Using Aspect Ratio to Classify Red Blood Images

Sherna Aziz Toma
College of Medicine Universiti of Baghdad -Iraq

Abstract: *In automated medical diagnosis, shape plays a key role in image processing and pattern matching. In particular, microscopic visual examination, as used in this paper, extensively uses shape to diagnose anemia using the shapes of Red Blood Cells (RBCs). This automated process depends entirely on the ability of the mammalian RBCs to change shape causing some types of anemia, which makes use of cell-discrimination on these varying RBC shapes an effective method of diagnosing anemia. In this automated diagnosis, image processing and pattern matching follows four steps: shape extraction, shape representation, shape size normalization and application of Fourier descriptors to obtain shapes of normal and abnormal RBCs. In testing the results, a client-server computer program applies invariant-moments to filter irrelevant shapes from the query, and aspect ratio of RBCs to improve on shape retrieval and verify results whereas increasing the geometrical number to reduces this accuracy by 10% to a promising 90%.*

Keyword: *Aspect Ratio, Fourier Descriptors, Invariant-Moments, Shape Extraction, Shape Representation, Image Processing, Red Blood Cells.*

I. INTRODUCTION

All Remarkable advancements made in the field of information technology have increasingly made available visual content in digital format, and along with it, an increasing need to make use of it. However, owing to a lack of a generalized description of visual information, search methods for visual content, especially images, are not as convincing as the present text-based search methods (Idris & Panchathan, 1997). As such, the existing image databases have limited search capabilities largely relying on image characteristics like color, size, texture and shape. However, of all these image characteristics, shape is emerging as the most salient characteristic of Content Based Image Retrieval (CBIR) system (Lam, Disney & Raicu, 2007).

Again, the advancements made in the fields of hematology, digital image processing and pattern recognition, have all contributed to the remarkable progress in the analysis and diagnosis of various illnesses in medicine. Particularly, as of interest to this study, the progress made in the analysis and classification of Red Blood Cells (RBCs) using their geometric features to diagnose of anemia (Bacus & Trobaugh, 1972; Engstrom & Lofvenberg 2012). Although there are two main shape representation and shape retrieval methods existing in scientific literatures, this paper will confine itself to the use of image processing and pattern recognition by the use of Fourier series in image processing and aspect ratio and moments method, in support of the results.

As stated, a careful examination of the geometry of RBCs can diagnose anemia. Similarly, Mauroy (2010) notes that RBCs in mammals have a biconcave discoid shape, which increases surface area for oxygen absorption and fluidity of movement in blood capillaries. He goes further to state that RBCs in the mammalian blood exhibit deformability, an ability to change shape to adopt various shapes. These irregular shapes make RBCs die prematurely or decrease their capacity to carry oxygen causing oxygen deprivation that lead to some types of anemia. Therefore, carrying out cell-discrimination on normal and abnormal shapes of RBCs is one method used to diagnose anemia. This follows the premise that the greater the number of deformed RBC, the greater the chance of a reduced supply of oxygen, a major feature of anemia (International Association for Pattern Recognition, 2003).

II. Extraction of Boundary Information (Preprocessing)

All This paper considers the outline shapes of RBCs in one-dimensional closed curves as illustrated in Figure 1 above. A digital camera fitted on an electron microscope examining a blood sample captures the initial RBCs images. From these RBCs shapes, boundary information is extracted in four steps as illustrated in Figure 2 below. The first step of extracting boundary information from the digitized RBC image converts the gray-scale RBC image into a binary image. This process of conversion is sensitive to noise, which corrupts the shape margins by introducing isolated pixels and regions (Zhang & Guojun, 2009). Therefore, the next step eliminates this noise from the outline of the binary image. Although, the noise elimination process removes isolated pixels and regions from the shape boundary, it still introduces gaps in the outline, which the third step addresses. The third step reconnects the shape boundary by filling these gaps between the boundary

points using m-connectivity connection method (Gonzalez & Woods, 1992). Finally, the fourth step traces the shape boundary using 8-connectivity contour method that first obtains the boundary coordinates and then traces the boundary by joining these coordinates (Pavlidis, 1982). These four steps successfully retrieve the shape of RBC.
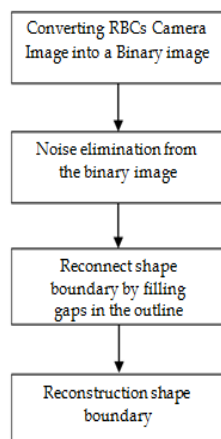

Figure1. Extraction of boundary information from RBC image

III.   Methods

A shape signature will then represent the shape obtained from the four-step process above. In this paper, Aspect Ratio shape signature will be used owing to its utility in general shape description as well as its prevalent in many recent Fourier Descriptors (FD) implementations. In defining the centroid-distance shape signature, let the boundary coordinates of the retrieved RBC shape be given by the function. Then the centroid-distance function will be expressed by the function

$$r(t) = \sqrt{[x(t) - x_c]^2 + [y(t) - y_c]^2}$$

Where

$$(x(t), y(t)), t = 0, 1, \ldots, L - 1,$$

and the distance from the central point (x_c, y_c) to the boundary points of the shape will be given by the function,

$$x_c = \frac{1}{L} \sum_{T=0}^{l-1} X(t)$$

$$y_c = \frac{1}{L} \sum_{T=0}^{l-1} X(t)$$

Owed to the expression of the function of the centroid- distance shape signature method by the distance from the centroid (central point) to the boundary points, this central point subtraction is almost constant throughout the shape, demonstrating that the centroid-distance shape signature is invariant to translation (Otterloo, 1991).

Having both reconstructed and represented RBC shape by the centroid-distance shape signature, the size of the RBC shape will need to be normalized. In normalization, the shape is sampled to a fixed number of points. Usually, objects have varying sizes that also makes the number of data points to vary. However, for the purposes of matching, sampling ensures that shape signatures have the same number of data points. Sampling process normalizes as well as eliminates any leftover noise from the shape boundary. Apart from the size of a shape, the length of the boundary also determines the number of sample points. A larger number of sample points give more details of the shape boundary that work to increase the accuracy of matching. Shape size normalization of the RBC shape will use the equal arc-length method, which Davies (1997) points out achieves the best equal space effect. For each extracted RBC shape, the shape is marked with varying contours points (32, 64, etc) with equal arc-length space between them (Otterloo, 1991). By successfully eliminating all the noise

from the shape boundary, normalization increases robustness of shape matching that is critical in shape description (Velasco, 1992).

Fourier Descriptors will then be used to carry out shape indexing on the normalized RBC shape derived from the centroid-distance shape signature. The Fourier shape descriptors are formed by the Fourier transformed coefficients, where the lower frequency FD contain information about the general shape features and the higher frequency contain the finer details of the shape (Raicu et al., 2007). Usually, the number of Fourier coefficients obtained from a transformation is large, however, only a subset of Fourier coefficients will be used to capture the overall features of the shape (Kauppinen, Seppanes & Pietikainen, 1995).

In the application of Fourier descriptor on the centroid-distance shape signature, for a given RBC shape signature as described above , the expression assumes that during sampling the shape signature is normalized to N points. Thus, Fourier Descriptor (FD)

$$u_n = \frac{1}{N} \sum_{t=0}^{N-1} s(t)\, exp\left(\frac{-j2\pi nt}{N}\right), n = 0,1,\dots,N-1$$

$$S(t)\,, t = 0,1,\dots L$$

$$u_n, n = 0,1,\dots,N-1$$

Fourier Descriptors (FD) of the shape, retrieval focuses only on the outline features of similar shapes, disregarding the position, size and rotation of the shape. However, to make the transformed RBC shapes and the database shapes comparable, shape description must be invariant to rotation, translation and scale. Owed to the fact that the centroid-distance shape signature is invariant under translation, its corresponding FDs are also invariant to translation. FDs can be made invariant to rotation by ignoring the phase information and taking only the magnitude values of the FDs. FDs can also achieve scale normalization by division of the magnitude values of the first half of FDs by the magnitude value of the first descriptor (DC component). The invariant feature used to index the shape is given by the function; For centroid-distance shape signature, its function is a real value, and there are only N/2 different frequencies in the Fourier transform. This implies that only half of the FDs are needed to index the shape. Thus, by dividing the magnitude of the first half of FDs by the DC component computes the Fourier transform series. Table 1 below shows four shapes of RBC obtained after application of Fourier series derived from centroid-distance shape signature on retrieved RBC image.

Table: 1 RBC OBTAINED after application of Fourier series and Aspect ratio

No.	RBC Shape	Description	Aspect Ratio $(in\ \mu m)$
1.		Normal shape of a red blood cell	7.8:2.5
2.		An abnormal spherical-shaped red blood cell.	7.4:2.2
3.		An abnormal oval-shaped red blood cell.	7.3:1.9
4.		An abnormal tear (sickle)-shaped red blood cell	6.3:1.6

III. Test Results

The four shapes of RBC in Table 1 above are the results of the application of Fourier series on the centroid-distance shape signature. The shapes are closed curves in 2-Dimensional, taken from the top view of RBC. The first shape represents a normal RBC while the other three represent abnormal shapes of RBC that have been demonstrated in medical literature to cause anemia. By successfully matching these retrieved shapes with RBC shapes found in the database with the use of invariant moments and aspect ratio of RBC, the percentage of abnormal shaped RBC in a blood sample can be established, hence, diagnosing anemia.

These shapes are then matched with RBC shapes found in the database to determine the accuracy of the shape retrieval method used. Pattern matching follows a client-server program where database is queried by example. In this program, the example image is the retrieved RBC shape that is fed into the BCR system The client executes a query that the server processes. By combining Fourier Descriptors with invariant moments (moments that are invariant to rotation, translation and scale), shapes that do not meet this criteria are filtered and got rid of the query (Niblack et al. 1993). In addition, using the aspect ratio of RBC, that is the relation of the width and height of RBC, refines retrieval of the results. Additionally, by varying the number of geometrical

features used, four and eight, the method can determine the accuracy of comparative analysis of the retrieved RBC shapes with the database shapes. Refer to appendix for the screenshots from the program.

IV.   Conclusion

Following the discussion in this paper, shape occupies a central role vital in image processing and pattern matching. From an input RBC shape represented by an appropriate shape signature, Fourier series is derived to obtain a robust shape used to query a database for comparative analysis. Combining FD with invariant methods filters and rids of irrelevant RBC shapes from the query. Then applying aspect ratio refines the process image retrieval. By changing the number of geometrical features in the query, the experiments establishes that the use of four geometrical features in pattern matching gives a conclusive result of 92% in determining the spherical shape of RBCs, whereas increasing the geometrical number to reduces this accuracy by 10% to a promising 90%.

References

[1]   Bacus, W., J., & Trobaugh, E., F. (1972). An Automated Classification of Peripheral Blood Cells by Means of Digital Image Processing. Sao Paulo, Brazil, 47(48).
[2]   Davies, E., R. (1997). Machine Vision: Theory, Algorithms, Practicalities. Academic Press.
[3]   Engstrom, G., & Lofvenberg, E. (2012). Treatment of Myeloproliferative Disorders with Hydroxyurea: Effects on Red Blood Cell Geometry and Deformability. American Society of Hematology
[4]   Gonzalez, R., C., & Woods, R., E. (1992). Digital Image Processing. Addison-Wesley
         Idris, H., & Panchathan, T. (1997). Image and Video Indexing Techniques. Journal of Communication and Image Representation. 8(2), 146-166 International Association for Pattern Recognition. (2003). What is Pattern Recognition. IAPR Newsletter, 25(1), 1.
[5]   Kauppinen, H., Seppanes, T., & Pietikainen. (1995). An Experiment Comaprison of Fourier-Based Descriptors in 2D Shape Classification. IEEE Trans, 17(2), 210-207.
[6]   Lam, M., Disney, T., & Raicu, J. (2007). Content Based Image Retrieval: Medical Informatics and Imaging,65(16).
[7]   Mauroy, B. (2010). Red Blood Cells in a Capillary. Labaratory MSC, University Paris 7/ CNRS – France
[8]   Niblack, W. et al. (1993). The QBIC Project: Querying Images by Content using Color, Texture and Shape. SPIE conference, 19(8), 173-187.
[9]   Otterloo, D., J. (1991). A Contour-Based Approach to Shape Analysis. London: Prentice Hall
[10]   Pavlidis, T. (1982). Algorithms for Image Processing and Pattern Matching. Computer Science Press.
[11]   Persoon, E., & Fu, K. (1995). Shape Discrimination Using Fourier Descriptors. IEEE Trans. On Sytems, Man and Cybernetics, 7(3).
[12]   Raicu, D., et al. (2007). Semantics and Image Content Integration for Pulmonary Nodule Interpretation in Thoracic Computer Tomography.
[13]   Ritendra, D., Joshi, D., & Wang, J., Z. (2008). Image Representation: Ideas, Influences and Trends. ACM Computing Surveys, 40(2), 1-60.
[14]   Smith, J., R., & Chang, S., F. (1996). An Automated Content-Based Image Retrieval System. ACM Multimedia, Boston.
[15]   Tieng, M., Q., & Boles, W., W. (1997). Recognition of 2D Object Contours using the Wavelength Transform Zero-Crossing Representation IEEE Trans, 19(8).
[16]   Velasco, G., E. (1992). Mathematical Analysis: Fourier Series. American Mathematical Monthly, 99(5), 427.
[17]   Yong, R., Huang, S., T., & Chang, S.(1991). Image Retrieval: Current Techniques, Promising Directions and Open Issues.
[18]   Zhang, D., & Guojun, L. (2009). A Comparative Study on Shape Representation Using Fourier Descriptors with Different Shape Signatures. Churchil: Australia.