数理解析研究所講究録
第 1744 巻 2011 年 185-188

Image Recognition and Retrieval by Using Distance Information

Jinhee Chun* Natsuda Kaothanthong* Takeshi Tokuyama*

Abstract

Image retrieval system is a system for retrieving images from a large database of digital images such as searching images on the Web. In this paper proposes an image retrieval system to perform shape-based image query by using distance description refined by an exact geometric matching method.

In geometric pattern recognition and information retrieval, shape descriptions based on geometric invariants play an important roles to differentiate one shape to the others. Distributional pairwise distances is one of such description methods that is applied in image recognition [5, 7, 9]. However, it is known that the set of pairwise distances is not enough to describe the differeces of one shape to the others. Though the pairwise distance is unable to uniquely describe the object, efficiency in pattern matching by using distance distribution is an attractive feature in practical use. Suppose that we have a database contaning a set of N shapes, and we would like to find all shapes that resemble to a given query data Q. Then, the shape description using distance can find the set $X(Q)$ of shapes with the same shape description as Q in O(log N) time or O(1) average time.

Our method needs $O(\mathrm{K}t+|X(Q)|)$ time, where t is the time complexity for pattern matching and K is the number of candidates retrieved, which improves naive $\mathrm{O}(Nt)$ time using solely the geometric matching method without losing the output quality.

1 Introduction

Image recognition has been a topic of active research in image search and pattern recognition for several decades. Content-base image retrival(CBIR) has become an active topic of image search research fild over the last few year[3].There is a growing interest in finding images in large collections or from a remote database [8]. A content-based image query is a system that searches similar images according to the feature of the input image. Color and texture are features that are basically used. In this paper, we use object's shape boundary as a query to search for similar images. The distribution of pairwise distances among boundary points is a basic method used for describing the shape. It can also be applied to a data structure such as hash table in order to support efficient query. However, there is a pairwise distance example that the original points location is not able to uniquely recofigure as shown in Figure 1. In addition, Skiena *et al.*[6] showed counter examples for the one-dimensional case.

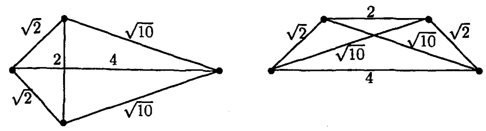


Figure 1: A counter example of 4 points set

From such ambiguity of the distance distribution, we propose an image retrieval system which integrates geometric pattern matching in order to measure the similarity of the two shapes boundary. In order to avoid comparing the query's shape boundary to every object's shape boundary in the data set, our propose system consists of two stages as shown in Figure 2. *Screening Process* is used for retrieving a set of candidates according to the distance distribution; and *Refining Process* is used for computing the similarity of the two boundary point sets by applying a geometric pattern matching.

*Graduate School of Information Science, Tohoku University, (jinhee,natsuda,tokuyama)@dais.is.tohoku.ac.jp.

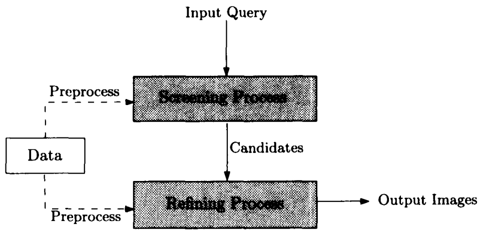


Figure 2: System Outline

2 Propose Method

We propose a two stages algorithm for retrieving a set of images according to the query boundary points. The algorithm consists of a screening process and a refining process. In the screening process, the set of similar shapes is retrieved from the hash table according to the shape signature which are computed from the distance distribution. Since the hash table is used in the screening process, the shape signature for each image in the database is preprocessed. The similarity of the query's shape boundary and the boundary from the result obtained from the screening process is computed by applying the alignment scheme for the geometric pattern matching which is computed in the refining process.

2.1 Preprocessing

In order to obtain the distance multiplicity vector, $n(n-1)/2$ pairwise distances among the given boundary point P where $P = p_1, ..., p_n$ and $|P| = n$ is computed. Then the distance is counted and stored according to the length interval of k bins which are equally divided. The distance multiplicity vector is denoted by $M(P) = (M_p(1), \ldots, M_p(k))$. The ***shape signature***, which represents a key in the dictionary, is denoted by $\mathrm{W}(P) = \{\mathrm{w}_1, ..., \mathrm{w}_k\}$ where

$$\mathrm{w}_i = \lfloor \log(M_p(i) \rfloor, i = 1, \ldots, k$$

While computing $M(P)$, the set of c sample edges $S = \{s_1 s'_1, ..., s_c s'_c\}$ are randomly selected and stored with corresponse to the length interval of each bin.

The time complexity for computing the shape signature for all images in the data set is $O(Nn^2)$, where N is the number of images.

2.2 Screening Process

The shape signature of the query boundary $\mathrm{W}(Q)$ is computed from the distance multiplicity vector $M(Q)$. $\mathrm{W}(Q)$ is used for retrieving the distance multiplicity vector of images which the key is $\mathrm{W}(Q)$. Then, the set of elements which is denoted by $X(Q) = \{x_1, \ldots, x_i\}$, $|X(Q)| = i$ is retrieved from the hash table. $M(x_1), \ldots, M(x_i)$ represents the distance multiplicity vector of the elements in $X(Q)$. To select the candidates set $Y(Q)$, L_∞ distance is used to compute the similarity between $M(Q)$ and the distance multiplicity vector of each element in $X(Q)$ which is computed by

$$L_\infty = max_{1 \leq i \leq k}(|M_q(i) - M_{x_l}(i)|)$$

where $l = 1, \ldots, i$. The K^{th} elements with the shortest L_∞ distance are selected as the candidate set $Y(Q)$.

The time complexity for obtaining a set of candidates from the dictionary is $O(1)$. To compute the similarity of the query's distance multiplicity vector to each distance multiplicity from $X(Q)$ takes $O(|X(Q)|)$ time.

2.3 Refinig Process

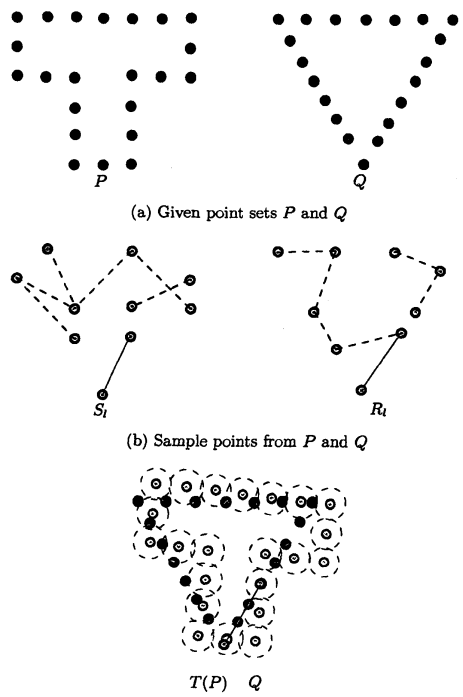
In this process, the randomized alignment scheme [1] of the geometric pattern matching method is used to measure the similarity of the two boundary point sets P and Q by counting the number of coincide points in order to remove the irrelevant images from the candidates $Y(Q)$.

First, the sample edge sets $S = \{s_1 s'_1, ..., s_c s'_c\}, S \in P$ and $R = \{r_1 r'_1, ..., r_c r'_c\}, R \in Q$ are selected according to the bin B_j where

$$j = \arg(\min_{1 \leq i \leq k} |M_p(i) - M_q(i)|)$$

The selected bin B_j is denoted by B^{sel}. Then, an edge ss' is aligned with an edge rr', in this way, the transformation $T = T[s, s'; r, r']$ is obtained. The boundary point P is transfromed according to the transformation T as shown in Figure3(c). Finally, the number of coincide points $mult_{s,r}$ is computed by counting the boundary points of Q that fall in the α radius from each point in $T(P)$ as shown in Figure3(c).

$$mult_{s,r} = \sum_{i=1}^{n} \sum_{j=1}^{m} \mathrm{coincide}(\mathrm{p_i}, \mathrm{q_j}),$$



(a) Given point sets P and Q

(b) Sample points from P and Q

(c) Coincide points between $T(P)$ and Q

Figure 3: Refining candidates algorithm

$$p_i \in T(P) \text{ and } q_j \in Q$$

$$\mathrm{coincide(p_i, q_j)} = \begin{cases} 1, \mathrm{dist(p_i, q_j)} \leq \alpha, \\ 0, \mathrm{Else} \end{cases}$$

An edge ss' is aligned to every edge in R. For all edges in S, the largest number of coincide points is the similarity measurement of the boundary points P and Q.

In an approximiate pattern matching approach, the sample points S is transformed according to $T = T[s, s'; r, r']$. The number of coincide points is counted from each point in $T(S)$.

The time complexity for computing the similarity measurement between $T(P)$ and Q of the pattern matching is $O(mn)$ where $|P| = n, |Q| = m$. However, the total time complexity can be reduced to $O(m)$ by using geometric hashing or $O(m \log n)$ time by using nearest neighbor search.

Theorem 2.1 *The pattern matching can be solved in $O((n+m)|B^{sel}|^2)$ time.*

Theorem 2.2 *The approximate pattern matching can be solved in $O(m|B^{sel}|^2)$ time. If we take a constant number of sample edges, the time complexity is reduced to $O(m|B^{sel}|)$.*

Let t be time for pattern matching, where $t = O(m|B^{sel}|)$ as shown in Theorem 2.2. If we consider a pattern matching between the two boundary point sets, $O(Nt)$ time complexity is needed to compare a query boundary point to every point set in the data set. Since the refining process is only work on $Y(Q)$ set which the size is δN, the time complexity for matching the two point sets is $O(\delta N t)$.

Theorem 2.3 *The image retrieval by using the pattern matching can be computed in $O(\delta N t)$ in two processes.*

3 Experimental Results

In this experiment, 1,134 images selected from MPEG7 Shape-1 Part B data set(http://www.imageprocessingplace.com/root_files_V3/image_databases.htm). The selected images are from 59 classes. 256 boundary points are uniformly selected from the outer boundary which is extracted from the binary image. The programming language is C#.

Table 1 shows the experimental result from the screening process and Table 2 shows the experimental result from the refining process. The F-measure denotes the test's accuracy. When comparing the result from both process, the F-measure is 13.93% increased from the screening process. Figure 4 shows examples of result images from the screening process. Figure 5 shows result from the refining process of the query in Figure 4.

Table 1: Result from retrieving images from hash table

K	δ	Precision	Recall	F-measure
$\lvert X(Q)\rvert$	0.29	6.98%	59.22%	12.48%
200	0.16	8.92%	55.23%	15.35%
150	0.12	10.19%	52.00%	17.04%
100	0.08	12.46%	47.22%	19.71%
$z-1$	0.016	28.71%	28.71%	28.71%

4 Conclusion

We propose a two stages image retrieval system which applied a geometric pattern matching. The first stage is called a screening process which is used

Table 2: Refining result from the candidates retrieving from the screening process

K	Threshold θ		
	Precision	$\tilde{\gamma}$	F-measure
$\lvert X(Q)\rvert$	29.09%	56.35%	38.37%
200	33.33%	53.73%	41.13%
150	35.89%	51.21%	42.20%
100	38.88%	47.22%	42.64%

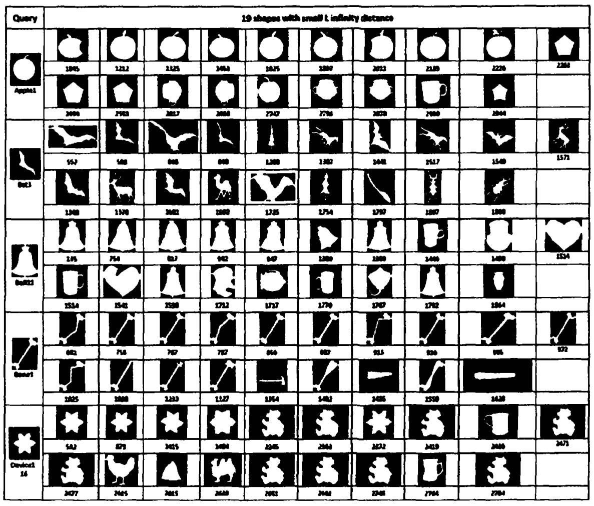


Figure 4: Example of results from the screening process

for retrieving a set of candidates from the hash table. The second stage is called a refining process which is used to remove irrelevant images from the set of candidates obtained from the screening process. The total time complexity for all process is $O(\mathrm{K}t + |X(Q)|)$. Our propose method improve the naive $O(Nt)$ time using solely the geometric matching method.

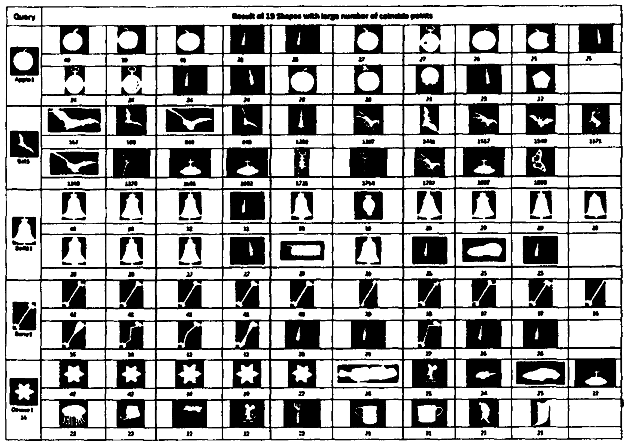


Figure 5: Examples of shapes with large number of coincide points

References

[1] T. Akutsu, H. Tamaki, and T. Tokuyama, *Distribution of distances and triangles in a point set and algorithms for computing the largest common point sets*, Discrete and Computational Geometry, 20 (1998), pp. 307–331.

[2] M. Boutin and G. Kemper, *On reconstructing n-point configurations from the distribution of distances or areas*, Advances in Applied Mathematics, 32 (2004), pp. 709 - 735.

[3] R. Datta, D. Joshi, J. Li, and J. Wang, *Image retrieval: Ideas, influences, and trends of the new age*,ACM Comput. Surv., vol.40(2)(2008),pp.5:1–5:60.

[4] L. Guth and N. H. Katz, *On the Erdos Distinct Distance Problem in the Plane.* preprint, Nov.

[5] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin, *Shape distributions*, ACM Trans. Graph., 21 (2002), pp. 807–832.

[6] S. S. Skiena, W. D. Smith, and P. Lemke, *Reconstructing sets from interpoint distances (extended abstract)*, in Proceedings of the Sixth Annual Symposium on Computational Geometry, 1990, pp. 332–339.

[7] S. Yang, *Symbol recognition via statistical integration of pixel-level constraint histograms: A new descriptor*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 27 (2005), pp. 278–281.

[8] D. Zhang and G. Lu, *Review of shape representation and description techniques*, Pattern Recognition, 37 (2004), pp. 1–19.

[9] J. Zhang and L. Wenyin, *A pixel-level statistical structural descriptor for shape measure and recognition*, International Conference on Document Analysis and Recognition, 0 (2009), pp. 386–390.