PAPER

Matching Handwritten Line Drawings with Von Mises Distributions

Katsutoshi UEAOKI[†], ***Student Member***, **Kazunori IWATA**[†a)], **Nobuo SUEMATSU**[†], ***and* Akira HAYASHI**[†], ***Members***

SUMMARY A two-dimensional shape is generally represented with line drawings or object contours in a digital image. Shapes can be divided into two types, namely ordered and unordered shapes. An ordered shape is an ordered set of points, while an unordered shape is an unordered set. As a result, each type typically uses different attributes to define the local descriptors involved in representing the local distributions of points sampled from the shape. Throughout this paper, we focus on unordered shapes. Since most local descriptors of unordered shapes are not scale-invariant, we usually make the shapes in an image data set the same size through scale normalization, before applying shape matching procedures. Shapes obtained through scale normalization are suitable for such descriptors if the original whole shapes are similar. However, they are not suitable if parts of each original shape are drawn using different scales. Thus, in this paper, we present a scale-invariant descriptor constructed by von Mises distributions to deal with such shapes. Since this descriptor has the merits of being both scale-invariant and a probability distribution, it does not require scale normalization and can employ an arbitrary measure of probability distributions in matching shape points. In experiments on shape matching and retrieval, we show the effectiveness of our descriptor, compared to several conventional descriptors.

key words: *shape descriptor, shape matching, shape retrieval, von Mises distribution, scale invariance*

1. Introduction

Shape matching means considering the correspondence between shapes, each of which is generally represented with line drawings or object contours in a digital image. Thus, shape matching is a fundamental issue in digital image processing, line drawing interpretation, and character handwriting recognition [1]–[4].

Shapes can be divided into two types, namely ordered and unordered shapes, where the former is an ordered set of points, and the latter is an unordered set. In general, ordered and unordered shapes employ different shape attributes to define their local descriptors. For instance, curvature, which is frequently used as a local descriptor [5]–[8], is applicable to ordered shapes only, because it depends on the ordering of points. On the other hand, unordered shapes make use of attributes that do not depend on the ordering of points. Local descriptors for unordered shapes are, therefore, more versatile in practice, because we do not have to assume an order of the points in the shape. In this paper, we focus on unordered shapes, and henceforth, they are referred to simply as the shapes.

Recently, several local descriptors for (unordered) shapes have been proposed [9], [10]. However, since the descriptors are not scale-invariant, as a preprocess we need to apply scale normalization to the shapes in an image data set to ensure that they are the same size. This is useful for such descriptors in matching shapes if the original whole shapes are similar in shape, but not if parts of each original shape are drawn with different scales. Accordingly, in this paper, we present a scale-invariant descriptor that can deal with shapes of this latter type and does not require scale normalization. Since this descriptor is a probability distribution described by a mixture of von Mises distributions, we can employ an arbitrary measure of probability distributions as the matching cost function of points. Using various shapes in the drawing and gesture data sets, we show that our descriptor is more effective in shape matching and retrieval than several conventional descriptors.

This paper is organized as follows. First, we explain the basics of shape matching with local descriptors in Sect. 2. Our descriptor is presented in Sect. 3, and compared with the conventional descriptors in Sect. 4. Finally, we give our conclusion in Sect. 5.

2. Matching with Local Descriptors

A shape is represented by line drawings or object contours. To reduce the computational cost of matching shapes, we usually sample some points from the shape and focus on these points. A set of points sampled from a shape is called a sample set, and an element in the set is called a sample point. Since a shape is described as a set of points, shape matching is, in fact, finding the correspondence between the sample set of one shape and that of another. Shape recognition relies on this correspondence.

2.1 Dissimilarity between Shapes

Let $\mathcal{S}_1$ and $\mathcal{S}_2$ denote shapes, and $\hat{\mathcal{S}}_1$ and $\hat{\mathcal{S}}_2$ their respective sample sets. We generally define a correspondence between the shapes by a many-to-one map $M : \hat{\mathcal{S}}_1 \to \hat{\mathcal{S}}_2$. Given a matching cost function (MCF) $C : \hat{\mathcal{S}}_1 \times \hat{\mathcal{S}}_2 \to \mathbb{R}$, the optimal correspondence in terms of C is expressed as

$$M^* \stackrel{\text{def}}{=} \operatorname*{argmin}_{M \in \mathcal{M}(\hat{\mathcal{S}}_1, \hat{\mathcal{S}}_2)} \frac{1}{|\hat{\mathcal{S}}_1|} \sum_{p \in \hat{\mathcal{S}}_1} C\left(p, M(p)\right), \qquad (1)$$

Manuscript received March 28, 2011.

Manuscript revised July 25, 2011.

[†]The authors are with the Graduate School of Information Sciences, Hiroshima City University, Hiroshima-shi, 731–3194 Japan.

a) E-mail: kiwata@hiroshima-cu.ac.jp

DOI: 10.1587/transinf.E94.D.2487

where

$$\mathcal{M}\left(\hat{\mathcal{S}}_1, \hat{\mathcal{S}}_2\right) \stackrel{\text{def}}{=} \left\{M \mid M : \hat{\mathcal{S}}_1 \to \hat{\mathcal{S}}_2\right\}. \tag{2}$$

Intuitively, the MCF in (1) quantifies how different the local distributions of sample points around p and $M(p)$ are. Using the optimal correspondence, the dissimilarity between the shapes is defined as

$$d\left(\mathcal{S}_1, \mathcal{S}_2\right) \stackrel{\text{def}}{=} \frac{1}{\left|\hat{\mathcal{S}}_1\right|} \sum_{p \in \hat{\mathcal{S}}_1} C\left(p, M^*(p)\right) + \sigma\left(\left|\hat{\mathcal{S}}_1\right|, \left|\hat{\mathcal{S}}_2\right|\right), \tag{3}$$

where the second term gives a non-negative real number, which tends to a large value as the difference between the numbers of sample points increases. Dissimilarity is fundamental to shape recognition.

2.2 Local Descriptors and their Cost Functions

A number of shape matching methods have been proposed [6]–[12]. The difference between these methods lies essentially in their choice of MCF. A local descriptor representing the local distribution of sample points around one point in the shape is involved in the MCF. In this section, we briefly introduce several local descriptors and their MCFs, which are typically used in shape matching.

(1) Distance Set

The distance set (DS) [9] expresses the local distribution around a sample point of the shape by a set of distances between that point and others. If the number of elements in the set is fixed as n, the set is called the n-DS. For any shape $\mathcal{S}$, the n-DS around $p \in \hat{\mathcal{S}}$ is defined by

$$s_n(p) \stackrel{\text{def}}{=} \left\{l_1(p), \ldots, l_n(p)\right\}, \tag{4}$$

where $l_i(p)$ denotes the Euclidean distance between p and the i-th nearest point in $\hat{\mathcal{S}}$. For all $p \in \hat{\mathcal{S}}_1$ and $q \in \hat{\mathcal{S}}_2$, the MCF of the n-DS is given by

$$C_{\mathrm{DS}}\left(p, q\right) \stackrel{\text{def}}{=} \min\left\{\frac{1}{n} \sum_{i=1}^{n} \frac{\left|l_i(p) - l_{\phi(i)}(q)\right|}{\max\left\{l_i(p), l_{\phi(i)}(q)\right\}} \,\middle|\, \phi \in \Phi_n\right\}, \tag{5}$$

where $l_i(p) \in s_n(p)$ and $l_{\phi(i)}(q) \in s_n(q)$ hold for all i, and

$$\begin{aligned}\Phi_n \stackrel{\text{def}}{=} \{\phi \mid \phi &: \{1, \ldots, n\} \to \{1, \ldots, n\},\\ &\forall i, j\ [\phi(i) = \phi(j) \Rightarrow i = j]\}.\end{aligned} \tag{6}$$

In short, the DS around a sample point quantifies the distance between that point and another. Since the DS depends on the distance, it is not scale-invariant.

Example 1 (DS): Consider the sample set $\hat{\mathcal{S}}$={(2, 1), (1, 2), (3, 4), (4, 3), (5, 1), (3, 1)} of a shape $\mathcal{S}$. Let p = (2, 1). Since the Euclidean distances between p and (1, 2), (3, 4), (4, 3), (5, 1), (3, 1) are $\sqrt{2}$, $\sqrt{10}$, $2\sqrt{2}$, 3, 1, respectively, the 3-DS around p is written as $s_3(p) = \left\{1, \sqrt{2}, 2\sqrt{2}\right\}$.

(2) Shape Context

The shape context (SC) [10] expresses the local distribution around a sample point of the shape with a histogram. For any shape $\mathcal{S}$, the SC around $p \in \hat{\mathcal{S}}$ with respect to the b-th bin is defined by

$$h_p(b) \stackrel{\text{def}}{=} \left|\left\{p' \in \hat{\mathcal{S}} \backslash \{p\} \,\middle|\, p' - p \in \mathrm{bin}(b)\right\}\right|, \tag{7}$$

where \ denotes the difference between the sets, and bin(b) means the b-th bin for $b = 1, \ldots, B$. B denotes the total number of bins. We should note that $p' - p$ is a vector quantity with magnitude and direction. For all $p \in \hat{\mathcal{S}}_1$ and $q \in \hat{\mathcal{S}}_2$, the MCF of the SC is given by

$$C_{\mathrm{SC}}\left(p, q\right) \stackrel{\text{def}}{=} \frac{1}{2} \sum_{b=1}^{B} \frac{\left(\bar{h}_p(b) - \bar{h}_q(b)\right)^2}{\bar{h}_p(b) + \bar{h}_q(b)}, \tag{8}$$

with the convention $0/0 = 0$, where

$$\bar{h}_p(b) \stackrel{\text{def}}{=} \frac{h_p(b)}{\sum_{b=1}^{B} h_p(b)}, \tag{9}$$

$$\bar{h}_q(b) \stackrel{\text{def}}{=} \frac{h_q(b)}{\sum_{b=1}^{B} h_q(b)}. \tag{10}$$

In short, the SC around a sample point depends not only on the distance between the point and another, but also on the angle between the x-axis and the vector formed by the two points. This dependence on distance implies that the SC is not scale-invariant either.

Example 2 (SC): A bin, illustrated in Fig. 1, represents a plane region partitioned by some lines passing through the origin and by circles centered on the origin (see [10] for the details). In the figure, the broken line represents the partition between bins, the dot depicts a sample point of the shape, and the number in a bin indicates the bin number. Table 1 gives the SC around p in Fig. 1.

2.3 Scale Effect and Scale Normalization

Usually the raw shape data obtained from line drawings of

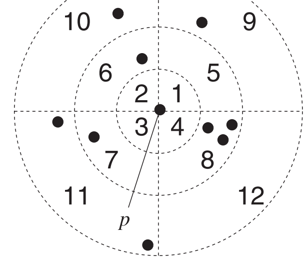
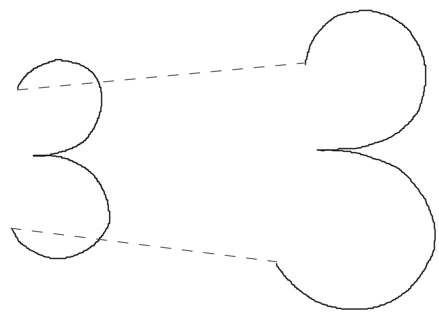


Fig. 1 Bins around a sample point p.



(a) Case in which the whole drawings are similar.

(b) Case in which corresponding parts of the drawings are similar.

Fig. 2 Line drawings of the digit "3".

Table 1 SC around p.

b	1	2	3	4	5	6	7	8	9	10	11	12
$h_p(b)$	0	0	0	0	0	1	1	3	1	1	2	0

an image data set are different in scale. Since most local descriptors, such as the DS and SC, are not scale-invariant, as a preprocess we first need to make the shapes the same size by scale normalization, and then we can sample various points of the shapes to create their sample sets. Scale normalization simply means magnifying or reducing the entire shape. It is effective for such descriptors if the original whole shapes are similar, but not so effective if parts of each original shape are represented by different scales. For example, consider the line drawings in Fig. 2, which appear to depict the same digit. Since the left and right drawings in Fig. 2 (a) have similar proportions, we can readily find the correspondence between the drawings through scale normalization. The right drawing in Fig. 2 (b) has been deformed by uniformly reducing the upper part of the left drawing and by uniformly magnifying the lower part. This means that different parts of the shape are expressed using different scales. Although these whole drawings may differ in similarity, they are the same digit in human recognition. In this case, it is difficult for non-scale-invariant descriptors to yield a correct correspondence even with scale normalization. As a result, such descriptors frequently fail to find the correct correspondence between their shapes. Accordingly, the purpose of this paper is to provide a scale-invariant descriptor that copes well with this problem.

3. A Novel Local Descriptor

We have explained that conventional descriptors cannot deal with shapes, parts of which are drawn with different scales, even when scale-normalization is applied in advance. In this section, we present a novel local descriptor to manage such shapes.

3.1 Mixture of von Mises Distributions

Neither the DS nor the SC is scale-invariant, because they depend on the distance between sample points of a shape.

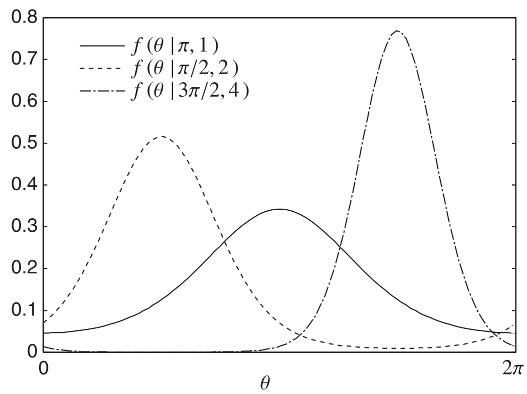


Fig. 3 vMDs with different parameters.

Our descriptor refers only to an angle of deviation, and hence it is obviously scale-invariant. The proposed descriptor expresses the local distribution around a sample point by a mixture of von Mises distributions (vMDs) [13], [14]. With θ measured in radians, let f be the vMD defined as

$$f(\theta|\theta_0, m) \overset{\text{def}}{=} \frac{\exp(m\cos(\theta-\theta_0))}{\int_0^{2\pi}\exp(m\cos\theta')\,d\theta'}, \tag{11}$$

for any θ_0 and $m > 0$. As shown in Fig. 3, the vMD is periodic with period 2π, and its parameters θ_0 and m correspond to the mean and concentration of the distribution, respectively. For any shape $\mathcal{S}$, the mixture of vMDs around $p \in \hat{\mathcal{S}}$ under a fixed m is given by

$$v_m(\theta|p) \overset{\text{def}}{=} \frac{1}{\left|\hat{\mathcal{S}}\right| - 1} \sum_{q\in\hat{\mathcal{S}}\setminus\{p\}} f\left(\theta \,\middle|\, \operatorname{ang}(q-p), m\right), \tag{12}$$

where ang $(q-p)$ denotes the angle of deviation determined by the angle between the x-axis and the vector $q-p$. Figure 4 illustrates the angle of deviation. Since the mixture of vMDs is a probability distribution, we can use an arbitrary measure of probability distributions to calculate the MCF. Now, we employ the symmetric divergence (for an example, see [15]) to define the MCF of the mixture of vMDs as follows. For all $p \in \hat{\mathcal{S}}_1$ and $q \in \hat{\mathcal{S}}_2$,

$$C_{\mathrm{MV}}(p,q) \overset{\text{def}}{=} \frac{1}{2} D\left(v_m(\cdot|p)\,\|\,v_m(\cdot|q)\right)$$

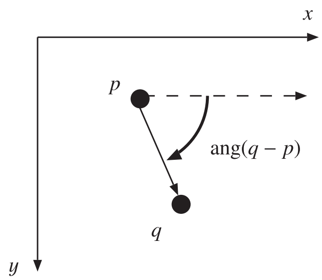


Fig. 4 Angle of deviation.

$$+\frac{1}{2}D\left(v_m\left(\cdot\,|q\right)\|v_m\left(\cdot\,|p\right)\right), \tag{13}$$

where D is the information divergence, for example,

$$D\left(v_m\left(\cdot\,|p\right)\|v_m\left(\cdot\,|q\right)\right) \stackrel{\text{def}}{=} \int_0^{2\pi} v_m\left(\theta\,|p\right)\log\frac{v_m\left(\theta\,|p\right)}{v_m\left(\theta\,|q\right)}d\theta. \tag{14}$$

Other typical measures of probability distributions should be the relative entropy [14] and the RDSP [15]. In fact, we tested these measures in the experiments in Sect. 4, and obtained the almost same results as the symmetric divergence. Hence, in this paper, we use the symmetric divergence to represent a measure involved in the MCF.

Since the mixture of vMDs is scale-invariant, if the original whole shapes are similar, then they are the same in terms of matching with the mixture of vMDs. Furthermore, since the mixture of vMDs is a local descriptor, even if parts of each original shape are drawn with different scales, it is effective for matching shapes with such similar parts.

Remark 1 (Measure Selection): The optimality of measures is strongly application dependent. So, whenever considering an application of shape matching, we need to select an appropriate measure. If there is some knowledge about ideal properties of the MCF in the application, then we might select an effective measure using the most of the properties. For example, if the MCF needs to be symmetric taking an application into account, then we have to use a symmetric measure such as the symmetric divergence. In addition to being symmetric, if the MCF needs to satisfy the triangle inequality, then the RDSP should be better. Thus, we could select an appropriate measure when ideal properties of the MCF is available.

Remark 2 (Rotation Invariance): The mixture of vMDs is not a rotation-invariant descriptor. However, we can provide a rotation-invariant matching by sliding the mixture of vMDs along the θ-axis, that is, by defining the MCF by

$$C_{\mathrm{MV}}\left(p,q\right) \stackrel{\text{def}}{=} \min\left\{\frac{1}{2}D\left(v_m\left(\cdot+\theta'\,|p\right)\|v_m\left(\cdot\,|q\right)\right)\right.$$
$$\left.+\frac{1}{2}D\left(v_m\left(\cdot\,|q\right)\|v_m\left(\cdot+\theta'\,|p\right)\right)\,\middle|\,\theta'\in[0,2\pi)\right\}, \tag{15}$$

instead, where for example, $D\left(v_m\left(\cdot+\theta'\,|p\right)\|v_m\left(\cdot\,|q\right)\right)$ means

$$D\left(v_m\left(\cdot+\theta'\,|p\right)\|v_m\left(\cdot\,|q\right)\right) \stackrel{\text{def}}{=} \int_0^{2\pi} v_m\left(\theta+\theta'\,|p\right)\log\frac{v_m\left(\theta+\theta'\,|p\right)}{v_m\left(\theta\,|q\right)}d\theta. \tag{16}$$

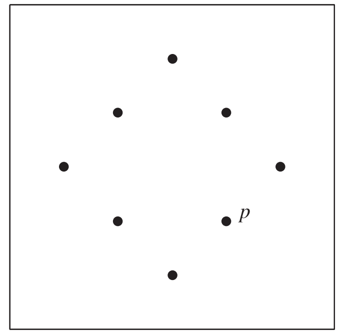


Fig. 5 Sample points of a diamond-shaped line drawing.

3.2 Implementation

Since the mixture of vMDs in (12) is a continuous function, it needs to be discretized in the implementation. Since the vMD has period 2π, we consider it over the discretized interval defined by

$$\Theta_T \stackrel{\text{def}}{=} \left\{0, \frac{2\pi}{T}, \frac{4\pi}{T}, \ldots, \frac{2(T-1)\pi}{T}\right\}. \tag{17}$$

Substituting Θ_T for the domain of the vMDs gives the following approximations of (11) and (14):

$$f\left(\theta\,|\theta_0, m\right) \approx \frac{T}{2\pi}\frac{\exp\left(m\cos\left(\theta-\theta_0\right)\right)}{\sum_{\theta'\in\Theta_T}\exp\left(m\cos\theta'\right)}, \tag{18}$$

$$D\left(v_m\left(\cdot\,|p\right)\|v_m\left(\cdot\,|q\right)\right) \approx \frac{2\pi}{T}\sum_{\theta\in\Theta_T} v_m\left(\theta\,|p\right)\log\frac{v_m\left(\theta\,|p\right)}{v_m\left(\theta\,|q\right)}. \tag{19}$$

These approximations are used in the experiments.

Example 3 (Mixture of vMDs): Figure 5 depicts the sample points of a diamond-shaped line drawing. The mixture of vMDs around sample point p depicted in the figure is obtained by making a mixture of seven different vMDs. Figure 6 plots this using $m = 10$ and $T = 60$.

3.3 Examples of Shape Matching

Using line drawings of the same class, we show that our descriptor can deal with shapes, parts of which differ in scale. The scale normalization and sampling preprocesses used here are as follows.

Scale Normalization: Scale normalization is performed by magnifying or reducing the whole shape so as to make its bounding box a fixed area of 40000 in keeping with the aspect ratio of the box. This is done before sampling.

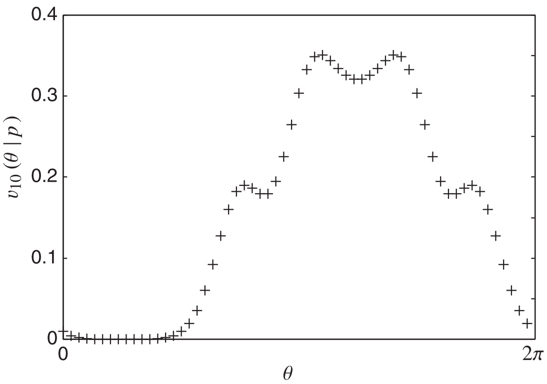
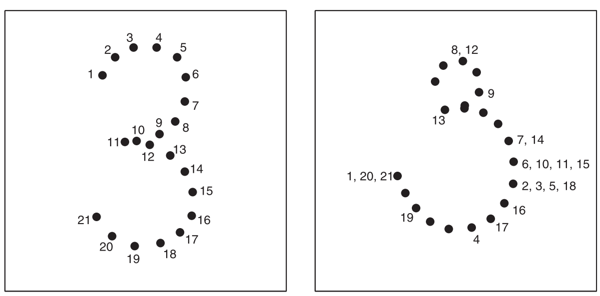
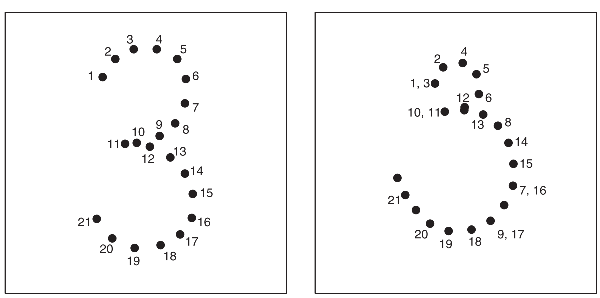


Fig. 6 Mixture of vMDs around p.



(a) Sample points of a query shape. (b) Sample points of a database shape.

Fig. 7 Correspondence produced by the DS MCF.



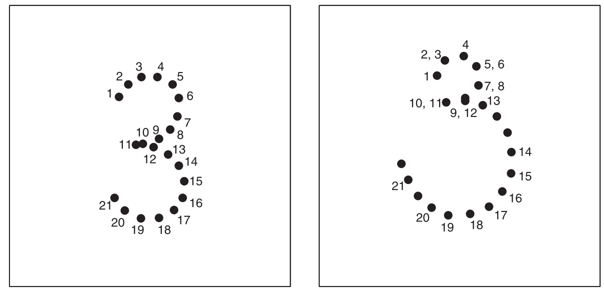
(a) Sample points of a query shape. (b) Sample points of a database shape.

Fig. 8 Correspondence produced by the SC MCF.

Sampling: Sampling implies uniformly extracting 21 points from each shape to create its sample set.

Figures 7 and 8 show examples of sample points obtained through both scale normalization and sampling of the drawings in Fig. 2 (b) with the DS and SC, respectively, applied to these. Meanwhile, since the mixture of vMDs is scale-invariant, it does not require shapes to be scale-normalized. Thus, the mixture of vMDs was applied to the sample points shown in Fig. 9, obtained through sampling only.

The left shape in each example is called a query shape. The right shape, which is a deformation of the query shape, is called a database shape. We employed the MCFs for the 10-DS, SC, and mixture of vMDs to obtain the optimal correspondence in terms of (1) between the query and database shapes. The parameters for the 10-DS, SC, and mixture of vMDs were set so as to achieve the best respective performance. Concretely, the parameters in each descriptor were determined by running a number of trials with different parameter values. In computing the mixture of vMDs, $m = 10$ and $T = 30$ were used.



(a) Sample points of a query shape. (b) Sample points of a database shape.

Fig. 9 Correspondence produced by the MCF for the mixture of vMDs.

The resulting correspondences are shown in Figs. 7, 8, and 9, respectively. The sample points of the query shape are labeled with successive numbers between 1 and 21. The numbering of a sample point on the database shape implies a correspondence with the point with the same number on the query shape. Unnumbered points on a database shape have no correspondence with points on the query shape. We see that only the mixture of vMDs yields the correct correspondence. Although the scale-normalized shapes were used differently from those to which the mixture of vMDs was applied, we confirm that the 10-DS and SC failed to find the correct correspondence (see, for example, points labeled 3, 7, and 9 in Figs. 7 (b) and 8 (b)). One reason for the poor result using the DS is that this descriptor is basically meant to be used to find a one-to-one correspondence, and not the many-to-one correspondence discussed here. These results suggest that our descriptor copes with the scale effect problem explained in Sect. 2.3.

4. Experiments

Using two shape data sets, we compared our descriptor with several conventional descriptors.

4.1 Drawing Data Set

The drawing data set available from [16] is a set of symbol shapes. Figure 10 shows all the line drawings in the data set, which consists of 100 shapes, each of which is classified into one of 25 classes. There are four shapes in each class. We employed the MCFs for the 10-DS, SC, and mixture of vMDs to calculate the dissimilarity in (3). In addition, for reference purposes, we compared all the above descriptors and the integral invariant (INI) descriptor [11]. This comparison examines the differences between matching methods for ordered and unordered shapes. Since shapes used in

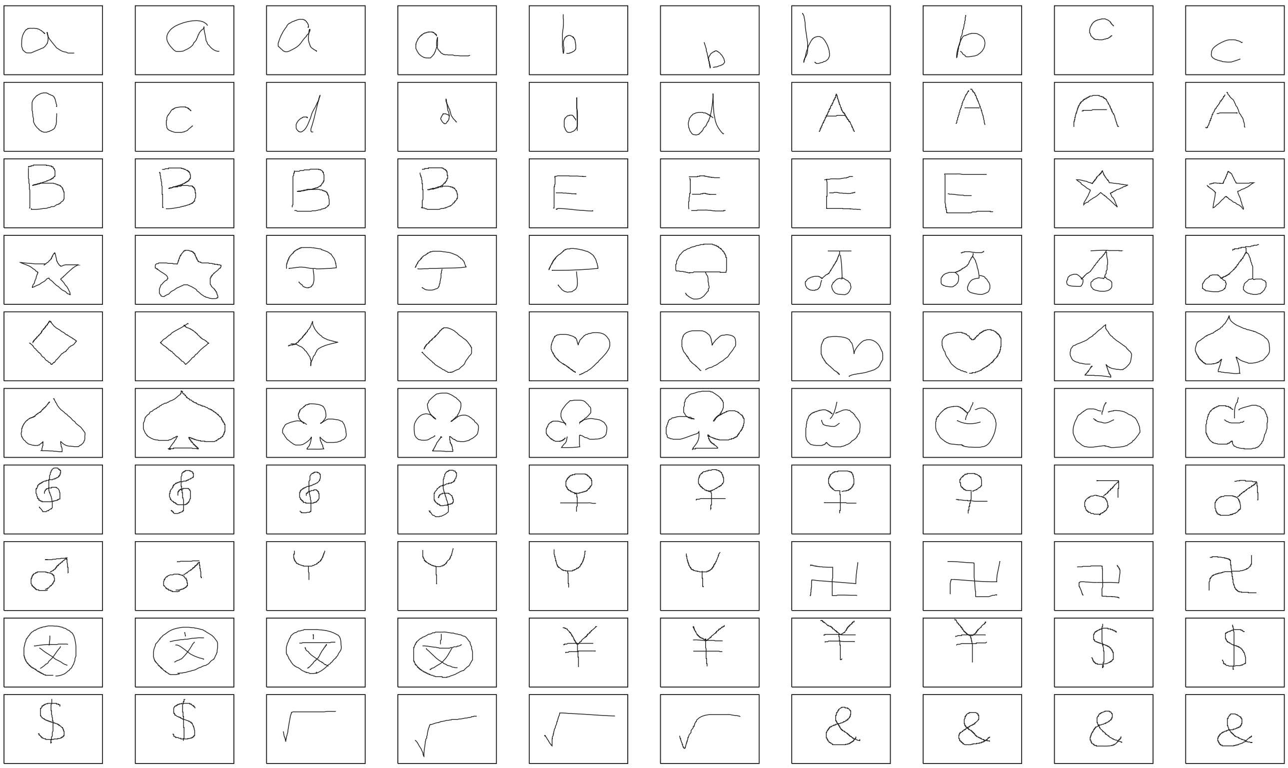


Fig. 10 Line drawing data.

the INI are required to be ordered and closed, we provided a reasonable order of the shape points using the recorded time of each point and made the shapes closed by simply connecting the first and last points of the shape points in keeping with the INI. In contrast to the matching schema based on a search with MCFs, matching with the INI is based on dynamic programming. The scale normalization and sampling performed here are the same as those discussed in Sect. 3.3. We should note that the second term in (3) is always zero in this case, because all the shapes have the same number of sample points. Moreover, the parameters for the 10-DS, SC, INI, and mixture of vMDs were tuned to achieve the best respective performance. In computing the mixture of vMDs, $m = 10$ and $T = 30$ were used.

Each shape was used in turn as the query shape, and matched against all 100 shapes. The retrieval rate was computed by counting the number of shapes in the same class that were found in the first five most similar matches (top five matches). Since the number of shapes in the same class is at most four, the total number of correct matches, when all the shapes are selected in turn as queries, is at most 400. Thus, the overall retrieval rate for the top five was calculated as the ratio of the number of actual correct matches to 400. The retrieval rate for the top ten was defined similarly. The effectiveness of the local descriptor is measured by the retrieval rate, which is frequently employed in shape retrieval evaluation (see [9], for example).

Table 2 gives the resulting retrieval rates for the four descriptors. The values for the 10-DS, SC, and INI in parentheses indicate the retrieval rates derived from the shapes

Table 2 Retrieval rates for the drawing data set.

descriptor	retrieval rate (%) for the top five	for the top ten
DS	69.25	80.5
	(66.5)	(75)
SC	93.25	95.25
	(89)	(92.5)
INI	84.5	88.25
	(79.25)	(83)
Mixture of vMDs	95.25	98.25

obtained through sampling only. These values imply that scale normalization improves shape matching. We confirm from the table that our descriptor is more effective in shape retrieval than the 10-DS, SC, and INI.

Figures 11, 12, and 13 show examples of the resulting correspondence obtained by the 10-DS, SC, and mixture of vMDs, respectively. The database shape in each example is a deformed query shape drawn by transverse stretch. Again, we see that only the mixture of vMDs produced the correct correspondence. Thus, our descriptor is relatively insensitive to a shape being stretched, because it does not depend on the distance between sample points, but exploits only the angle of deviation in (12).

4.2 Gesture Data Set

The gesture data set available from [17] consists of 17 query shapes and 980 database shapes of hand contours. Figure 14 shows the query shapes, each of which depicts a gesture.

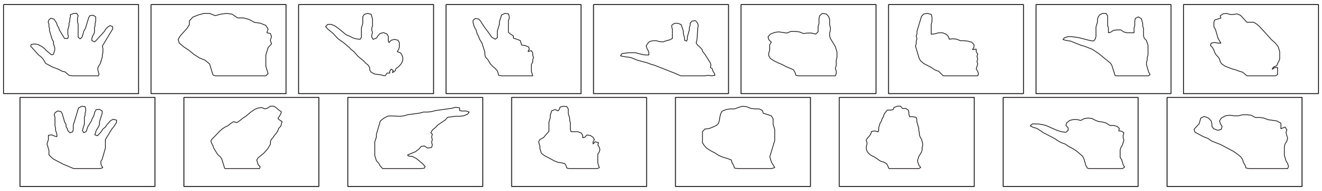
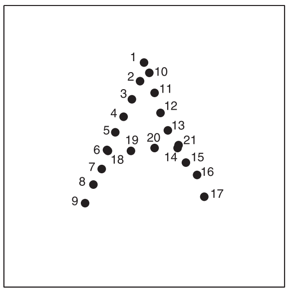
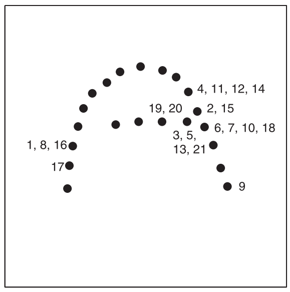


Fig. 14 Query shapes in gesture data set.

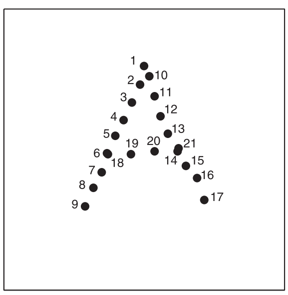


(a) Sample points of a query shape.

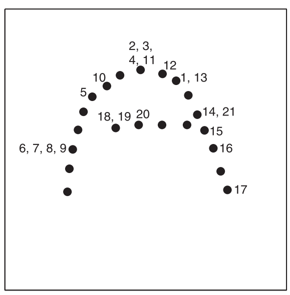


(b) Sample points of a database shape.

Fig. 11 Correspondence obtained by the DS MCF.

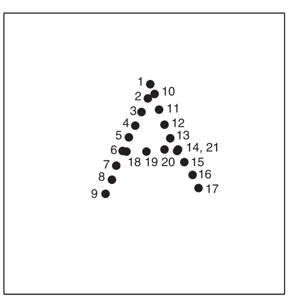


(a) Sample points of a query shape.

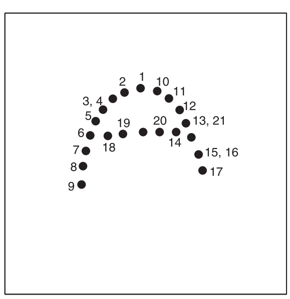


(b) Sample points of a database shape.

Fig. 12 Correspondence obtained by the SC MCF.



(a) Sample points of a query shape.



(b) Sample points of a database shape.

Fig. 13 Correspondence obtained by the MCF for the mixture of vMDs.

Each database shape is classified into one of the 17 gesture classes represented by the query shapes. In this data set, the numbers of database shapes in each of the classes are different. We employed the MCFs for the 20-DS, SC, INI, and mixture of vMDs to calculate the dissimilarity in (3). The scale normalization is the same as that discussed in Sect. 3.3, and we uniformly extracted 40 points from each shape in sampling. Moreover, the parameters for the four descriptors were tuned to achieve the best respective performance. In computing the mixture of vMDs, $m = 9$ and $T = 30$ were used.

Each query shape was matched against all 980 database shapes. The retrieval rate was computed by counting the number of database shapes of the same class that were found in the first 30 most similar matches. The overall retrieval rate for the top 30 was calculated as the ratio of the number of actual correct matches to the total number of possible correct matches. The retrieval rates for the top 40 and 50 were defined similarly. The effectiveness of the local descriptor is measured by the retrieval rates.

Table 3 Retrieval rates for gesture data set.

descriptor	retrieval rate (%)		
	for the top 30	for the top 40	for the top 50
DS	84.7	73.74	66.26
SC	95.49	88.64	83.55
INI	92.54	85.54	79.53
Mixture of vMDs	95.49	89.52	84.04

Table 3 gives the resulting retrieval rates for the four descriptors. In addition, similar correspondences to Figs. 7–9 and Figs. 11–13 were observed. Again, we confirm from the table that our descriptor is more effective in shape retrieval than the others.

5. Conclusion

We presented a scale-invariant descriptor defined as the mixture of vMDs. Using various shapes in the drawing and gesture data sets, we showed that our descriptor can handle shapes, parts of which are drawn with different scales, and it is more effective in shape matching and retrieval than several conventional descriptors.

Just as the parameter settings of the conventional descriptors, such as the DS, SC, and INI, have a great influence on their performances, so that of the mixture of vMDs has a great influence on its performance. So, establishing an effective method for finding good parameters of the mixture of vMDs is an important future work.

Acknowledgments

This work was supported in part by Grant-in-Aid 22700152 for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

[1] I.L. Dryden and K.V. Mardia, Statistical Shape Analysis, Wiley series in probability and mathematical statistics, Wiley, Chichester, UK, 1998.
[2] S. Loncaric, "A survey of shape analysis techniques," Pattern Recognit., vol.31, no.8, pp.983–1001, Aug. 1998.
[3] S. Ablameyko and T. Pridmore, Machine interpretation of line drawing images: technical drawings, maps, and diagrams, Springer, Berlin, 2000.
[4] R. Davies, C. Twining, and C. Taylor, Statistical Models of Shape: Optimisation and Evaluation, Springer, London, 2008.
[5] F. Mokhtarian and A.K. Mackworth, "A theory of multiscale, curvature-based shape representation for planar curves," IEEE Trans. Pattern Anal. Mach. Intell., vol.14, no.8, pp.789–805, Aug. 1992.
[6] T.B. Sebastian, P.N. Klein, and B.B. Kimia, "On aligning curves," IEEE Trans. Pattern Anal. Mach. Intell., vol.25, no.1, pp.116–125, Jan. 2003.
[7] M. Cui, J. Femiani, J. Hu, P. Wonka, and A. Razdan, "Curve matching for open 2D curves," Pattern Recognit. Lett., vol.30, no.1, pp.1–10, Jan. 2009.
[8] K. Iwata and A. Hayashi, "Matching between piecewise similar curve images," Collection of Technical Reports of the 12th Workshop on Information-Based Induction Sciences, Fukuoka, Japan, pp.128–135, Oct. 2009. http://ibis-workshop.org/2009/ibis09technical.pdf
[9] C. Grigorescu and N. Petkov, "Distance sets for shape filters and shape recognition," IEEE Trans. Image Process., vol.12, no.10, pp.1274–1286, Oct. 2003.
[10] S. Belongie, G. Mori, and J. Malik, "Matching with shape contexts," in Statistics and Analysis of Shapes, ed. H. Krim and A. Yezzi Jr., Modeling and Simulation in Science, Engineering and Technology, pp.81–105, Birkhäuser, Boston, 2006.
[11] S. Manay, D. Cremers, B.W. Hong, A.J. Yezzi Jr., and S. Soatto, "Integral invariants for shape matching," IEEE Trans. Pattern Anal. Mach. Intell., vol.28, no.10, pp.1602–1618, Oct. 2006.
[12] C. Xu, J. Liu, and X. Tang, "2D shape matching by contour flexibility," IEEE Trans. Pattern Anal. Mach. Intell., vol.31, no.1, pp.180–186, Jan. 2009.
[13] K.V. Mardia and P.E. Jupp, Directional Statistics, Wiley series in probability and mathematical statistics, Wiley, Chichester, 1999.
[14] C.M. Bishop, Pattern Recognition and Machine Learning, Information science and statistics, Springer, New York, 2006.
[15] K. Iwata and A. Hayashi, "A redundancy-based measure of dissimilarity among probability distributions for hierarchical clustering criteria," IEEE Trans. Pattern Anal. Mach. Intell., vol.30, no.1, pp.76–88, Jan. 2008.
[16] Y. Ouchi and K. Iwata, "Data set of drawings," Feb. 2009. http://www.prl.info.hiroshima-cu.ac.jp/˜kiwata/ouchiiwata/
[17] E.G. Petrakis, "Shape datasets and evaluation of shape matching methods for image retrieval," Nov. 1998. http://www.intelligence.tuc.gr/˜petrakis/miscellaneus.html

Katsutoshi Ueaoki received a B.E. and an M.E. degree from Hiroshima City University, Hiroshima, Japan, in 2009 and 2011, respectively. He joined Yokogawa Solutions Corporation in April 2011.

Kazunori Iwata received a B.E. and an M.E. degree from Nagoya Institute of Technology, Aichi, Japan in 2000 and 2002, respectively, and a Ph.D. degree in Informatics from Kyoto University, Kyoto, Japan, in 2005. He was a JSPS research fellow from April 2002 to March 2005. He has been with Hiroshima City University, Hiroshima, Japan, since April 2005. His research interests include machine learning, statistical inference, and information theory. He received the IEEE Kansai-Section Student Paper Award in February 2005.

Nobuo Suematsu received a B.S. and an M.S. degree in Physics from Kyushu University, Fukuoka, Japan, in 1988 and 1990, respectively. He received a Ph.D. degree in Engineering from Osaka University, Osaka, Japan, in 2000. He joined Fujitsu Laboratories LTD in 1990. Since 1994, he has been with Hiroshima City University, Hiroshima, Japan. Currently, he is an Associate Professor at the Graduate School of Information Sciences, Hiroshima City University. His research interests include machine learning and pattern recognition.

Akira Hayashi received a B.S. degree in Mathematics from Kyoto University, Kyoto, Japan, in 1974, an M.S. degree in Computer Science from Brown University, Providence, RI, in 1988, and a Ph.D. degree in Computer Science from the University of Texas at Austin, in 1991. He joined IBM Japan in April 1974. He was a visiting Associate Professor at Kyushu Institute of Technology, Fukuoka, Japan, until March 1994. Currently, he is a Professor at the Graduate School of Information Sciences, Hiroshima City University, Hiroshima, Japan. His research interests include machine learning and pattern recognition.