

Comparison of Shape Signature Sub-Sampling Methods for Cell Tracking

Michael Majurski, Christopher Zheng, Joe Chalfoun, Alden Dima, Mary Brady
Software and Systems Division, Information Technology Lab
National Institute of Standards and Technology
Gaithersburg, USA
michael.majurski@nist.gov

Abstract— New microscope technologies are enabling the acquisition of large volumes of live cell image data. Accurate temporal object tracking is required to facilitate the analysis of this data. One principal component of cell tracking is correspondence, matching cells between consecutive frames. This component can be enhanced by incorporating shape metrics into the tracking model. The measure of shape similarity between two objects can be accomplished using Fourier descriptors, derived from a one dimensional shape signature. The type of one dimensional shape signature affects the quality of the resulting Fourier descriptors when dealing with noisy object boundaries. We present a comparison of several different sampling methods for converting a two dimensional object boundary into a one dimensional shape signature suitable for computing Fourier descriptors. The Fourier descriptors are evaluated on two shape datasets with ground truth and used as a similarity feature for performing object tracking on a time-lapse series of NIH 3T3 cells. Experimental results show that for noisy object boundaries, Fourier descriptors constructed from the R-Theta Binning centroid distance shape signature presented here perform better than Fourier descriptors constructed from the other evaluated shape signatures.

Index Terms—Fourier Descriptor, Shape Analysis

I. INTRODUCTION

Object shape is a familiar aspect of the human visual perception system. Shape information is often valuable as a component of object tracking systems, both as a means of improving the accuracy of object trajectories as well as a filter for creating a subset of objects to study. This work is motivated by the inclusion of a cell shape feature descriptor in the cost function of an overlap-based cell tracker which operates on a time-lapse series of live cell images [1]–[3]. Shape descriptors are usually extracted using computer algorithms that convert object geometrical pixel information into one dimensional numerical vectors. Figure 1 shows a simplified workflow from raw cell images to Fourier descriptors. The raw cell images (a) are segmented to produce a mask (b-c) from which a cell boundary is extracted (d). The boundary is converted into a 1D shape signature (e) from which Fourier descriptors (f) are constructed.

Fourier descriptors (FD) are often used in the literature due to their invariance to translation, rotation, and scaling. Two surveys by Loncaric [1] and by Zhang and Lu [2] present shape analysis techniques including FD. Hu and Li present a novel shape signature and compare the performance of FD based on

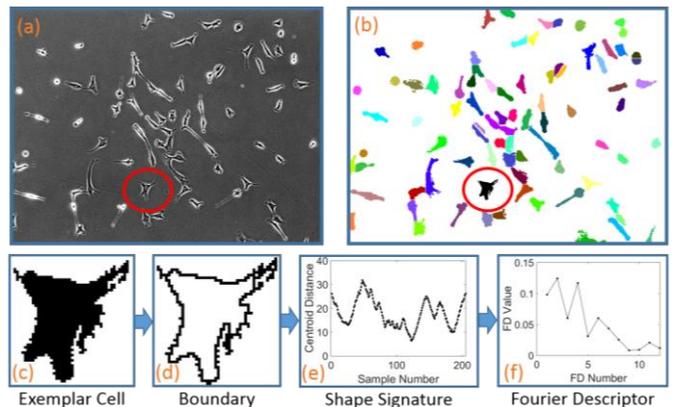


Figure 1: A simplified workflow from raw images to Fourier descriptors. (a) NIH 3T3 cells raw image at time point 238, (b) segmented image with exemplar cell circled, (c) exemplar cell mask, (d) cell boundary, (e) extracted shape signature, (f) Fourier descriptor.

different shape signatures [3]. Zhang and Lu evaluate several methods of constructing FD from shape signatures and conclude that shape signatures based on centroid distance are optimal for shape representation [4]. In [5] Zhang and Lu evaluate an expanded pool of shape signatures and confirm their prior conclusions on centroid distance shape signatures.

From the literature listed above, we will use Fourier descriptors constructed from centroid distance shape signatures for similarity measurement. However, cell images present a high uncertainty on the boundary location as discussed in Dima [6]. This translates into noisy boundary pixel locations that can affect accuracy of shape descriptors. In order to address this problem we compared several boundary sampling techniques with regard to their ability to reduce noise and increase the accuracy of shape metric comparison in cellular images.

The rest of this paper is organized as follows: Section II introduces the cell tracking application dataset. Section III covers the definition of centroid distance and the boundary sampling methods. Section IV covers the construction of FD from 1D shape signatures. Section V details the methodology used to evaluate the FD. Section VI discusses the evaluation results.

II. APPLICATION DATASET

The cell tracking application dataset consists of live National Institutes of Health (NIH) 3T3 cells imaged in phase

contrast at 15 minute intervals; see [7] for more details. To visually demonstrate shape signatures, an informative exemplar cell was selected from the last time point in the NIH 3T3 dataset. The grayscale image of the NIH 3T3 cells is shown in Figure 1 (a), next to the corresponding segmented mask with the exemplar cell circled (b), and an enlarged view of the exemplar cell (c).

III. SHAPE SIGNATURE CONSTRUCTION

From Figure 1 image segmentation converts the input image into a set of 8-connected pixels forming regions (cells). Boundary pixels can be extracted from each region as an ordered sequence of pixel locations. This set of 2D boundary pixel locations is then converted into a 1D vector called a shape signature using the centroid distance function described below. The shape signature is sub-sampled to reduce noise and then converted into a set of Fourier descriptors.

A. Centroid Distance Function

The centroid distance function is used to convert a set of 2D pixel locations into a 1D vector from which Fourier descriptors are computed. The centroid distance function, shown in Equation 1, computes the Euclidian distance of each boundary pixel with coordinates (x_i, y_i) to the centroid of that object with coordinates (\bar{x}, \bar{y}) .

$$CD_i = \sqrt{(x_i - \bar{x})^2 + (y_i - \bar{y})^2} \quad (1)$$

where $i = 1, \dots, N_b$, N_b is the number of boundary pixels in the segmented object. The centroid of a segmented object, Equation 2, is defined as the average of the object's pixel locations, where N is the total number of pixels in the segmented object.

$$\bar{x} = \frac{1}{N} \sum_{k=1}^N x_k \quad \bar{y} = \frac{1}{N} \sum_{k=1}^N y_k \quad (2)$$

Shifting an object's coordinate system to its centroid makes the shape representation invariant to translation. The next section demonstrates how to construct the shape signatures.

B. Boundary Sampling

We explored three methods for sampling the boundary pixels: (1) Every Pixel, (2) Interpolate N Points, and (3) R-Theta Binning. The first two methods require a list of boundary pixels ordered using a boundary tracing; like the Moore algorithm [8] which sorts the boundary pixels in a clockwise or counter clockwise order. The last method, R-Theta Binning, avoids the boundary tracing by sorting the pixel locations directly.

1) Every Pixel (EPSig)

A simple method for sampling the boundary pixels is to take every pixel on the boundary without omitting any information or sub-sampling.

2) Interpolate N Points (INPsig)

This shape signature creates a set of N points interpolated from the 1D centroid distance function. The collection of interpolation methods tested includes nearest neighbor, linear, cubic, and spline. The number of sample points N is adjustable.

3) R-Theta Binning (RTBsig)

This shape signature was created to avoid a perceived pitfall in the previous methods, the distortion of object shape in polar coordinates. Since the purpose of the shape signature is to create a periodic 1D function as input for a Fourier transform, the effective shape is implicitly represented by polar coordinates. RTBsig is created by converting the edge pixels (x_i, y_i) into polar coordinates (r_i, θ_i) and sorting them by angle (θ_i) . This 1D function is then binned into N bins to produce a shape signature containing N sample points. The maximum value within a bin is considered the value of that bin. For bins without a value, the nearest non-empty bin value is used.

IV. FOURIER DESCRIPTOR CONSTRUCTION

The conversion of a shape signature (1D vector of distances) into a set of Fourier descriptors starts with performing a discrete Fourier transform on the shape signature. This produces a vector of Fourier coefficients that can be normalized to construct Fourier descriptors. The Fourier transform of the shape signature $s(k)$, shown in Equation 3, produces Fourier coefficients $F(n)$ with the Fourier coefficient index $n = 0, \dots, N - 1$ as given by:

$$F(n) = \frac{1}{N} \sum_{k=0}^{N-1} s(k) e^{-2\pi i n k / N} \quad (3)$$

Rotational invariance of the Fourier coefficients can be achieved by ignoring phase information and only considering magnitude values, $F(n) = |F(n)|$. In ignoring phase information, the Fourier descriptors are symmetric so only the first half are kept, $F(0) \dots F(\frac{N}{2})$. Size invariance is achieved by normalizing the Fourier descriptors with respect to the magnitude of the DC component, $F(n) = F(n)/|F(0)|$. By definition, the first normalized descriptor $\frac{F(0)}{F(0)}$ will be equal to 1 and is ignored. Thus the Fourier coefficients have been transformed into Fourier descriptors, shown in Equation 4, having been made invariant to translation, rotation, and scale.

$$FD = \frac{|F(1)|}{|F(0)|}, \frac{|F(2)|}{|F(0)|}, \dots, \frac{|F(\frac{N}{2})|}{|F(0)|} \quad (4)$$

Low order Fourier descriptors contain information about an object's general shape while higher order Fourier descriptors contain more detailed information. Generally the first few Fourier descriptors are sufficient to capture overall object shape. The literature survey suggests that the first 15 descriptors are sufficient, so descriptors beyond that are discarded [2], [4], [5].

V. EVALUATION METHODOLOGY

To evaluate FD performance when constructed from different shape signatures we used the precision and recall metrics on two datasets with ground truth. The MPEG-7 CE Shape-1 Part B shape dataset [9] and a Vehicle Silhouettes dataset [10]. The images in the MPEG-7 dataset are grouped into 70 classes of perceptually similar objects and the images in the Vehicle Silhouettes dataset are grouped into 4 classes.

Datasets with ground truth shape classification enable the comparison of the FD shape classification with the ground truth shape classification. To evaluate FD performance, precision-recall curves are generated for each evaluated shape signature. Precision is the fraction of retrieved records that are relevant and recall is the fraction of relevant records that are retrieved. Given the set of relevant records R and the set of all records retrieved S , Equation 5 defines Precision and Recall.

$$Precision(R, S) = \frac{|R \cap S|}{|S|} \quad Recall(R, S) = \frac{|R \cap S|}{|R|} \quad (5)$$

To generate the precision and recall data, each shape in the dataset was queried against the remainder of the dataset. The similarity between two shapes is computed as the Euclidean distance between the respective two Fourier descriptors. The relevant records are the other members of the queries ground truth class. For this evaluation the number of RTBsig bins (boundary sample points) is set 360, or one sample per degree in the polar coordinate representation. This value is selected by exploring its effect on FD precision and recall in the following section.

In addition to validating the Fourier descriptors on ground truth datasets, a visual evaluation is performed using the time-lapse NIH 3T3 dataset and a cell tracking application which mimics using the cell shape feature as part of the cost function to determine cell correspondence between time points.

VI. RESULTS

This section is divided in two parts: (1) precision and recall results on the two datasets with ground truth classification and (2) visual evaluation of NIH 3T3 cell association tracking application.

A. Datasets with Ground Truth

Average precision and recall curves were generated for each shape signature type and displayed by dataset in Figure 4. The higher the precision of the retrieval results, the higher the fraction of correct retrieved shapes by the descriptor. For both datasets the INPsig precision for identical recall values differed from EPsig precision by less than $\pm 1.5\%$ across all interpolation methods. For visual clarity only EPsig is shown compared to RTBsig.

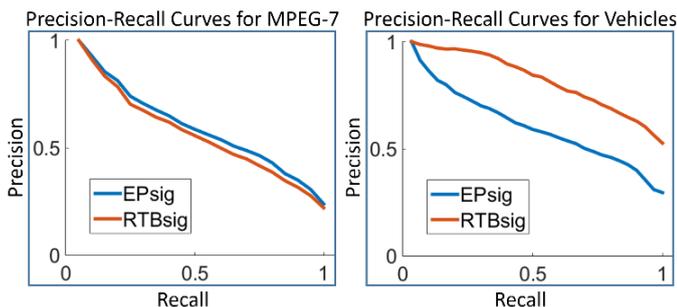


Figure 4: Precision-Recall curves plotted by shape signature. MPEG-7 CE Shape-1 Part B (left). Vehicle Silhouettes (right).

For the MPEG-7 dataset there is little to differentiate the performance of the shape signatures. EPsig and RTBsig vary by up to $\pm 5\%$. However, on the Vehicle Silhouette dataset the



Figure 5: Sedan Example

RTBsig performed better (up to 25 % higher precision) than the techniques which rely on a boundary tracing. We hypothesize that this effect arises in the Vehicle Silhouette dataset as a result of segmentation errors causing the boundary walk to distort the shape signature. This is best demonstrated using an example sedan image from the Vehicle Silhouettes dataset shown in Figure 5 with the distortion marked. The sedan’s EPsig, polar representation, and Fourier descriptors are shown in Figure 2. The same image with the RTBsig is shown in Figure 3 with the distortion marked.

The segmentation error in the front wheel of the sedan causes a significant dip in the EPsig (shown by the red arrow).

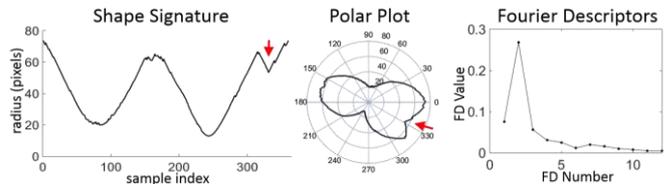


Figure 2: EPsig (left), with its polar representation created by treating the shape signature as a 1D function and rescaling its x values into 0-360 degrees and plotting in polar space (center), and its Fourier Descriptors (right). The effect of the segmentation error is marked with an arrow.

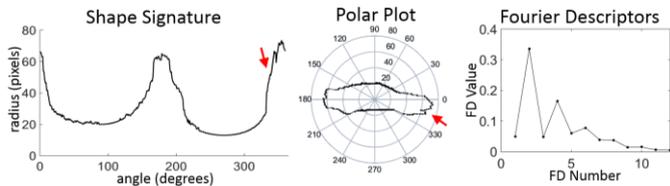


Figure 3: RTBsig (left), with its polar representation (center), and its Fourier Descriptors (right). Effect of the segmentation error is marked with an arrow.

However, this segmentation error does not distort the human perception of the general shape of the sedan. Therefore, while the EPsig captures the segmented mask’s shape, the mask is misrepresenting the object’s shape due to segmentation errors. The RTBsig corrects this by ignoring regions where the shape boundary doubles back on itself as long as boundary points with larger radius values for the same angle exist. With our data this effect manifests only in the Vehicle Silhouettes because the MPEG-7 shapes all have smooth, clean boundaries where any difference in the boundary tracing can be attributed to difference in shape.

The cell tracking application contains a segmentation step where errors similar to the front tire of the sedan are possible. The exemplar cell contains just such segmentation errors, shown circled in Figure 6. This suggests that the RTBsig will perform better than the other two shape signatures for the cell tracking application.

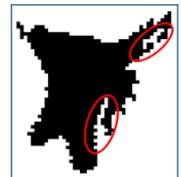


Figure 6: Exemplar cell mask with segmentation errors circled.

Sub-sampled shape signatures (INPsig) were evaluated to discover that no sub-sampling method produced higher precision-recall curves than EPsig. Therefore INPsig cannot reduce the effects of boundary noise introduced by segmentation error. Interestingly, for both datasets using INPsig with only 10 % of the total number of

boundary samples produced precision-recall values that differed by less than $\pm 2\%$ from EPsig.

Adjusting the number of samples in RTBsig affects the FD accuracy. Using 360 ± 90 RTBsig samples results in precision-recall curves that vary by less than $\pm 1\%$ from RTBsig with 360 points. For sample counts less than 270, decreasing the number of samples decreases the resulting FD precision and recall.

B. NIH 3T3 Cell Tracking

To perform a visual evaluation of the Fourier descriptor performance for the cell tracking application, a set of m cells are randomly selected from time point t within the NIH 3T3 time-lapse image dataset and the n cells in time point $t + 1$ with the closest shape are displayed. The Fourier descriptors are used as a similarity feature to perform matching between adjacent time points. An example set of results is shown in Table 1 where 6 cells from time point (image number) $t=194$ are shown with the nearest 5 cells from time point $t=195$, the time point within the sequence was selected randomly. The distance between the query and each result is included below each shape in the table.

TABLE I. NIH 3T3 CELL ASSOCIATION TRACKING EXAMPLE

Query (time t)	Nearest n cells (time $t+1$) with distance from query				
					
	0.2280	0.4389	0.5767	0.6347	0.6938
					
	0.1303	0.1510	0.1539	0.1686	0.1708
					
	0.3393	0.4425	0.4660	0.5874	0.6047
					
	0.1054	0.1448	0.1668	0.1845	0.2053
					
	0.1812	0.2354	0.3184	0.3188	0.3868
					
	0.0818	0.1102	0.1141	0.1549	0.1747

Generally the more unique the cell shape, the easier it is for the shape descriptor to match the corresponding cell in the next time frame. The simpler shapes, for example the last row of Table 1, proved more difficult to match because cell deformation during the time between image acquisitions causes changes in the observed cell shape. A deformation of the same magnitude will have a larger effect on a simpler shape. The first row of Table 1 shows an example where the query cell deformed slightly between images but the closest match is correct.

VII. CONCLUSIONS

Three methods of constructing FDs based on centroid distance shape signatures were analyzed. We concluded that if noise from cell boundary segmentation error is a concern then the R-Theta Binning signature provides superior shape discrimination as part of a cell tracking framework for NIH 3T3 cells imaged and segmented in phase contrast. Conversely, for low noise object boundaries the 1D shape signature provides little differentiation. The number of sample points in the shape signature was analyzed to discover that sub-sampling alone cannot reduce the effect of boundary noise for the evaluated datasets.

The R-Theta Binning shape signature was utilized to establish correspondences between cells with noisy boundaries in adjacent time points of a time-lapse series of NIH 3T3 live cells. The results showed that the more unique cell shapes were more likely to find a correct match between time frames. With simpler shapes the descriptor had more difficulty correctly discerning correspondence because cell deformation between images became a dominant factor. Overall the R-Theta Binning based Fourier descriptor successfully provides a pool of potential cell associations between time points as part of the cell tracking cost function.

ACKNOWLEDGMENT

This work has been supported by NIST. We would like to acknowledge the team members of the computational science in biological metrology project for providing invaluable inputs to our work. We would like to acknowledge Michael Halter from NIST Material Measurement Lab for acquiring the 3T3 cells.

REFERENCES

- [1] S. Loncaric, "A survey of shape analysis techniques," *Pattern Recognit.*, vol. 31, no. 8, pp. 983–1001, 1998.
- [2] D. Zhang and G. Lu, "Review of shape representation and description techniques," *Pattern Recognit.*, vol. 37, no. 1, pp. 1–19, 2004.
- [3] Y. Hu and Z. Li, "An Improved Shape Signature for Shape Representation and Image Retrieval," *J. Softw.*, vol. 8, no. 11, pp. 2925–2929, 2013.
- [4] D. Zhang and G. Lu, "A Comparative Study on Shape Retrieval Using Fourier Descriptors with Different Shape Signatures," in *International Conference on Intelligent Multimedia and Distance Education*, 2001, vol. 1, pp. 1–9.
- [5] D. Zhang and G. Lu, "Study and evaluation of different Fourier methods for image retrieval," *Image Vis. Comput.*, vol. 23, no. 1, pp. 33–49, 2005.
- [6] A. Dima, J. T. Elliott, J. J. Filliben, M. Halter, A. Peskin, J. Bernal, M. Kociolek, M. C. Brady, H. C. Tang, and A. L. Plant, "Comparison of segmentation algorithms for fluorescence microscopy images of cells," *Cytometry. A*, vol. 79, no. 7, pp. 545–559, Jul. 2011.
- [7] M. Halter, D. R. Sisan, J. Chalfoun, B. L. Stottrup, A. Cardone, A. Dima, A. Tona, A. L. Plant, and J. T. Elliott, "Cell cycle dependent TN-C promoter activity determined by live cell imaging," *Cytom. Part A*, vol. 79 A, no. 3, pp. 192–202, 2011.

- [8] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital Image Processing Using Matlab*. New Jersey: Pearson Prentice Hall, 2004.
- [9] "MPEG-7 Core Experiment CE-Shape-1." Temple University.
- [10] N. Thakoor, J. Gao, and S. Jung, "Hidden Markov model-based weighted likelihood discriminant for 2-D shape classification," *IEEE Trans. Image Process.*, vol. 16, no. 11, pp. 2707–2719, 2007.