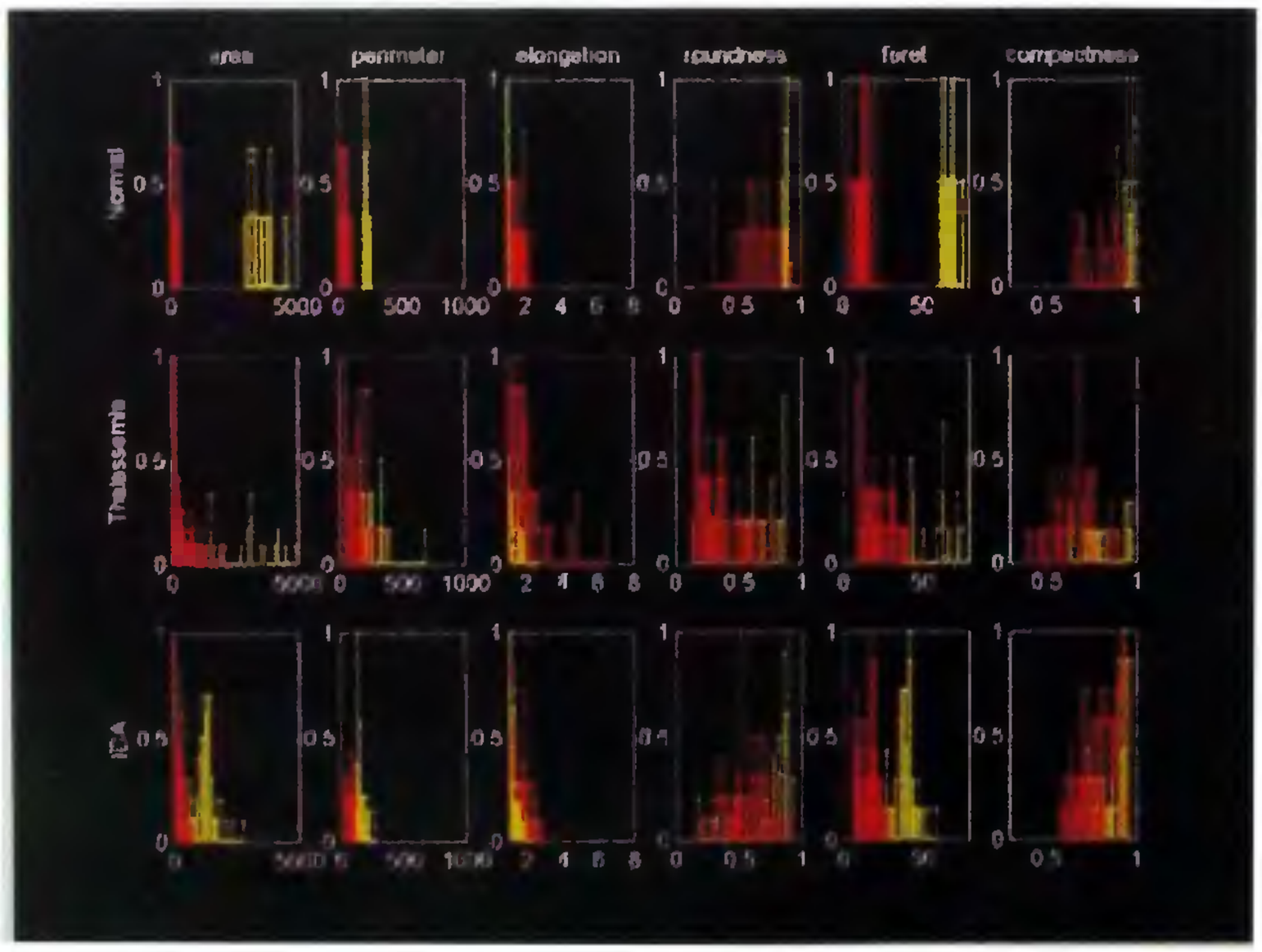

Figure 3. Distribution of morphological parameters that represent metric quantities like area or perimeter, or dimensionless entities like roundness, elongation or compactness. The red and yellow shades correspond to the outer contour and the inner zone of the RBCs, respectively.

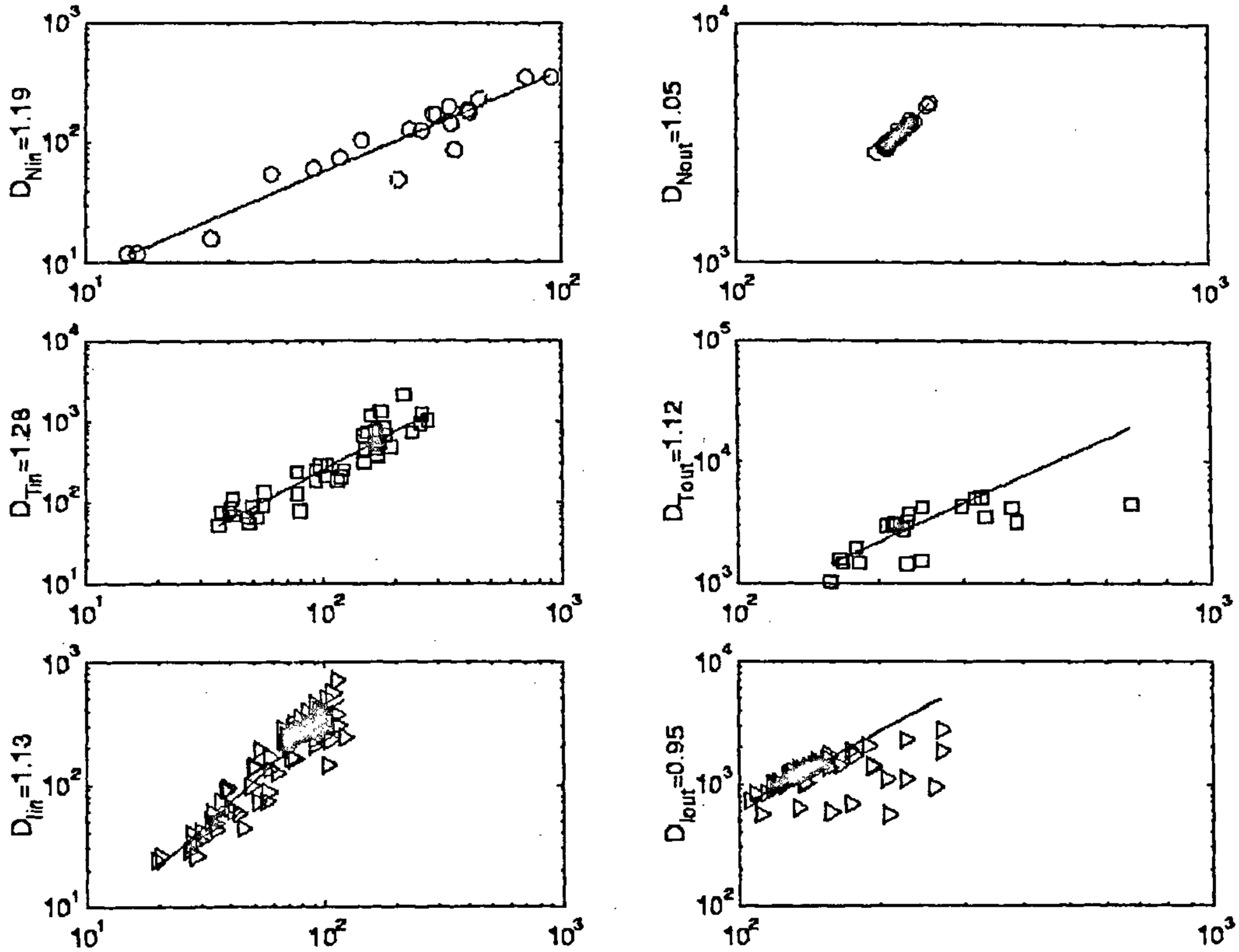

Figure 4. Log–log plot of area vs perimeter. The ordinate and abscissa represent the area and the perimeter values in the log–log scales respectively. *D* expresses the fractal dimension. Subscripts in and out correspond to inner and outer envelopes whereas subscripts N, T and I represent the normal, thalassaemic and IDA cells.

degree of overlap. IDA cells showed intermediate behaviour.

Figure 4 shows the log–log plots of area vs perimeter, the slope of which provides the fractal dimension. For normal cells, both the inner and outer zones show good linearity. For other cell types, the outer envelopes showed significant deviations from linearity beyond a critical length scale. The inner envelope however showed good linearity in each case. The inner zone *D* value for thalassaemia was 1.28.

Figure 5 *a* and *b* expresses the results of the learning model developed here. While simulations were performed for all pairs of morphological parameters, the roundness–compactness space provided the greatest discrimination. Interestingly, the classified data show a hierarchy, starting from the normal cell, to IDA and followed by thalassaemic cells. Such graded behaviour can also be observed in the image histograms (Figure 2) and suggests that this is a morphology-independent property of the image.

The red cell lifetime in thalassaemia is drastically shortened. This is attributed mainly to the oxidative damage to the cell membrane. Any change in the membrane organization is likely to affect the erythrocyte morphology[21]. Moreover, cross-linked membrane proteins result in greater rigidity and roughness, as occurring in the thalassaemic red cells[18]. A higher *D* value would in general correspond to increased roughness of the surface. The higher fractal dimension at the inner zone of thalassaemic red cells may thus be associated with increased membrane rigidity; the latter is believed to be the major contributing factor to the haemolytic process[22].

One lesson from the learning model developed here is noteworthy. Figure 5 *a* and *b* suggests that when cells are classified on the basis of the outer contours alone normal cells can be differentiated from those affected with IDA or thalassaemia. When they are classified on the basis of the concave inner zone, normal cells and IDA cells can be differentiated from thalassaemic cells.

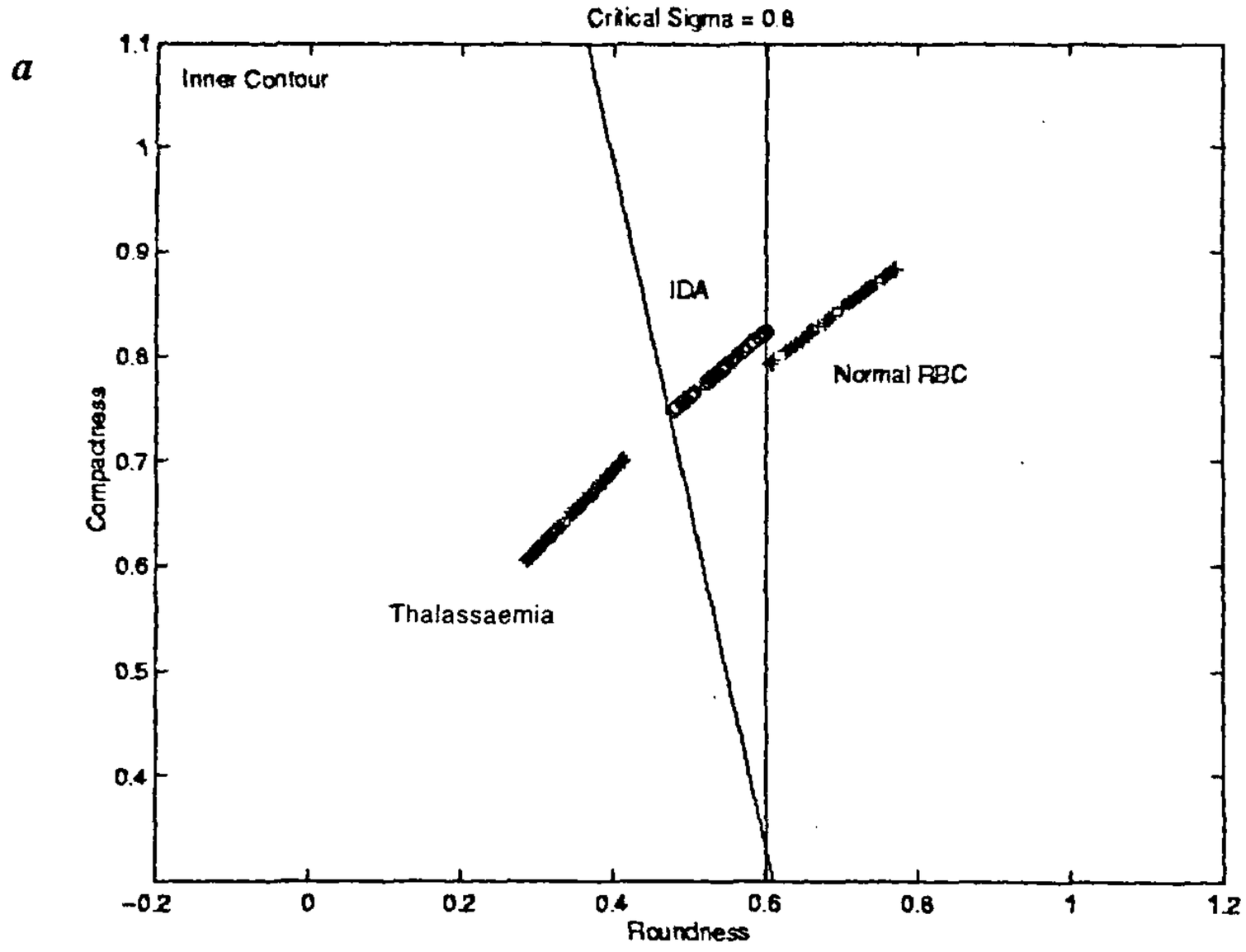

Figure 5 ***a***, ***b***. RBC classification based on critical linear separability method as described in the text. ***a*** and ***b*** correspond to the inner and outer zones of the RBC image.

A possible implication is that the biconcave depressed region is characteristically different in thalassaemia (a conclusion matching with the fractal analysis).

Finally, microscopic imaging and numerical processing of the image require simple and affordable equipment. It can be done with high reliability and at high speed with properly configured and automated software for diagnostic purposes without relying on the expertise of a haematologist and without the frailty of human judgement.

1. Lange, Y. and Steck, T. L., *J. Membrane Biol.*, 1984, **77**, 153–159.
2. Bossi, D. and Russo, M., *Mol. Aspects Med.*, 1996, **17**, 171–188.
3. Zeman, K., Engelhard, H. and Sackmann E. Hardeman, *Eur. Biophys. J.*, 1990, **18**, 203–219.
4. Wang, X., Wu, Z., Song, G., Wang, H., Long, M. and Cai, S., *Clin. Haemorpheol. Microcirc.*, 1999, **21**, 137–146.
5. Schmid-Schonbein, H., Heidtmann, H. and Grebe, R., *Blut.*, 1986, **52**, 149–164.
6. Dondorp, A. M., Angus, B. J., Chotivanich, K., Silamut, K., Ruangveerayuth, R., Hardeman, M. R., Kager, P. A., Vreeken, J. and White, N. J., *Am. J. Trop. Med. Hyg.*, 1999, **60**, 733–737.
7. Hardeman, M. R., Meinardi, M. M., Ince, C. and Vreeken, J., *Scand. J. Clin. Lab. Invest.*, 1998, **58**, 617–623.
8. Vaya, A., Martinez, M., Guillen, M. and Dalmau, J., *Clin. Haemorheol. Microcirc.*, 1998, **19**, 43–48.
9. Bacus, J. W., Belanger, M. G., Aggarwal, R. K., Trobaugh, F. E., Jr., *J. Histochem. Cytochem.*, 1976, **24**, 195–201.
10. Schonfeld, M. and Grebe, R., *Comput. Methods Programs Biomed.*, 1989, **28**, 217–224.
11. http://hms.medweb.harvard.edu/HS_Heme/AtlasTOC.htm.
12. http://pathy.med.nagoya-u.ac.jp/pathy.
13. Gonzalez, R. C. and Wintz, P. A., *Digital Image Processing*, Addison-Wesley, Boston, MA, 1979.
14. Thompson, C. M. and Shure, L., *Image Processing Toolbox User's Guide*, The Math Works Inc., Natick, MA, 1993.
15. Mandelbrot, B. B., *The Fractal Geometry of Nature*, Freeman, New York, 1983.
16. Stoyan, D. and Stoyan, H., *Fractals, Random Shapes and Point Fields: Method of Geometrical Statistics*, John Wiley & Sons, West Sussex, 1994.
17. Feder, J., *Fractals*, Plenum Press, New York, 1988.
18. Minsky and Papert, *Perceptrons*, MIT Press, Cambridge, MA, Expanded Edition, 1990.
19. Widrow, B. and Lehr, M. A., *Proc. IEEE*, 1990, **78**, 1415–1442.
20. Bishop, C. M., *Neural Networks for Pattern Recognition*, Clarendon, 1995.
21. Shinar, E., Shalev, O., Rachmilewitz, E. A. and Schrier, S. L., *Blood*, 1987, **70**, 158–164.
22. Kahane, I., Shifter, A. and Rachmilewitz, E. A., *FEBS Lett.*, 1978, **85**, 267–270.

ACKNOWLEDGEMENTS. We thank Prof. Utpal Chowdhury, Head Haematology Division, Calcutta Medical College, and Prof. P. K. Dasgupta, Department of Chemistry and Biochemistry, Texas Tech University, for their support, comments and help.

Received 23 October 1999; revised accepted 19 February 2000