Comparison of Shape Descriptors for Mice Behavior Recognition

Jonathan de Andrade Silva[1], Wesley Nunes Gonçalves[1], Bruno Brandoli Machado[1], Hemerson Pistori[2], Albert Schiaveto de Souza[3], and Kleber Padovani de Souza[2]

[1] Computer Science Department
University of São Paulo (USP) at São Carlos, Brazil
{jandrade,brandoli}@icmc.usp.br {wnunes}@ursa.ifsc.usp.br
[2] Research Group in Engineering and Computing
Dom Bosco Catholic University, Brazil
{pistori}@ucdb.br {kriowloo}@gmail.com
http://www.gpec.ucdb.br
[3] Department of Morphophysiology
Federal University of Mato Grosso do Sul, Brazil
{albertss}@hotmail.com

Abstract. Shape representation provides fundamental features for many applications in computer vision and it is known to be important cues for human vision. This paper presents an experimental study on recognition of mice behavior. We investigate the performance of the four shape recognition methods, namely Chain-Code, Curvature, Fourier descriptors and Zernike moments. These methods are applied to a real database that consists of four mice behaviors. Our experiments show that Zernike moments and Fourier descriptors provide the best results. To evaluate the noise tolerance, we corrupt each contour with different levels of noise. In this scenario, Fourier descriptor shows invariance to high levels of noise.

Key words: Computer Vision, Shape Descriptors, Mice Behavior

1 Introduction

Shape analysis is an important field of investigation that has been a subject of intense research for decades, further reinforced in recent years by applications in content-based information retrieval, human gait recognition, and medicine. Psychological studies have suggested that shape provides fundamental information for brain processing and are known to be important cues for human vision system [1]. In images, shape is considered one of the most important visual attributes to characterize objects. Large variations in shape, as well as varying scale, rotation and noise make the representation of shape particularly challenging from a computer vision perspective.

Several methods have been proposed in the literature and most of them can be classified into two categories: contour-based methods and region-based

methods. Contour-based methods explore boundary shape information. In this category, the most common methods are chain code [2], curvature [3], Fourier descriptors [4], and simple descriptors [1]. Though chain code and curvature methods demand low computational cost, they are easily influenced by noise. Otherwise, Fourier descriptor has low sensitivity to noise when low frequency Fourier coefficients are used [4]. However, Fourier descriptor does not provide local information as chain code and curvature methods.

Unlike the contour-based methods, region-based methods extract features from the whole shape region. Different moments, such as Zernike moments [5, 6] and Legendre moments [7] have been demonstrated to achieve excellent performance. Zernike moments has been suggested for shape description due to its superiority over other moments functions regarding to the robustness to deformations and noise [5].

This paper provides an experimental study of shape descriptor methods applied to mice behavior recognition, an important real-world application. This application aims at helping experts to determine the drug effect used for the medicine development and answering basic ethological questions [8]. Currently, behavioral analysis is carried out by means of visual inspection, which is considered a laborious and time consuming task. In order to enhance this procedure, we study automatic procedures through applying four shape methods (chain code, curvature, Fourier, and Zernike moments) on real database of mice behaviors. This database consists of four behaviors with similar contour characteristics, presenting an interesting challenge for the recognition task. According to the experimental results, Zernike moments and Fourier descriptors achieve the best recognition performances for mice behavior recognition. Further, results show that 20 descriptors seem to be sufficient to characterize the behaviors. Moreover, results from noised images show that Fourier Descriptor provides more invariance for high levels of noise.

This paper is organized as follows. Section 2 describes the fundamentals of the shape descriptor methods. In Section 3 experimental results and discussion have been provided. Finally, Section 4 concludes the paper.

2   Shape Descriptors

In this section we give a brief overview of the methods employed in this work. Before describing the methods, it is important considering that a parametric curve $c(t) = (x(t), y(t))$ consists of points belong to the contour of binary objects.

2.1   Fourier Descriptors

Fourier is one of the most popular descriptors for shape characterization. The basic idea of this method consists in applying the Fourier transform in a parametric curve. Initially, the complex signal $u(t) = x(t) + j\,y(t)$ is obtained from the parametric curve. Then the complex signal is expanded by a complex Fourier series defined as

$$U(s) = \frac{1}{L} \int_0^L u(t)e^{-j2\pi st/L} dt \tag{1}$$

where L is the curve perimeter.

The $FD(s)$ corresponds to the Fourier descriptors defined as $FD(s) = \frac{U(s)}{U(0)}$. Some properties from these descriptors can be discussed. For example, the 0-th component of U is associated with the centroid of the original shape. Moreover, the descriptors are also invariance to geometric transformations as a consequence of the properties of the Fourier series.

A useful property exhibited by this method is that the most of the contour information is concentrated along the first coefficients, since they contain most of the energy. Following this property, a feature vector with n features obtained from the Fourier descriptor consists of the n descriptors from $FD(s)$.

2.2 Curvature

Curvature has been identified as an important clue explored by the human visual system. Since visual information is highly redundant, curvature plays an important role in compressing information, such as straight line where the curvature is null.

The curvature $k(t)$ of a parametric curve $c(t) = (x(t), y(t))$ is defined according to Equation 2.

$$k(t) = \frac{\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)}{(\dot{x}(t)^2\dot{y}(t)^2)^{\frac{3}{2}}} \tag{2}$$

where $\dot{x}(t)$ and $\ddot{x}(t)$ are the first and second derivative of $x(t)$.

To extract useful information from $k(t)$, a strategy based on Fourier is used. Fourier series is applied on $k(t)$ and then a feature vector of dimension n is built with the n first coefficients. Since Fourier contain most of the signal's energy concentrated in the first descriptors, the strategy discussed above is useful to remove redundancy.

2.3 Zernike Moments

The Zernike moments from a binary image can be calculated in three steps: computation of radial polynomials, computation of Zernike basis functions, and projection of the image on to the basis functions. The computation of radial polynomials is defined according to Equation 3. Usually, the indexes n and m are called order and repetition.

$$R_{nm}(p) = \sum_{s=0}^{(n-|m|)/2} c(n, m, s)p^{n-2s} \tag{3}$$

where

$$c(n, m, s) = (-1)^s \frac{(n-s)!}{s!((n+|m|)/2 - s)!((n-|m|)/2 - s)!} \tag{4}$$

Using the radial polynomials, Zernike basis functions are described in Equation 5. These basis functions are defined within a unit circle and since they are orthogonal, there is no redundancy information between moments.

$$V_{nm}(p, \theta) = R_{nm}(p)exp(jm\theta), \quad |p| \leq 1 \tag{5}$$

Finally, the Zernike moments of order n and repetition m over an image $f(x, y)$ with size $N \times N$ is defined in Equation 6. The coordinates of the image must be normalized into $[0, 1]$. Thus, the magnitudes of the Zernike moments can be used as shape descriptors.

$$Z_{nm} = \frac{n+1}{\lambda_N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) V_{nm}^*(x, y)$$

$$= \frac{n+1}{\lambda_N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) R_{nm}(p_{xy}) exp(-jm\theta_{xy}) \tag{6}$$

where λ_N is a normalization factor related to the number of pixels in the unit circle, $p_{xy} = \frac{\sqrt{(2x-N+1)^2+(N-1-2y)^2}}{N}$, and $\theta_{xy} = tan^{-1}(\frac{N-1-2y}{2x-N+1})$.

2.4   Chain Code

The Freeman chain code is a compact way to represent the contour of an object. The method assigns an integer i, $0 \leq i \leq 7$ for each point of the contour according to its direction to a successor point. This representation is based on the connectivity definition of neighboring pixels, usually 8-connectivity. The direction of each movement is given by a vector $v_i \in \{(1,0), (1,1), (0,1), (-1,1), (-1,0), (-1,-1), (0,-1), (1,-1)\}$. For example, consider an starting point $s_0(3,3)$ and a successor point $s_1(3,4)$. The path from s_0 to s_1 is represented by the code 3 $(s_0 + v_3 \rightarrow s_1)$.

In the chain code method, the contour of a binary object is represented by a vector of codes. However, this strategy has two drawbacks: it depends on the starting point and it is not rotation invariant. These drawbacks can be overcome, for example, using Fourier transform. For this, the Fourier transform is applied in the vector of codes and then, the n descriptors for the chain code method is composed by the n first Fourier coefficients.

3   Experimental Results

This section presents two experiments that were conducted to evaluate shape recognition. The first experiment is devised to compare the performance of shape descriptor methods described in Section 2. In the second experiment, we analyze the influence of noise on the classification task.

In order to assess the performance of the methods using real data, experiments were performed using a database composed by images of mice. This database consists of 4 types of behavior, namely curved horizontal, self-cleaning, vertical and horizontal exploratory behaviors, each containing 50 images captured in an environment with artificial light. Each image was segmented using

background subtraction technique [9, 10]. The types of behavior have similar contour characteristics thus presenting an interesting challenge for classification tasks. Examples of the database can be seen in Fig. 1.
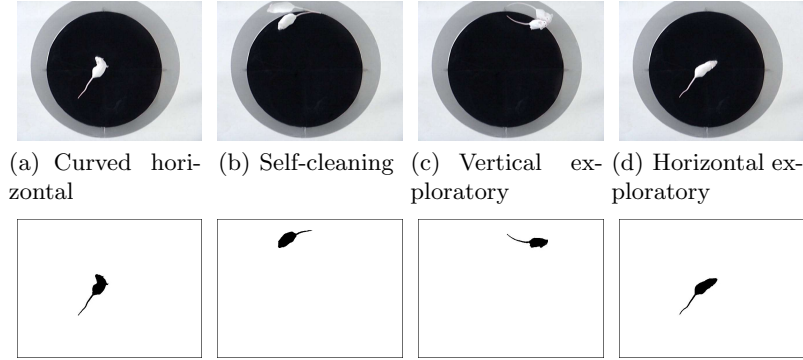

(a) Curved hori- (b) Self-cleaning (c) Vertical   ex- (d) Horizontal ex-
zontal   ploratory   ploratory

Fig. 1. Examples of the classes in the mice behavior database and its respectively segmented images.

The training and the test routines were done using a 10-fold cross-validation approach. For comparisons, we selected 4 classifiers that are popular in the data mining community were used: Decision Tree - DT , K-Nearest Neighbor - KNN, Support Vector Machine - SVM, and Radial Basis Function Network - RBF. We considered the classifiers from WEKA [4] with default parameter settings.

3.1   Comparison of Methods

Fig. 2 provides a comparison of the shape descriptor methods for each classifier. The horizontal axis is the number of descriptors and the vertical axis is the Average Correct Classification Rate (ACCR). From the figures, two observations can be made: First, the ACCR did not significantly improve with more than 20 descriptors approximately, which indicates that most of the shape information concentrates on the first descriptors. Second, Zernike moments and Fourier descriptors achieved the best performance in the most of cases. Using KNN classifier, Zernike moments and Fourier descriptors achieved $89.90\%(\pm 6.24\%)$ and $85.93\%(\pm 5.78\%)$, respectively. With regard to the classifiers, KNN classifier seems to provide the highest ACCRs, followed by the decision tree classifier. On average (see Fig. 5(a)), the best result of $80.82\%(\pm 6.98\%)$ has been obtained by the Zernike moments with 10 descriptors.

In order to evaluate the influence of noise, each contour of the database has been corrupted with different levels of noise. Since we are interested in evaluating shape descriptors and do not evaluate algorithms for contour extract, we decided

---

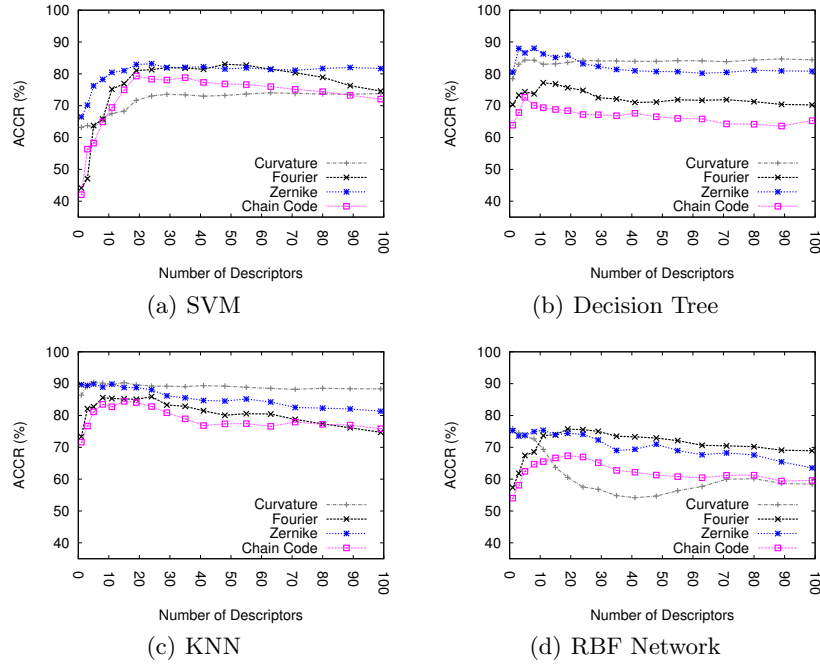[4] Weka data mining software: http://www.cs.waikato.ac.nz/ml/weka/

Fig. 2. Experimental results for each shape descriptor method using four classifiers.

to insert noise only into the contour instead of inserting it into the image. For all the points of contour, a Gaussian noise with mean 0 and standard deviation σ ranging from 0.1 to 5.0 is added to the original point, according to Equation 7. To illustrate, Fig. 3 shows examples of levels of noise.

$$
\begin{aligned}
x_\sigma &= x + N(0, \sigma) \\
y_\sigma &= y + N(0, \sigma)
\end{aligned}
\tag{7}
$$

where (x_σ, y_σ) is the noised point, (x, y) is the original point and $N(0, \sigma)$ is a random number generated from the normal distribution with mean 0 and standard deviation σ.
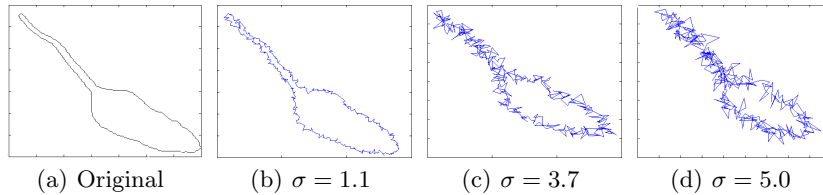

Fig. 3. Examples of level of noise.

The results for the database with noised contour are shown in Fig. 4. The horizontal axis is the level of noise σ and the vertical axis is the ACCR. As expected, the ACCR decreases as the level of noise increases. Again, Zernike moments and Fourier descriptors provided the best results. From the Fig. 5(b) which plots the average from the classifiers, Fourier descriptors' ACCR decreases from $75.04\%(\pm$ $6.63\%)$ to $70.85\%(\pm\ 6.34\%)$. On the other hand, Zernike moments' ACCR decreases from $82.10\%(\pm\ 6.75\%)$ to $66.51\%(\pm\ 2.31\%)$. These results suggest that Fourier descriptor method is more invariant to high levels of noise. Further, it is found that the KNN classifier provided the highest ACCR.
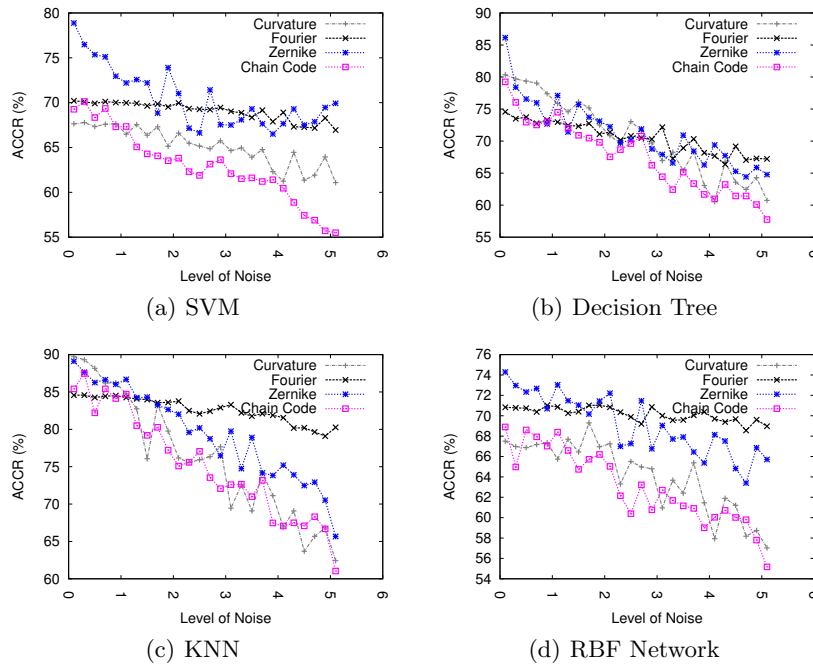

Fig. 4. Results for the employed classifiers on the noised contours.

4   Conclusion and Future works

In this paper we have presented a comparative study of shape descriptors methods. Promising results have been obtained on mice behavior images of relatively high complexity. Experimental results on this database indicate that Zernike moments and Fourier descriptors provided better performance than curvature and chain code methods. We also demonstrated how the noise levels affect shape recognition. Results for this problem suggest that Fourier descriptor method is more invariant to high levels of noise.
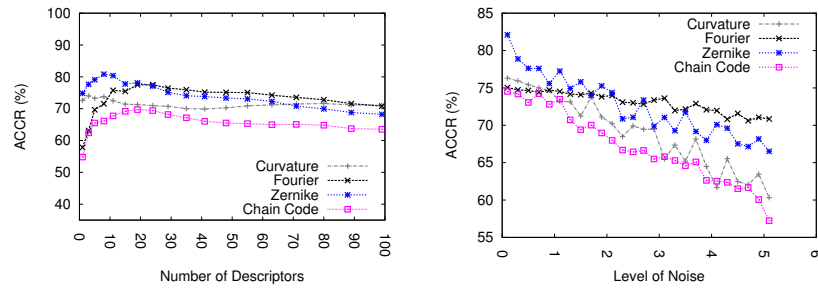
(a) Average values over four classifiers -   (b) Average values over four classifiers -
original contours.   noised contours.

Fig. 5. Average values for each shape descriptor method using four classifiers for original (a) and noised contours (b).

Regarding future work, it is proposed the analysis of other region-based methods. Moreover, further research could be focused on investigating multiscale version of the methods and the study of model selection criteria.

References

1. Luciano, Cesar, R.M.: Shape Analysis and Classification: Theory and Practice (Image Processing Series). CRC (2000)
2. Salem, A.B.M., Sewisy, A.A., Elyan, U.A.: A vertex chain code approach for image recognition. Graphics, Vision and Image Processing ICGST **05** (2005)
3. Zhang, D., Lu, G.: A comparative study of curvature scale space and fourier descriptors for shape-based image retrieval. J. Visual Commun. Image Represent. **14** (2003) 39–57
4. Zhang, D., Lu, G.: Study and evaluation of different fourier methods for image retrieval. Image Vision Comput. **23** (2005) 33–49
5. Hwang, S.K., Kim, W.Y.: A novel approach to the fast computation of zernike moments. Pattern Recognition **39** (2006) 2065–2076
6. Chong, C.W., Raveendran, P., Mukundan, R.: Translation invariants of zernike moments. Pattern Recognition (2003) 1765–1773
7. Yang, G.Y., Shu, H.Z., Toumoulin, C., Han, G.N., Luo, L.M.: Efficient legendre moment computation for grey level images. Pattern Recognition **39** (2006) 74–80
8. Morrow-Tesch, J., Dailey, J.W., JIang, H.: A video data base system for studying animal behavior. Journal of Animal Science **76** (1998) 2605–2608
9. Gonçalves, W.N., Saueia, V.A., Machado, B.B., de A. Silva, J., de Souza, K.P., Pistori, H.: Técnicas de segmentação baseadas em subtração de fundo e modelos de cores: Um estudo comparativo. XXVIII CILAMCE (2007)
10. Pistori, H., Odakura, V.V.V.A., Monteiro, J.B.O., Gonçalves, W.N., Roel, A.R., de A. Silva, J., Machado, B.B.: Mice and larvae tracking using a particle filter with an auto-adjustable observation model. Pattern Recog. Lett. **31** (2010) 337 – 346