MFSC: A NEW SHAPE DESCRIPTOR WITH ROBUSTNESS TO DEFORMATIONS

Lunshao Chai[a], *Zhen Qin*[b], *Honggang Zhang*[a], *Jun Guo*[a], *Bir Bhanu*[b]

[a]Beijing University of Posts and Telecomuunictions, Beijing, 100876, China
[b]University of California at Riverside, Riverside, CA 92521, USA

ABSTRACT

In this paper, we propose a new shape descriptor, *Multi-scale Fuzzy Shape Context (MFSC)*, highlighted by its robustness to deformations. A novel multi-scale fuzzy model is presented and applied on the widely used shape descriptor Shape Context to generate MFSC. The multi-scale fuzzy model can handle shape deformations of different scales, which makes MFSC robust to various deformations. Experiments on an articulated shape dataset demonstrate performance improvement gained by MFSC over existing methods. We also applied MFSC on a real-world application, *Content-Based Product Image Retrieval*, and the experimental results further validate its effectiveness. We make our code and experimental data publicly available for future reference.

Index Terms— Multi-scale Fuzzy Shape Context, Multi-scale Fuzzy Model, Shape Context

1. INTRODUCTION

The classical *Shape Context (SC)* [1] descriptor has played an important role in various multimedia and vision tasks. However, as a histogram-based descriptor, SC is not robust enough to shape deformations. In order to capture rich and precise information, a shape is divided into small and absolute sectors and each sector is represented by one bin in the SC histogram of the shape. In this way SC features can only precisely measure the similarity of two shapes based on the assumption that the they are sector-aligned, which is hardly the case in most applications because of shape deformations.

Targeting to this problem, many solutions following the idea of fuzzy model have been proposed. The basic idea of fuzzy model is to use one-to-many correspondence when sampled points from a shape are assigned to bins in a SC histogram, instead of using the one-to-one correspondence in crisp model. Liu et al. proposed a soft shape context descriptor in which one point is assigned to angular neighboring bins [2]. Wang et al. tried to address histogram distortion by angular blur [3]. They enlarged angular span, letting bins be overlapped in angular directions. But these two solutions only deal with one dimension of SC, namely the angular dimension. In this sense, Ayed et al. went much further by defining fuzzy rules on both angular and radial dimensions and got an improved fuzzy descriptor [4]. However, the fuzzy rules are empirically defined, whose performance may degrade when applied on new datasets.

Besides the weaknesses mentioned above, there is another critical drawback in these models: They are all designed to handle shape deformations of a certain scale, using a set of pre-defined fuzzy rules or points-to-bins assignments. But in practice the scales of deformations vary much. Thus this drawback limits the performance and validity of these models. In order to address this problem, we introduce a novel multi-scale fuzzy model and a *Multi-scale Fuzzy Shape Context (MFSC)* descriptor. The model formulates multi-scale shape deformations with multiple points-to-bins assignments and this feature makes MFSC more robust to deformations of different scales. To the best of our knowledge, it is the first multi-scale fuzzy model applied on SC (and even the first applied on multi-dimensional histogram-based descriptors).

2. MULTI-SCALE FUZZY SHAPE CONTEXT

We now describe in detail our proposed multi-scale fuzzy model and the MFSC descriptor that it generates. First we introduce the idea of fuzzy model on histogram-based descriptors. Then we present our multi-scale fuzzy model, followed by implementation details of MFSC. Finally we study the distance measure of MFSC.

2.1. Fuzzy Model

Let $\boldsymbol{x}_i$ represents a datapoint, whose domain is $\boldsymbol{\Omega}$. In histogram-based descriptors, $\boldsymbol{\Omega}$ is usually divided into a number of sub-domains $\{A_k\}$, each of which corresponds to one bin in feature histogram $\{B_k\}$. In fuzzy model, instead of using the accurate representation $\boldsymbol{x}_i$, a distribution $d(\boldsymbol{x}_i)$ is used to represent the datapoint to formulate possible distortions. Then each bin B_k in the feature histogram can be computed as

$$B_k = \sum_i \int_{A_k} d(\boldsymbol{x}_i). \tag{1}$$

However, in the above model there is a huge computational cost problem because of the many-to-many relationship between histogram bins and datapoints. In practice, usually fixed points-to-bins assignments are employed for efficiency. All datapoints falling into certain sub-domain A_k will be quantized to the center of the sub-domain $\boldsymbol{c}_k$. Then fixed assignments $a(\boldsymbol{c}_k)$ derived from $d(\boldsymbol{c}_k)$ can be used and a set of fixed values will be assigned to A_k's corresponding bin B_k and its surrounding bins. In this way the model is simplified to

$$B_k = \sum_i a(\overline{\boldsymbol{x}_i}), \tag{2}$$

where $\overline{\boldsymbol{x}_i}$ is the quantized $\boldsymbol{x}_i$.

2.2. Multi-scale Fuzzy Model

In Multi-scale Fuzzy Model, we use a set of normal distributions with covariance matrix

$$\boldsymbol{Cov}_m = \alpha_m \boldsymbol{I}, \quad m = 1, 2, ..., M \tag{3}$$

to formulate the distortions of datapoints. Following aforementioned fuzzy model, we can obtain a set of points-to-bins assignments and then a set of fuzzy histograms. These histograms compose a rich description which can formulate multi-scale datapoint distortions.

A straightforward way to build the fuzzy histograms is fuzzifying the original crisp histogram using fuzzy models with different αs, which we term *original level-based fuzzifying approach*. Nevertheless, when α gets bigger, the normal distribution gets more gradual, and fixed assignments are no longer good approximations. So we propose a *previous level-based fuzzifying approach*. We build the fuzzy histograms in an iterative manner with a constant small α, and in each iteration we fuzzify the output fuzzy histogram of the previous iteration. This approximately equals fuzzifying the original crisp feature histogram using a gradually increasing α and a wider assignment range.

2.3. Multi-scale Fuzzy Shape Context

We apply multi-scale fuzzy model on SC and term the new descriptor *Multi-scale Fuzzy Shape Context (MFSC)*. Since SC is a two-dimensional descriptor, we use a two dimensional normal distributions with covariance matrix

$$\boldsymbol{Cov} = \alpha \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{4}$$

and iteratively construct fuzzy histograms. The first fuzzy histogram $\boldsymbol{h_1}$ is generated based on the crisp SC histogram $\boldsymbol{h_0}$. Then the second fuzzy histogram $\boldsymbol{h_2}$ is generated based on the $\boldsymbol{h_1}$, and so on until most deformations can be formulated in these fuzzy histograms. Thus there are two parameters in MFSC: α in covariance matrix and the number of fuzzy scales M.

2.4. Distance Measure of MFSC

Given MFSC features $\boldsymbol{H}$ and $\boldsymbol{G}$ from two shapes, we describe them by histogram arrays, namely $\boldsymbol{h_0 h_1 h_2 \ldots h_M}$ for $\boldsymbol{H}$, and $\boldsymbol{g_0 g_1 g_2 \ldots g_M}$ for $\boldsymbol{G}$. The distance of $\boldsymbol{H}$ and $\boldsymbol{G}$ is measured as

$$Dis_{MFSC}(\boldsymbol{H}, \boldsymbol{G}) = \sum_{i=0}^{M} Dis_{SC}(\boldsymbol{h_i}, \boldsymbol{g_i}), \tag{5}$$

where $Dis_{SC}(\boldsymbol{h_i}, \boldsymbol{g_i})$ can be any traditional distance measures of SC, such as χ^2 statistic and L_2 norm. We also design and study other distance measures. Besides the above one, another one that performs well is defined as

$$Dis_{MFSC}(\boldsymbol{H}, \boldsymbol{G}) = \sum_{i=0}^{M} (Dis_{SC}(\boldsymbol{h_s}, \boldsymbol{g_i}) + Dis_{SC}(\boldsymbol{h_i}, \boldsymbol{g_s})), \tag{6}$$

where s is a specific scale. However, this distance measure is much more time-consuming than the first one. A comparison between these two distance measures can be found in the experiments section.

3. EXPERIMENTS

In this section we test MFSC for two tasks. The first one is on the articulated shape database and the second one is on a real-world application, *Content-based Product Image Retrieval (CBPIR)*, both of which demonstrate the effectiveness of MFSC.[1]

3.1. Articulated shape database

The articulated shape database [5] is designed and used for testing articulation, which is an important case of deformation [5, 6].

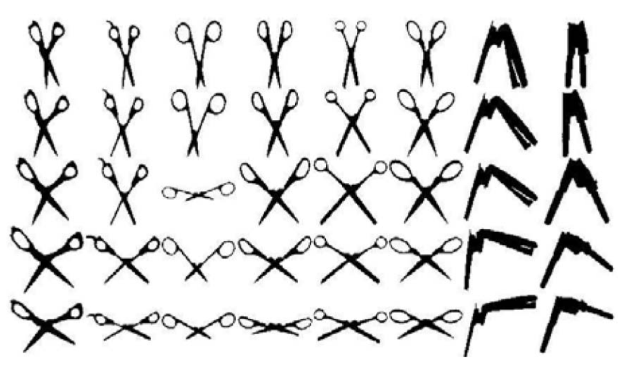

Fig. 1. Articulated shape database. Each column contains five images from the same object with different articulation.

[1] Our Matlab implementation of MFSC can be downloaded for research usage at https://dl.dropboxusercontent.com/u/57435211/icme2013.zip.

We use the following setting for all the features: 200 points sampled from every shape; 5 radial bins and 12 angular bins in SC features; and α in the covariance matrix is 0.5. MFSC scales are indicated in each set of comparison respectively. The recognition result is evaluated as the following: For each image, the 4 most similar matches are found from other images in the dataset. Then the retrieval results are summarized as the numbers of the 1st, 2nd, 3rd and 4th most similar matches that come from the correct object.

The experiment consists of three sets of comparisons. First MFSC is compared with the original Shape Context [1] and one-dimensional fuzzy Shape Context [4]. The following two sets of comparisons are on two technical aspects of our model respectively: the fuzzifying approach and distance measure.

Table 1. Retrieval results on the articulated dataset with Shape Context, fuzzy Shape Context and MFSC. MFSC scales are marked in brackets. The reported time is all the time spent, including feature extraction time, retrieval time and ranking time.

Feature	1st	2nd	3rd	4th	Time
SC	20	10	9	5	385.91s
fuzzy SC	23	13	10	7	398.21s
MFSC (1)	25	22	12	8	409.36s
MFSC (2)	29	18	12	9	425.15s
MFSC (3)	31	17	12	9	441.18s
MFSC (4)	31	20	12	7	449.30s

Table 2. Retrieval results on the articulated dataset with different fuzzifying approaches. MFSC scales are marked in brackets.

fuzz. approach	1st	2nd	3rd	4th	Time
prev.-based (2)	29	18	12	9	425.15s
orig.-based (2)	28	19	12	8	434.17s
prev.-based (3)	31	17	12	9	441.18s
orig.-based (3)	30	18	12	8	481.20s

Table 3. Retrieval results on the articulated dataset with different distance measures.

Measure	1st	2nd	3rd	4th	Time
Meas.1	31	17	12	9	441.18s
Meas.2(s=0)	30	21	10	7	1344.08s
Meas.2(s=1)	31	21	10	8	1393.14s

Table 1 shows that MFSC overperforms SC and fuzzy SC while spends a similar time. It is noteworthy that more scales can improve the performance of MFSC, but MFSC with 3 and above scales have similar performance, which indicates that 3 fuzzification scales are enough to formulate deformations in this dataset.

Table 2 shows comparison between the proposed previous level-based fuzzifying approach and the original level-based fuzzifying approach. And the results of different distance measures are demonstrated in Table 3. *Meas. 1* denotes the distance measure in *Equation (5)* and *Meas. 2* denotes the distance measure in *Equation (6)*. From the table we can see that *Meas. 2* with an appropriate s has a slightly better performance but is much more time-consuming.

3.2. Content-based Product Image Retrieval

To test MFSC under various types of deformation, we applied MFSC on *Content-based Product Image Retrieval(CBPIR)* task. *CBPIR* is an emerging application-oriented field of Content-based Image Retrieval with the prevalence of E-Commerce sites such as Amazon and eBay. Shape features can be an important cue in this task because the background of product images are simple and the shapes of the products can be easily extracted.

We conduct the experiment on two datasets, the *Product Image Categorization (PI100)* dataset [7] and the *CPImage10* dataset [8, 9]. We choose 10 categories from PI100 (see example images in Fig. 2 upper two rows). There are 100 images in each category's target gallery and 20 query images for each category, which are not included in the target gallery. CPImage10 is a subset of the CPImage dataset (see example images in Fig. 2 lower two rows). There are 10 categories with 100 images in each category, and we randomly choose 20 of them as query images and perform retrieval in the whole dataset.

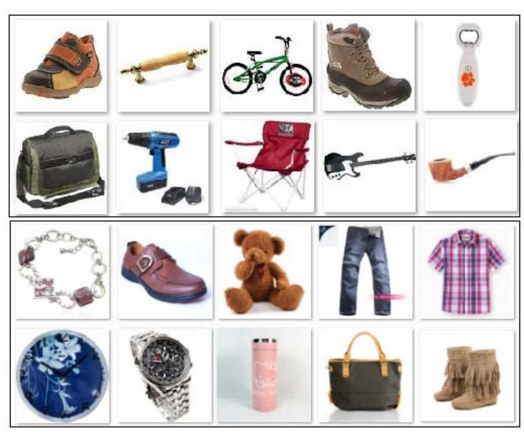

Fig. 2. Example images from the selected PI100 and CPImage databases

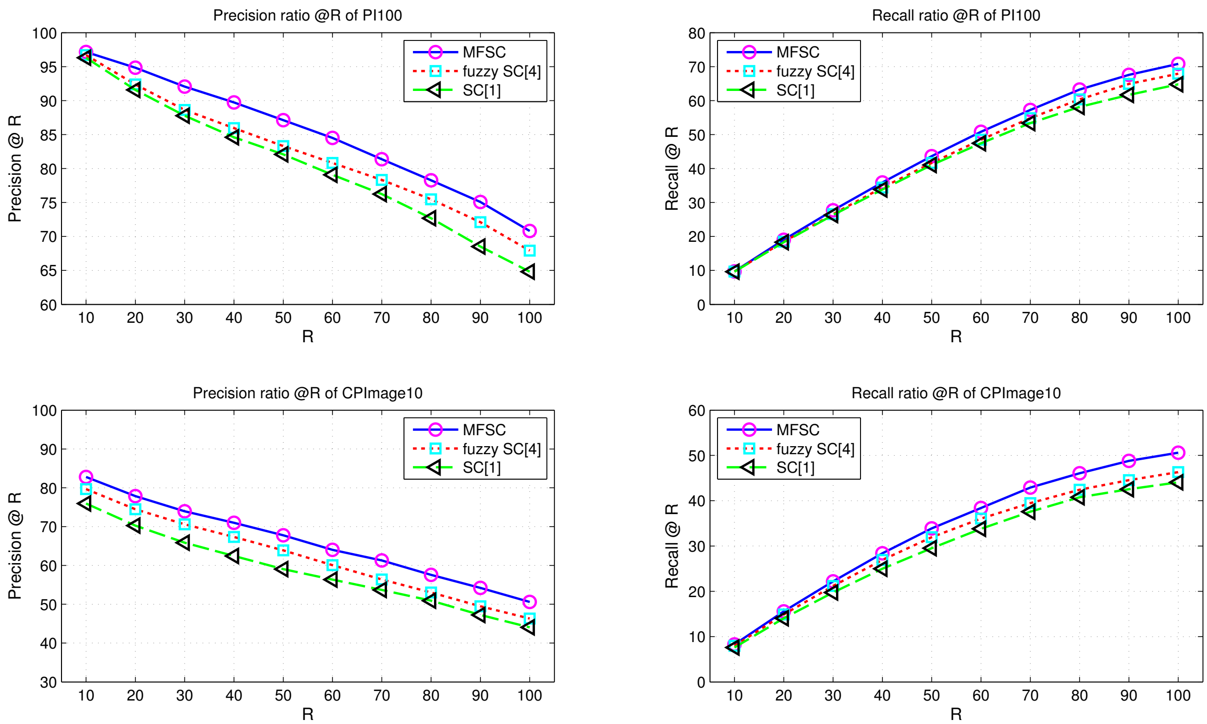

Fig. 3. Retrieval results on PI100 and CPImage10 datasets

In this experiment, 100 points are sampled from every shape for efficiency. There are 5 radial bins 12 and angular bins in SC features. In MFSC, the number of scales is 3 and α is 0.5. Precision and Recall ratiois used to evaluate the retrieval result. Fig. 3 shows the precision and recall at the first R retrieval results.

4. CONCLUSION AND FUTURE WORK

In this paper we propose a Shape Context-based descriptor, *Multi-level Fuzzy Shape Context (MFSC)*, which is generated from a new multi-scale fuzzy model. The model formulates different scales of shape deformation, which makes MFSC more robust and effective. Experiments on an articulated shape dataset and a real-world application CBPIR validate our model and demonstrate the effectiveness of MFSC.

We are interested in generalizing our model to other widely used HBDs, such as SIFT and HOG. We also plan to go further with MFSC, including finding better fuzzifying approach, and exploring the possibility of using Bag-of-Feature-like model to speed up the matching.

5. ACKNOWLEDGEMENT

This work was partially supported by National Natural Science Foundation of China under Grant No.61273217, 61175011 and 61171193, the 111 project under Grant No.B08004, and the Fundamental Research Funds for the Central Universities.

6. REFERENCES

[1] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape context," in *TPAMI*, 2002.

[2] D. Liu and T. Chen, "Soft shape context for iterative closest point registration," in *ICIP*, 2004.

[3] L. Wang, J. Shi, G. Song, and I. Shen, "Object detection combining recognition and segmentation," in *ACCV*, 2007.

[4] A. B. Ayed, M. Kardouchi, and S. Selouani, "Rotation invariant fuzzy shape contexts based on eigenshapes and fourier transforms for efficient radiological image retrieval," in *ICMCS*, 2012.

[5] H. Ling and D. W. Jacobs, "Using the inner-distance for classification of articulated shapes," in *CVPR*, 2005.

[6] H. Ling and K. Okad, "Diffusion distance for histogram comparison," in *CVPR*, 2006.

[7] X. Xie, L. Lu, M. Jia, H. Li, F. Seide, and W.g Ma, "Mobile search with multimodal queries," in *Proc. of the IEEE*, 2008.

[8] L. Chai, H. Zhang, Z. Qin, J. Yu, , and Y. Qi, "Multi-feature content-based product image retrieval based on region of main object," in *ICICS*, 2011.

[9] L. Chai, Z. Qin, H. Zhang, J. Guo, and C. Shelton, "Re-ranking using compression-based distance measure for content-based commercial product image retrieval," in *ICIP*, 2012.