LETTER

Fast Shape Matching Using Statistical Features of Shape Contexts

Moon-Jai LIM[†], **Chan-Hee HAN**[†], **Si-Woong LEE**[†], ***Nonmembers***, ***and*** **Yun-Ho KO**[††a)], ***Member***

SUMMARY A novel fast algorithm for shape matching using statistical features of shape contexts is presented. By pruning the candidate shapes using the moment-based statistical features of shape contexts, the required number of matching processes is dramatically reduced with negligible performance degradation. Experimental results demonstrate that the proposed algorithm reduces the pruning time up to $1/(r \cdot n)$ compared with the conventional RSC algorithm while maintaining a similar or better performance, where n is the number of sampled points of a shape and r is the number of randomly selected representative shape contexts for the query shape.
key words: *shape context, generalized shape context, shape matching, fast algorithm*

1. Introduction

Object recognition has been one of the main issues in computer vision and image processing research. Diverse algorithms for object recognition have been studied, and a lot of shape-based approaches have recently been reported [1]–[5]. Belongie et al. introduced the *shape context* descriptor, which characterizes a particular point location on the shape [1]. The shape matching using shape context shows superior performance in matching accuracy as reported in [3]. However, that approach needs heavy computations since n^2 times of matching operations are required to obtain the similarity between any two shapes each of which is sampled into n points. In addition, the number of matching operations and the resultant computation time increase dramatically as the cardinality of the database of reference shapes increases. Therefore, a fast matching algorithm which enables fast searching of candidate shapes of high similarity with the query shape is requisite.

In [4], a *representative shape contexts* (RSCs) method, which uses a subset of shape contexts of a query shape for quick pruning, was proposed. Although it can achieve a quick construction of a shortlist of candidate shapes, it may suffer performance degradation since RSCs do not fully reflect the entire shape characteristics. As a solution to this problem, an alternative fast pruning method which uses the moment-based statistical features of shape contexts is proposed in this paper. Instead of using an individual shape context, the proposed method utilizes moments of shape contexts such as a mean shape context or a variance of shape contexts for each shape in the pruning stage. Since the moment of shape contexts closely reflects the statistical properties of the entire shape, more efficient pruning results can be obtained as well as a reduced pruning time.

Manuscript received February 14, 2011.
Manuscript revised June 9, 2011.
[†]The authors are with Hanbat National University, Daejon, Korea.
[††]The author is with Chungnam National University, Daejon, Korea (corresponding author).
a) E-mail: koyh@cnu.ac.kr
DOI: 10.1587/transinf.E94.D.2056

2. Background

2.1 Shape Context

For a shape whose boundary is sampled into n points, a shape context of a point p_i is defined as a histogram of the relative polar coordinates of all other points on the shape as shown in Fig. 1:

$$H_i(k) = \#\{p \neq p_i : (p - p_i) \in bin(k)\} \tag{1}$$

To compute the histogram in Eq. (1), the uniform bin in log-polar space is used [1]. For n points on the shape, a total of n shape contexts is generated. In order to compute the similarity between two shape contexts of a point p_i on the first shape and a point q_j on the second shape, it is natural to use χ^2 test statistics:

$$C_{ij} \equiv C\left(p_i, q_j\right) = \frac{1}{2}\sum_{k=1}^{d}\frac{\left[h_i(k) - h_j(k)\right]^2}{h_i(k) + h_j(k)} \tag{2}$$

where d is the number of bins in a shape context, and $h_i(k)$ and $h_j(k)$ denote the k-bin normalized histogram of p_i and q_j, respectively.

As an extension of shape contexts, *generalized shape contexts* are proposed based on oriented edges [4]. The descriptor for a point p_i is the histogram h_i that is the sum of the tangent vectors for all points falling in the bin:

$$\hat{h}_i(k) = \sum_{q_j \in Q} t_j, \text{where } Q = \left\{q_j \neq p_i, (q_j - p_i) \in \text{bin}(k)\right\} \tag{3}$$

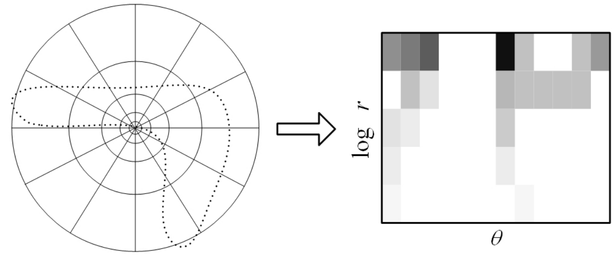

Fig. 1 Shape context of a point on a shape [1].

When the descriptors for two points are compared, this d-bin histogram is converted to a $2d$-dimensional vector $\hat{v}_i$, and compared using the L^2 norm:

$$\begin{aligned} \hat{v}_i &= \left\lfloor \hat{h}_i^x(1), \hat{h}_i^y(1), \hat{h}_i^x(2), \hat{h}_i^y(2), \cdots, \hat{h}_i^x(d), \hat{h}_i^y(d) \right\rfloor \\ d_{GSC}\left(\hat{h}_i, \hat{h}_j\right) &= \left\| \hat{v}_i - \hat{v}_j \right\| \end{aligned} \tag{4}$$

where $\hat{h}_i^x(k)$ and $\hat{h}_i^y(k)$ are the x and y components of $\hat{h}_i(k)$, respectively.

2.2 Representative Shape Contexts

The RSCs method is a fast matching framework that uses a two stage approach to object recognition: fast pruning followed by detailed matching. Given a large set of known shapes, the pruning stage quickly constructs a shortlist of candidate shapes that includes the best matching shape [4].

For each of the known shapes S^i, a large number n (about 100) of shape contexts $\{SC_j^i : j = 1, 2, \cdots, n\}$ are pre-computed. However, for the query shape, only a small number $r (\ll n)$ of randomly selected representative shape contexts are computed. Then, the distance between a query shape Q and a reference shape S^i is computed as an averaged distance of the best k RSCs that have the smallest distances:

$$\begin{aligned} d_S\left(Q, S^i\right) &= \frac{1}{k} \sum_{u \in G_i} \frac{d_{GSC}\left(SC_u^Q, SC_{m(u)}^i\right)}{N_u} \\ \text{where } m(u) &= \arg\min_j d_{GSC}\left(SC_u^Q, SC_j^i\right) \end{aligned} \tag{5}$$

where G_i denotes a set of indices of the k best RSCs and N_u is a normalizing factor that measures how discriminative the representative shape context SC_u^Q is. The shortlist is determined by sorting these distances.

The RSCs method reduces the computation time by sub-sampling shape contexts of the query shape. Thus, it cannot avoid the *aliasing* problem basically. In other words, if the RSCs are under-sampled to reduce the computational load, they are apt to fail in representing the characteristics of the query shape. On the contrary, if the size of RSCs is increased, the constraint of fast pruning becomes unsatisfactory.

3. Proposed Method

In object recognition based on shape matching, the matching feature should capture the information from the entire shape. Therefore, we adopted moment-based statistical features to fully reflect all the shape contexts in the pruning stage.

Each shape context for n sampled points can be considered as a d-dimensional vector, where d is the number of bins in a shape context:

$$\mathbf{h}_i = [h_i(1), h_i(2), \cdots, h_i(d)] \tag{6}$$

where $h_i(j)$ is the histogram value of the jth bin in the shape context of a point p_i. Then, for a shape of n sample points, a d-dimensional vector space that consists of n vectors is composed. Each vector in the vector space represents an individual shape context of the given shape. From the vector space, we compute three kinds of moment-based shape features for each shape:

- mean vector

$$\mathbf{v} \equiv \overline{\mathbf{h}} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{h}_i \tag{7}$$

- variance vector

$$\begin{aligned} \mathbf{v} &\equiv \sigma_h^2 = [s_{11}, s_{22}, \cdots s_{dd}] \\ \mathbf{S} &= \frac{1}{n-1} \sum_{i=1}^{n} \left(\mathbf{h}_i - \overline{\mathbf{h}}\right)^T \left(\mathbf{h}_i - \overline{\mathbf{h}}\right) \end{aligned} \tag{8}$$

- mean-std vector

$$\mathbf{v} \equiv \left[\overline{\mathbf{h}}, \sigma_h\right] \tag{9}$$

Note that the variance vector is composed of diagonal elements of an unbiased covariance matrix of shape context vectors. It also be noted that the dimension of the mean-std vector is $2d$, while that of the others is d. Using one of the matching features, the matching cost between the query shape Q and a reference shape S^i is computed as the Euclidean distance of their feature vectors:

$$d\left(Q, S^i\right) = \left\| \mathbf{v}_Q - \mathbf{v}_{S^i} \right\| \tag{10}$$

The shortlist is determined by sorting these distances. The proposed method has merits in two aspects:

- Since similar shapes have similar statistics in their shape contexts, statistical features gather the general information of the entire shape and can be a good discriminator in the pruning stage.
- The proposed method needs only one matching operation to compute the matching cost between two shapes, while $(r \cdot n)$ times of matching operations are needed in the RSCs method. This enables a dramatic reduction of computation time. Table 1 shows a rough comparison of the required matching operations of the three methods where \$ is the set of all reference shapes. For example, in the case of $n = 100$, $r = 16$, $|\$| = 250$, and *length(shortlist)* $= 5$, the proposed method requires 50,250 matching operations, while the full search and the RSCs method requires 2,500,000 and 450,000 matching operations, respectively.

Table 1 Comparison of numbers of matching operations.

Matching method	# of matching operations
Full search	$n \times n \times \|\$\|$
RSCs	$r \times n \times \|\$\| + n \times n \times$ *length(shortlist)*
Proposed	$\|\$\| + n \times n \times$ *length(shortlist)*

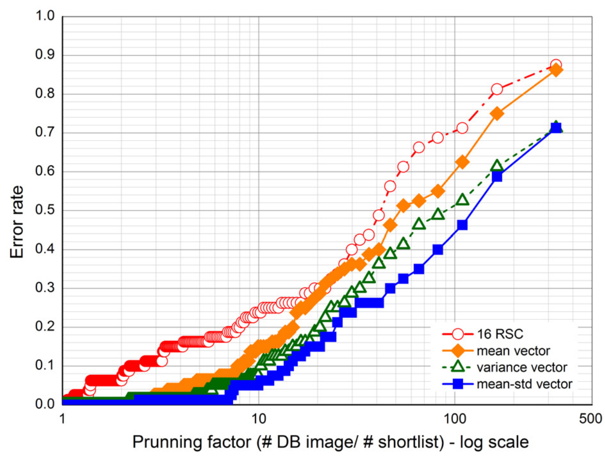

Fig. 2 Error rate vs. pruning factor on ETH-80 dataset.

Fig. 3 Shortlists for the ETH-80 dataset using the proposed method.

4. Experimental Results

ETH-80 database was used as a test set in our experiment [6]. The database consists of 80 unique objects, each from one of 8 classes. Each object is represented by 41 views spaced evenly over the upper viewing hemisphere. We used 328 images (8×41) by selecting one object from each of 8 categories as database shapes, and used the images of the remaining 9 objects from each categories as query shapes. For the RSCs method, pruning was done using 16 representative shape contexts, and 12 best RSCs were used to compute the matching cost. For the proposed method, all three kinds of feature vectors were used in the experiment.

Figure 2 shows the graph of the error rates vs. the pruning factor. Error is counted when the shortlist obtained by the pruning method doesn't include the best shape obtained by the full search method. The pruning factor is defined as the ratio of the number of database images and the size of shortlist. The proposed three methods show the superior performances compared to the RSCs method. Note that, among the three kinds of statistical feature vectors, the mean-std vector reveals the best results. This result is somewhat natural since the mean-std vector has two times of discriminating feature elements compared to the others.

Figure 3 shows an example of shortlists for several queries obtained by the proposed method. The first column is the query object and the remaining 5 columns show the closest matches. Some errors including images of other objects in the shortlist are observed. However, those matching errors in the shortlist are not a problem in this scheme, because a detailed matching procedure to find the best matching will follow.

5. Conclusion

A fast shape matching algorithm based on the statistical features of shape contexts is presented. By adopting the statistical features of the entire shape contexts to the pruning process, more reliable and computation efficient shape matching could be achieved.

Acknowledgements

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2009-0076943).

References

[1] S. Belongie, J. Malik, and J. Puzicha, "Shape context: A new descriptor for shape matching and object recognition," Advances in Neural Information Processing Systems 13: Proc. 2000 Conf., pp.831–837, 2000.
[2] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," IEEE Trans. Pattern Anal. Mach. Intell., vol.24, no.4, pp.509–522, April 2002.
[3] G. Mori, S. Belongie, and J. Malik, "Shape contexts enable efficient retrieval of similar shapes," Proc. IEEE Conf. Computer Vision and Pattern Recognition, vol.1, pp.723–730, 2001.
[4] G. Mori, S. Belongie, and J. Malik, "Efficient shape matching using shape contexts," IEEE Trans. Pattern Anal. Mach. Intell., vol.27, no.11, pp.1832–1837, Nov. 2005.
[5] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," IEEE Trans. Pattern Anal. Mach. Intell., vol.27, no.10, pp.1615–1630, Oct. 2005.
[6] http://www.mis.tu-darmstadt.de/Research/Projects/eth-80.html