

Structure Integral Transform Versus Radon Transform: A 2D Mathematical Tool for Invariant Shape Recognition

Bin Wang, *Member, IEEE*, and Yongsheng Gao, *Senior Member, IEEE*

Abstract—In this paper, we present a novel mathematical tool, Structure Integral Transform (SIT), for invariant shape description and recognition. Different from the Radon Transform (RT), which integrates the shape image function over a 1D line in the image plane, the proposed SIT builds upon two orthogonal integrals over a 2D K -cross dissecting structure spanning across all rotation angles by which the shape regions are bisected in each integral. The proposed SIT brings the following advantages over the RT: 1) it has the extra function of describing the interior structural relationship within the shape which provides a more powerful discriminative ability for shape recognition; 2) the shape regions are dissected by the K -cross in a coarse to fine hierarchical order that can characterize the shape in a better spatial organization scanning from the center to the periphery; and 3) it is easier to build a completely invariant shape descriptor. The experimental results of applying SIT to shape recognition demonstrate its superior performance over the well-known Radon transform, and the well-known shape contexts and the polar harmonic transforms.

Index Terms—Structure Integral Transform, Radon Transform, shape description, shape analysis, MPEG-7, image retrieval.

I. INTRODUCTION

IN HUMAN perception, the shape of an object is known to play an important role for object identification and can often provide enough visual information for correctly recognizing the object [21]. In computer vision and image analysis, characterizing objects using their shape features is a fundamental problem in the literature, which has found many applications such as trademark image retrieval [1], logo recognition [18], patent binary image retrieval [13], [22], medical image retrieval [19], and human activity recognition [20].

The central issue in shape analysis is how to effectively describe the essence of the shape [16], [17]. Building a completely invariant shape descriptor with powerful discriminative ability remains a challenging task, for which some sophisticated mathematical tools are being explored for addressing this

issue [15]. Radon transform [14] has been widely employed as a mathematical tool for developing image analysis algorithms [1], [2], [25]–[29] (Hough transform is one of its derivatives [5]). It integrates the image function over all the possible straight lines defined in its domain. With Radon transform, the image is transformed into another “transformed image” which is a function of (θ, λ) that characterizes each line. However, as Radon transform is mathematically only a set of line integrals of the function, the current shape analysis methods built on its theory cannot appropriately describe the structure relationship within the target object.

In this paper, we propose a novel integral transform, Structure Integral Transform (SIT), for invariant shape description and matching. Integrating the shape image function over the 2D dissecting structure in various directions can capture not only the geometrical information of the shape, but also the interior structural relationship of the shape regions cut by the K -cross dissecting structure at all dissecting levels (i.e., $i = 1, 2, 3, \dots, K$), which will largely contribute to the yielding of the concise shape descriptor with powerful discriminative ability.

The rest of the paper is organized as follows: A brief review of related work is presented in Section II. Section III describes the details of the concept of 2D dissecting structure and the proposed Structure Integral Transform (SIT). A discrete SIT algorithm is developed for invariant shape recognition. In Section IV, we report the experiments that have been conducted to validate the feasibility and effectiveness of the proposed method. Finally, the paper concludes in Section V.

II. RELATED WORK

Shape description and matching techniques can be classified into two main categories. One is contour based methods [6]–[8], [10], [32]–[45], which analyze the silhouette of the shape as a sequence of landmarks. Due to the 1D nature of the boundary curve of the object, dynamic programming approaches can be used for the shape matching by searching for curve point or curve segment correspondences with the minimization of a cost. Since the dynamic programming scheme is used to take advantage of the sequential information of the boundary, these methods have very high discriminative ability for analyzing those contour shapes as in MPEG-7 CE-1 database (see Fig. 1(a)). However, they cannot consider the interior content of a shape and fail to function for those region shapes as in MPEG-7 CE-2 database (see Fig. 1(b)).

Manuscript received March 3, 2016; revised July 10, 2016; accepted August 26, 2016. Date of publication September 15, 2016; date of current version October 7, 2016. This work was supported in part by the Australian Research Council under Grant DP0877929 and Grant DP140101075, and in part by the National Natural Science Foundation of China under Grant 61372158. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Yonggang Shi.

B. Wang is with the School of Information Engineering, Nanjing University of Finance and Economics, Nanjing 210023, China (on leave) and the School of Engineering, Griffith University, Brisbane, Qld 4111, Australia (e-mails: wangbin@njue.edu.cn; bin.wang@griffith.edu.au).

Y. Gao is with the School of Engineering, Griffith University, Brisbane, Qld 4111, Australia (e-mail: yongsheng.gao@griffith.edu.au).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2016.2609816

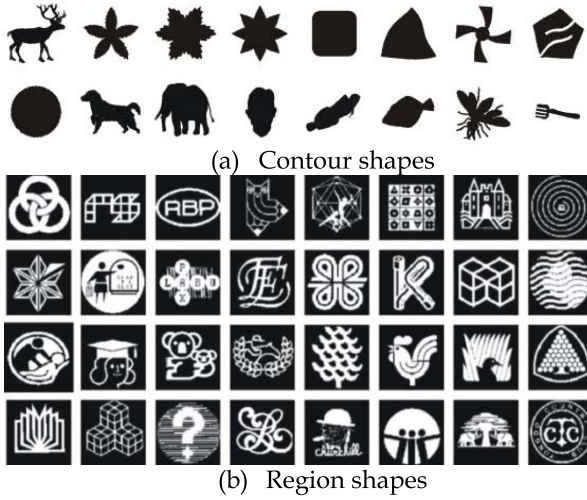


Fig. 1. Contour shapes vs. region shapes. (a) Sample contour shapes from MPEG-7 CE-1 database. (b) Sample region shapes from MPEG-7 CE-2 database.

The other is region based methods [1], [3], [9], [11], [24], [29]–[31], which can analyse more challenging shapes containing complex interior contents, disjoint regions, and unordered landmarks. Radon transform is a mathematical tool often used for describing and recognizing region shapes.

Radon transform is a line integral transform to convert the shape image into another “image”, which is a function of parameters (λ, θ) characterizing the line (see Fig. 2(a)). Tabbone *et al.* [50] and Hoang and Tabbone [51] developed a method that can remove the translation and scaling parameters in Radon transform by using an integral operation on λ . Fourier transform is then applied to the obtained one-dimensional function of angle θ for removing the rotation parameter. This method has been successfully applied to human activity recognition [20], [52]. In [2] and [4], the Fourier-Mellin transform are applied to the Radon transformed “image” to achieve invariant description of shapes, which obtained highly competitive performances in region shape recognition.

Radon composite features (RCF) [1] is another favorable shape descriptor derived from Radon transform. In this method, three morphological features and a spectral feature vector are extracted from the Radon plane to effectively represent both the statistical and the visual attributes of a shape. A remarkable point of RCF is that it possesses ability to distinguish partial shape characteristics by providing more precise information of inner structures of the shape. When applied to image retrieval, the RCF method can achieve a hierarchical representation through some retrieval strategies.

III. THE PROPOSED STRUCTURE INTEGRAL TRANSFORM

Radon transform is a mathematical tool often used to capture the shape information through line integral along a varying orientation [14]. Given a function $f(x, y)$ on the 2D plane R^2 , its Radon transform is mathematically expressed as

$$R_f(\lambda, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \cos \theta + y \sin \theta - \lambda) dx dy. \quad (1)$$

In other words, $R_f(\lambda, \theta)$ is the line integral of function $f(x, y)$ over the straight line defined by

$$L(\lambda, \theta) = \{(x, y) \in \mathbb{R}^2 | x \cos \theta + y \sin \theta = \lambda\}, \quad (2)$$

where $\lambda \in (-\infty, \infty)$ is the signed perpendicular distance from the origin to the line, $\theta \in [0, \pi)$ is the angle between the line and the y-axis, and $\delta(\cdot)$ is the Dirac delta function. In shape analysis, the Radon transform describes the lengths of all the possible lines fall in the shape region (see Fig. 2). It is clear, as shown in Fig. 2(b), that the distribution of integral lines is biased resulting uneven description capability at different regions irrelevant to the characteristics of shape. More importantly, the structural relationship information within the shape regions is not encoded in the integral over its 1D line. In this section, we present a new Structure Integral Transform that can address above problems.

A. 2D Dissecting Structure and Structure Integral Transform

Given a shape image $f(x, y)$ and K crosses, a cross at an angle $\theta = 0^\circ$ is used to partition the shape into two equal-sized parts by its horizontal line and two equal-sized parts by its vertical line, respectively. Then a second cross ($i = 2$) is added to dissect the upper half shape cut by the first cross. The two crosses form a 2-cross *Dissecting Structure*. Continue the above process by adding one more cross at a time to the dissecting structure till $i = K$, resulting in a K -cross dissecting structure (Fig. 3).

Mathematically, the above 2D K -cross dissecting structure at angle θ can be expressed as

$$\Gamma(\theta) = \bigcup_{i=1}^K \left\{ L(\lambda^i(\theta), \theta), L(\mu^i(\theta), \theta + \pi/2) \right\} \quad (3)$$

where $L(\lambda^i(\theta), \theta)$ and $L(\mu^i(\theta), \theta + \pi/2)$ are the horizontal and vertical lines of the i th cross in a dissecting structure at angle θ .

The i th cross $\{L(\lambda^i(\theta), \theta), L(\mu^i(\theta), \theta + \pi/2)\}$ is uniquely determined by three parameters of θ , λ and μ . θ is the orientation of the dissecting structure. λ and μ are the signed perpendicular distances from the origin to the two orthogonal lines of the cross, which can be computed by

$$\lambda^i(\theta) = \arg \min_{\hat{\lambda}} \left\{ \begin{aligned} & \iint_{x \cos \theta + y \sin \theta \leq \hat{\lambda}} f_{\theta}^{(i-1)}(x, y) dx dy \\ & = \frac{1}{2} \iint_{-\infty}^{\infty} f_{\theta}^{(i-1)}(x, y) dx dy \end{aligned} \right\} \quad (4)$$

and

$$\mu^i(\theta) = \arg \min_{\hat{\mu}} \left\{ \begin{aligned} & \iint_{x \cos(\theta + \frac{\pi}{2}) + y \sin(\theta + \frac{\pi}{2}) \leq \hat{\mu}} f_{\theta}^{(i-1)}(x, y) dx dy \\ & = \frac{1}{2} \iint_{-\infty}^{\infty} f_{\theta}^{(i-1)}(x, y) dx dy \end{aligned} \right\}. \quad (5)$$

$f_{\theta}^{(i-1)}$ denotes the upper half of the shape cut by the line $L(\lambda^{i-1}(\theta), \theta)$ in the $(i-1)$ th cross, which can be calculated

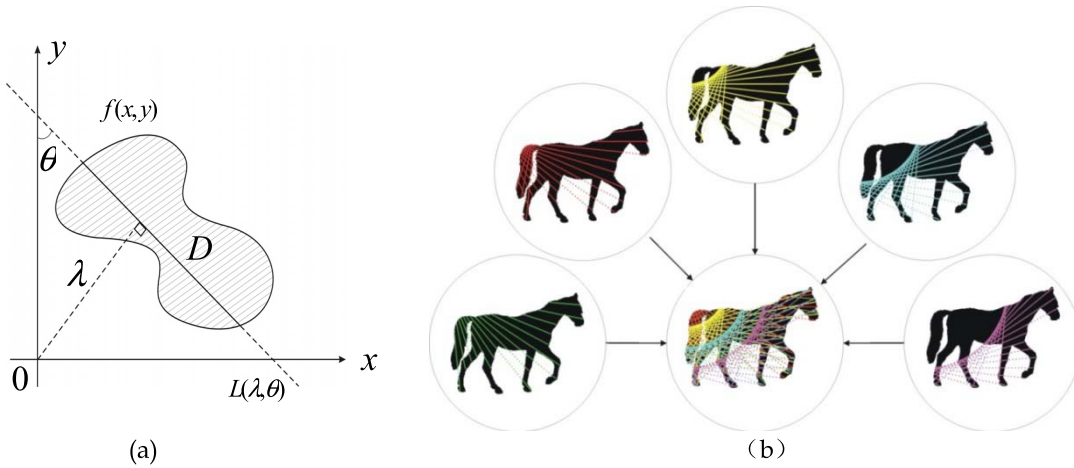


Fig. 2. Radon transform. (a) Integral on a line of given λ and θ passing through shape region D. (b) Integral on lines with varying $\lambda = 0, 50, 100, 150$, and 200 (shown in 5 different colors) and $\theta \in [0, \pi)$.

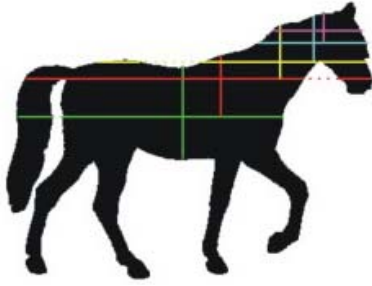


Fig. 3. Construction of a 2D K -cross dissecting structure.

as

$$f_{\theta}^{(i)}(x, y) = \begin{cases} f_{\theta}^{(i-1)}(x, y) & \text{if } x \cos \theta + y \sin \theta \leq \lambda^i(\theta) \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

When $i = 0$, $f^{(0)}(x, y)$ is the original shape image function of $f(x, y)$. Applying Eq. (4) and Eq. (6) iteratively, we can obtain all the $f_{\theta}^{(i-1)}(x, y)$ for $i = 1, 2, 3, \dots, K$. Because the shape $f_{\theta}^{(i-1)}(x, y)$ may consist of unconnected regions (see Fig. 8), a line that equally partitions the shape may appear anywhere between two unconnected regions. To guarantee the uniqueness of the line location in the dissecting structure, the minimum operator is used in Eq. (4) and Eq. (5).

The proposed Structure Integral Transform (SIT) is defined as the orthogonal integrals of function $f(x, y)$ over the 2D K -cross dissecting structure $\Gamma(\theta)$:

$$\Psi_f(i, \theta) = \frac{1}{\sqrt{S}} \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\theta}^{(i-1)}(x, y) \delta(x \cos \theta + y \sin \theta - \lambda^i(\theta)) dx dy \right. \\ \left. + j \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\theta}^{(i-1)}(x, y) \delta(x \cos(\theta + \pi/2) + y \sin(\theta + \pi/2) - \mu^i(\theta)) dx dy \right), \\ i = 1, \dots, K, \quad \theta \in [0, 2\pi) \quad (7)$$

where S is the area of the shape regions:

$$S = \iint_{-\infty}^{\infty} f(x, y) dx dy. \quad (8)$$

The SIT is a complex valued function whose real and imaginary parts are integrals of function $f(x, y)$ over two orthogonal directions, which encode the structural relationship in the shape function of $f(x, y)$ at each dissecting level i .

By rotating the K -cross dissecting structure $\Gamma(\theta)$ across its range of $\theta \in [0, 2\pi)$, the SIT coefficients describe the orthogonal length pairs of all possible crosses in the 2D K -cross dissecting structure $\Gamma(\theta)$ that fall in the shape regions. It can be seen from Fig. 4 that, for a small i , the SIT calculates the integrals of a large shape area over its corresponding cross. For example, when $i=1$, the whole shape area is considered for integral computation. When i increases, the SIT becomes the integrals of a smaller part of the shape over its corresponding cross, which captures finer details of the shape. Together, the SIT hierarchically describes the structural characteristics of the shape from coarse to fine and from the center of the shape to its periphery in a much better organization than the Radon transform (see Fig. 4 in comparison to Fig. 2(b)). A graphical illustration of SIT on three leaf shapes is given in Fig. 5. When $i = 1$, the SIT curves (for both real and imaginary components) of the three shapes are almost same because they are globally similar. When $i = 7$, the SIT curves of shape b are clearly different to those of shapes a and c due to the fine details at the edge of the leaves. This shows how SIT effectively separates and characterizes shapes in a coarse to fine and from the center to border manner controlled by i .

B. Discrete Structure Integral Transform for a Shape Image

Here we present the algorithm of computing discrete Structure Integral Transform for a shape image (see Algorithm 1).

Because K may be larger than $\log M$, where M is the number of pixels in the shape, the Step 3 of the Algorithm 1 ensures the validity of shape partitions for all $i = 1, \dots, K$ by setting $\Psi(i, l) = 0$ for $\log M < i \leq K$.

We now analyze the computational complexity of Algorithm 1. Because Steps 6, 10 and 12 take time $O(M)$, $O(M \log M)$ (using fast sorting algorithm) and $O(K + M)$ respectively, the algorithm has a complexity of $O(N(M + M \log M + K + M))$. As $K < \log M$ (Step 3), the computational complexity of Algorithm 1 is $O(NM \log M)$.

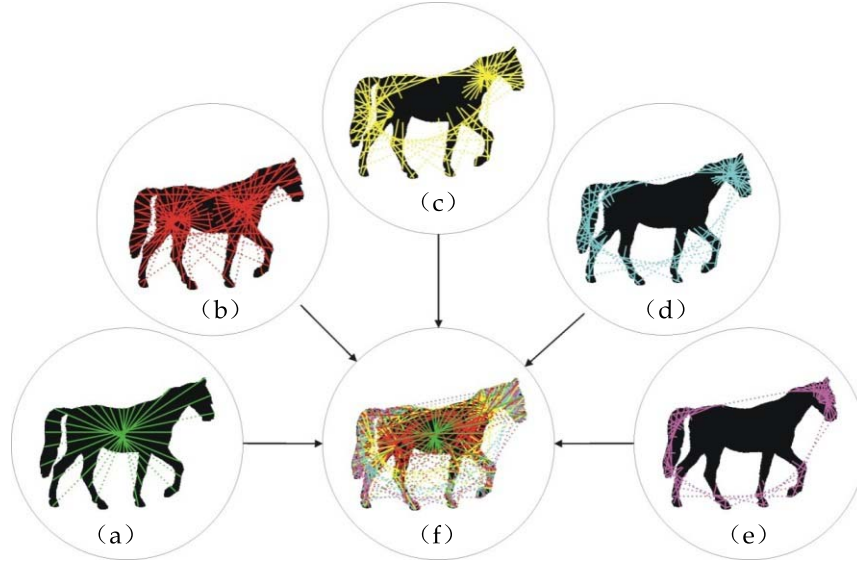


Fig. 4. Structure Integral Transform. (a)-(e) Integral on the i th ($i = 1, \dots, 5$) cross in the dissecting structure; (f) Integral on the dissecting structure in Fig. 3 with varying $\theta \in [0, 2\pi)$.

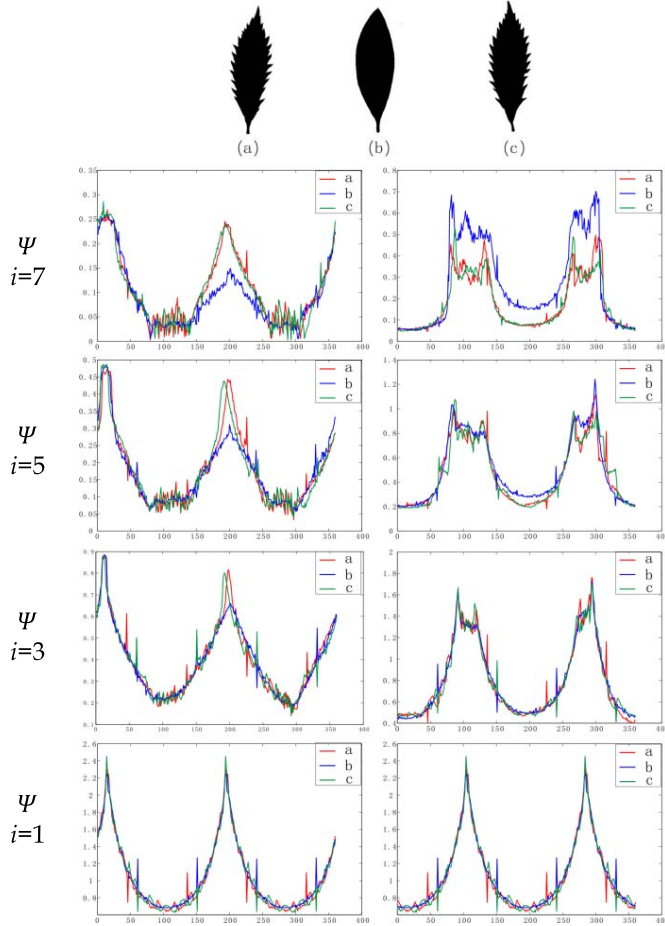


Fig. 5. A graphical illustration of the Structure Integral Transform $\Psi(i, \theta)$ applied on three leaf shapes (shown on the top row and labeled by a, b and c respectively). The left and right columns compare the real and imaginary parts of these SITs versus θ , respectively.

C. Invariant Feature Construction and Application to Shape Recognition

In this subsection, we use the proposed Structure Integral Transform (SIT) to build an invariant shape descriptor and

apply it for shape recognition. A desirable shape descriptor should not only have a high discriminative ability, but also be invariant to rotation, scaling, and translation.

The proposed Structure Integral Transform (SIT), in comparison to Radon transform (RT), has the following properties (see proofs in Appendix A):

- Translating the function $f(x, y)$ by x_0 and y_0 , i.e. $g(x, y) = f(x - x_0, y - y_0)$, we have

$$\text{SIT: } \Psi_g(i, \theta) = \Psi_f(i, \theta), \quad (P_1)$$

$$\text{RT: } R_g(\lambda, \theta) = R_f(\lambda - x_0 \cos \theta + y_0 \sin \theta, \theta)$$

- Scaling the function $f(x, y)$ by α , i.e. $g(x, y) = f(\alpha x, \alpha y)$, we have

$$\text{SIT: } \Psi_g(i, \theta) = \Psi_f(i, \theta), \quad (P_2)$$

$$\text{RT: } R_g(\lambda, \theta) = \frac{1}{\alpha} R_f(\alpha \lambda, \theta)$$

- Rotating the function $f(x, y)$ by angle θ_0 , i.e. $g(x, y) = f(x \cos \theta_0 + y \sin \theta_0, -x \sin \theta_0 + y \cos \theta_0)$, we have

$$\text{SIT: } \Psi_g(i, \theta) = \Psi_f(i, \theta - \theta_0), \quad (P_3)$$

$$\text{RT: } R_g(\lambda, \theta) = R_f(\lambda, \theta - \theta_0)$$

These properties show that SIT is intrinsically invariant to translation and scaling, while RT is variant to both translation and scaling. Rotation only leads to a shift in the variable θ for both SIT and RT. This makes SIT much more attractive than Radon transform when building a completely invariant shape descriptor, in addition to its stronger discriminative ability.

For a given i , $\Psi_f(i, \theta)$ can be considered as a signal and the magnitude of its Fourier transform coefficients is invariant to rotation. To keep the obtained descriptor robust to noise and compact, the lowest N' order coefficients are used for describing the shape, where $N' \ll N$. Thus, the shape descriptor is a matrix $(\tilde{\Psi}_f(i, l))_{K \times N'}$, where $\tilde{\Psi}_f(i, l)$ denotes the magnitudes of the Fourier coefficients of $\Psi_f(i, \theta)$.

The dissimilarity between two given shapes f and g can be measured using the fast L_1 Minkowski distance of the two

Algorithm 1 Calculating Discrete Structure Integral Transform**Input:**

f: binary shape;
 K: the number of crosses in dissecting structure;
 N: the sampling number of angle in range $[0, 2\pi)$;

Output:

$(\Psi(i, l))_{K \times N}$: Structure Integral Transform of f;
 1: Initialize: $(\Psi(i, l))_{K \times N} \leftarrow 0$; $\Delta\theta \leftarrow 2\pi/N$;
 2: Let $A \leftarrow \{(x_t, y_t), t = 1, \dots, M\}$ be the set of pixels in shape f, where M is the total number of pixels in the shape;
 3: $K \leftarrow \min(\lceil \log M \rceil, K)$;
 4: **FOR** $l=1$ **TO** N **DO**
 5: $\theta \leftarrow (l-1)\Delta\theta$;
 6: **FOR** $t=1$ **TO** M **DO**
 7: $r_t \leftarrow \text{round}(x_t \cos \theta + y_t \sin \theta)$;
 8: $h_t \leftarrow \text{round}(x_t \cos(\theta + \pi/2) + y_t \sin(\theta + \pi/2))$;
 9: **ENDFOR**
 10: Let $U \leftarrow \{\tilde{r}_1, \dots, \tilde{r}_M\}$ and $V \leftarrow \{\tilde{h}_1, \dots, \tilde{h}_M\}$ be the ascending sort of $\{r_1, \dots, r_M\}$ and $\{h_1, \dots, h_M\}$, respectively; Let w_1, \dots, w_M be the indexes of $\tilde{r}_1, \dots, \tilde{r}_M$ in the original sequence $\{r_1, \dots, r_M\}$.
 11: $E \leftarrow M$;
 12: **FOR** $i=1$ **TO** K **DO**
 13: $m \leftarrow \lfloor E/2 \rfloor$;
 14: Let C_x be the number of the elements (i.e. pixels) in U whose values equal the median value \tilde{r}_m (This is the integral over line $L(\lambda^i(\theta), \theta)$ that partitions the region $f_{\theta}^{(i-1)}(x, y)$ into two equal-sized areas);
 15: Let C_y be the number of the elements (i.e. pixels) in V whose values equal the median value \tilde{h}_m (This is the integral over line $L(\mu^i(\theta), \theta + \pi/2)$ that partitions the region $f_{\theta}^{(i-1)}(x, y)$ into two equal-sized areas);
 16: $\Psi(i, l) \leftarrow C_x + j \cdot C_y$, where $j = \sqrt{-1}$;
 17: Delete the elements $\tilde{r}_{m+1}, \dots, \tilde{r}_E$ from the sequence U , and delete the elements $h_{w_{m+1}}, \dots, h_{w_E}$ from the sequence V (This is to delete those r_t and h_t that correspond to pixels in the bottom half of the region);
 18: $E \leftarrow m$;
 19: **ENDFOR**
 20: **ENDFOR**

matrixes:

$$d(f, g) = \sum_{i=1}^K \sum_{l=1}^{N'} \left| \tilde{\Psi}_f(i, l) - \tilde{\Psi}_g(i, l) \right|. \quad (9)$$

IV. EXPERIMENTAL RESULTS

A comparative experimental study between the proposed Structure Integral Transform and Radon Transform [1], [2] are conducted using the MPEG-7 CE-2 perspective transform test, subjective test and overall robustness test, as well as on the COIL-20 database and MPEG-7 CE-1 database. We also

TABLE I
THE BULLS-EYE TEST SCORES FOR THE CE-2 PERSPECTIVE TEST

Algorithm	Bulls-eye test score (%)
Proposed method	88.2
Radon composite features (RCF) [1]	63.7
Radon-Fourier-Mellin transforms (RFM) [2]	61.4
Polar harmonic transforms (PHTs) [3]	64.4
Shape Contexts [11]	72.2

compare the performance of the proposed SIT with another two region based methods, the well-known shape contexts [11] and the recent Polar Harmonic Transforms [3] methods.

A. MPEG-7 CE-2 Perspective Transform Test

MPEG-7 CE-2 A4 dataset is built for testing the robustness to perspective transforms of region-based shape recognition algorithms. In the MPEG-7 CE-2 perspective transform test protocol, all the 3101 region-based shape images in the A4 dataset are used as gallery images. Among them, 330 images of 30 classes with 11 images in each class are labeled as queries for the retrieval experiment. In each query class, one image is the original shape, and the other ten images are its perspective transformed versions (see Fig. 6).

The commonly used bulls-eye test score [6], [10]–[13] and precision-recall curve [1], [7], [8], [12] are both employed in performance measurement. In the bulls-eye test score measurement, each query shape image is matched against all the shape images in the dataset. The number of correct matches (that is the retrieved shape and the query shape belong to the same class) in the top $2N$ (where N is the number of images in the database that belongs to the same class of the query image) retrieved shapes that have the smallest dissimilarity values are counted. The average percentage of matched shapes out of N over all the queries is the bulls-eye test score.

Tables I summarizes the bulls-eye test scores obtained from the MPEG-7 CE-2 perspective transform test. It can be seen that the proposed method achieved encouraging accuracy increases of 24.5% and 26.8% over RCF [1] and RFM [2], demonstrating the effectiveness of the proposed SIT in describing structural relationship within a shape region and its superior discriminability over the existing 1D line integral strategy in Radon transform. While comparing with the other two region based methods, the proposed SIT achieves 23.8% and 16.0% higher accuracies than polar harmonic transforms (PHTs) [3] and shape contexts [11] respectively.

The precision-recall curve provides a combined performance measurement, in which precision is defined as the percentage of similar shapes retrieved with respect to the total number of retrieved shapes, and recall is defined as the percentage of similar shapes retrieved with respect to the total number of similar shapes in the database.

To finely compare the behavior of the proposed SIT against the benchmarks, their Precision-Recall (PR) curves of the



Fig. 6. Two sample classes of images used as queries in MPEG-7 CE-2 A4 dataset (in each class, the left most image is the original shape and the remaining 10 images are undergone various perspective transforms).

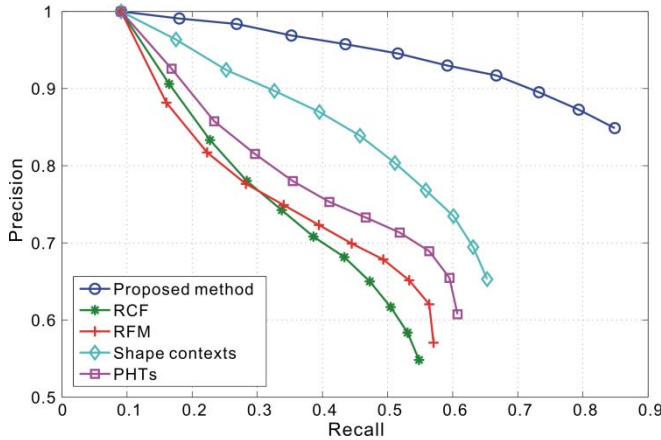


Fig. 7. The precision-recall curves for CE-2 perspective transform test.



Fig. 8. Samples from group 7 (the first row) and group 8 (the second row) in MPEG-7 CE-2 B dataset.

top 11 matches are provided in Fig. 7. The proposed approach obtains a significantly better Precision-Recall (PR) curve than the Radon methods and the other two region based methods, confirming the effectiveness of the proposed SIT in shape analysis.

A computational time analysis of the proposed method is conducted on the same dataset. Computing the SIT descriptors of all the 3101 shape images takes 187.79 seconds, that is, 0.06 second for each shape. After extracting the SIT descriptors, matching one query SIT descriptor with those of all images in the dataset takes 0.007 second in average, that is, 2.26×10^{-6} for matching two SIT descriptors. Our algorithm is implemented using MATLAB and is executed on a PC with Intel i7-4910 2.9 GHz CPU and 32 GB RAM under Windows 7(x64). These results demonstrated the attractive computation efficiency of the proposed method.

B. MPEG-7 CE-2 Subjective Test

MPEG-7 CE-2 B dataset is used for subjective test, which contains 2801 shape images used as galleries. Among them, 678 shape images are manually classified into 10 groups and are labeled as queries for the retrieval experiment. The numbers of images in groups 1 to 10 are 68, 244, 22, 28, 17,

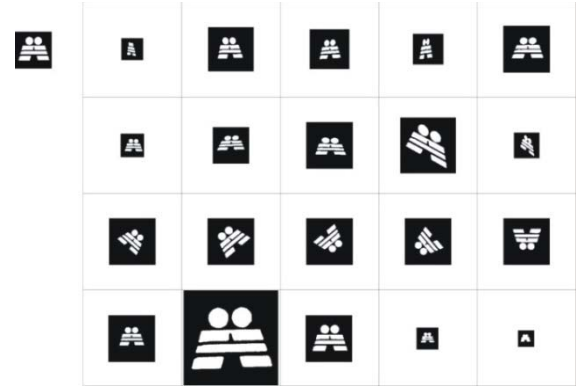


Fig. 9. An example of the complete 21 images from one class used as queries in the CE-2 overall robustness test. The top left: the original shape. The shapes in the first two rows: 10 perspective transformed versions. The third row: rotated versions. The last row: scaled versions.

22, 45, 145, 45, and 42. Fig. 8 shows 10 example images from groups 7 and 8 respectively.

Under the challenging MPEG-7 CE-2 subjective test protocol, the proposed method (see Table II) achieved a retrieval accuracy that is 23% and 6.2% higher than the Radon transform methods of RCF [1] and RFM [2], and is 2.7% and 23% higher than the polar harmonic transforms (PHTs) [3] and shape contexts [11], respectively. This may imply that the orthogonal integrals on 2D dissecting structure of the proposed Structure Integral Transform characterizes image shapes in a way closer to human perception in shape classification than the non-structural 1D line integral strategy of Radon transform.

C. MPEG-7 CE-2 Overall Robustness Test

The MPEG-7 CE-2 overall robustness test protocol utilizes the 3621 shape images contained in the complete CE-2 dataset as galleries for evaluating region-based shape matching algorithms. Among them, 651 images, which are organized into 31 classes with 21 images in each one, are labeled as queries for the retrieval experiment. Each class contains one reference shape, ten perspective transformed versions, five rotated versions, and five scaled versions. Fig. 9 shows an example of the complete 21 images from one class under the MPEG-7 CE-2 overall robustness test protocol.

Consistently, the proposed Structural Integral Transform achieved a higher score than Radon transform based methods (see Table III). The encouraging 91.5% of retrieval accuracy (over 24% higher than the two Radon transform based methods and more than 20% higher than the other two benchmark methods) reconfirms the discriminative ability of the proposed

TABLE II
THE BULLS-EYE TEST SCORES FOR THE CE-2 SUBJECTIVE TEST

Algorithm	Bulls-eye test score (%)										
	Group Number										
	1	2	3	4	5	6	7	8	9	10	Average
Proposed method	43.4	75.5	67.4	63.3	73.4	32.1	42.9	67.6	41.5	32.5	54.0
RCF [1]	27.8	47.5	21.7	30.1	20.4	17.6	13.5	46.3	56.2	28.7	31.0
RFM [2]	40.0	21.2	60.5	64.8	71.6	11.6	26.4	59.6	49.2	36.4	47.8
PHTs [3]	47.5	64.2	71.3	71.8	79.9	22.5	48.4	33.2	47.1	26.6	51.3
Shape Contexts [11]	20.1	43.1	27.5	31.6	27.0	22.3	14.3	49.3	55.8	22.9	31.0

TABLE III
THE BULLS-EYE TEST SCORES FOR THE CE-2 OVERALL ROBUSTNESS TEST

Algorithm	Bulls-eye test score (%)
Proposed method	91.5
Radon composite features (RCF) [1]	67.4
Radon-Fourier-Mellin transforms (RFM) [2]	60.9
Polar harmonic transforms (PHTs) [3]	64.1
Shape Contexts [11]	71.0

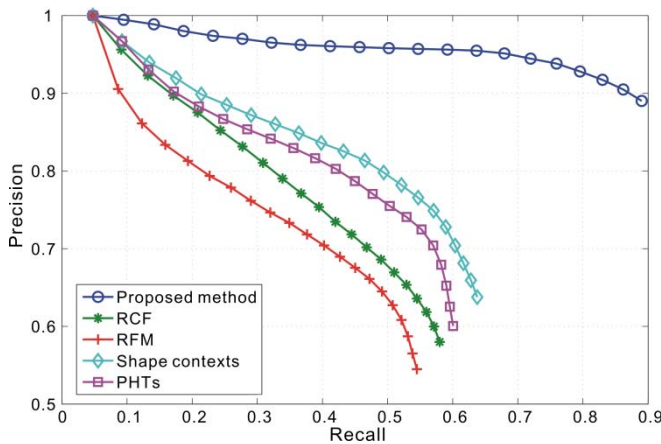


Fig. 10. The precision-recall curves for the CE-2 overall robustness test.

SIT in shape representation and analysis. To finely compare the behavior of the proposed SIT against benchmarks, their Precision-Recall (PR) curves of the top 21 matches are plotted in Fig. 10. It is noticed that the proposed approach obtains a much better Precision-Recall (PR) curve than the Radon transform methods with large margins in both overall robustness test and perspective transform test (PR curve is not available for the subjective test due to varying number of images in each group), demonstrating the superiority and robustness of SIT than Radon transform as an new and effective mathematical tool in shape analysis.



Fig. 11. 20 object samples from the COIL-20 database.



Fig. 12. Example views of one object in the COIL-20 database.

D. COIL-20 Database Test

The proposed SIT is also tested on the COIL-20 database [53] which contains 20 common household objects (see Fig. 11). The objects have a wide variety of complex geometric, appearance and reflectance characteristics. Each object was placed on a motorized turntable against a black background and images of the object were taken every 5 degrees for a total of 72 views by rotating the turntable through 360 degrees (Example views from one object are displayed in Fig.12). There are a total of 1440 images in the database (72 views per object).

All the views in the database are converted to region based binary shapes for our comparative experiments (See example in Figs. 13 and 14 for those images shown in Figs. 11 and 12). The same test protocol as used for MPEG-7 CE-2 database is adopted. All the binary shapes in the database are used in turn as queries for retrieving the similar shapes from the database. The bulls-eye test scores for the proposed SIT



Fig. 13. Binary shapes from the images shown in Fig. 11.



Fig. 14. Binary shapes from the images shown in Fig. 12.

TABLE IV
THE BULLS-EYE TEST SCORES OBTAINED ON THE COIL-20 DATABASE

Algorithm	Bulls-eye test score (%)
Proposed method	74.5
Radon composite features (RCF) [1]	64.3
Radon-Fourier-Mellin transforms (RFM) [2]	65.9
Polar harmonic transforms (PHTs) [3]	69.2
Shape Contexts [11]	67.9

and the benchmark approaches are summarized in Table IV. It can be seen that the proposed method achieves the best score of 74.5%, which is 10.2% and 8.6% higher than the two Radon transform based methods and is 5.3% and 6.6% higher than the polar harmonic transforms (PHTs) and shape contexts methods respectively. A detailed performance comparison of the proposed SIT against the benchmarks using their PR curves of the top 72 matches is shown in Fig. 15. The proposed method obtains a significantly better PR curves over the other comparative methods with significant margins.

E. MPEG-7 CE-1 Test

MPEG-7 CE-1 Part B database [6] contains 1400 binary shapes organized into 70 groups (see Figure 16) with 20 similar shapes in each group (see Fig. 17). It is widely used for evaluating contour based methods because they can be recognized by their contours without considering interior structures (most of them do not have interior content). In this subsection, we also use this database to compare the performance of the

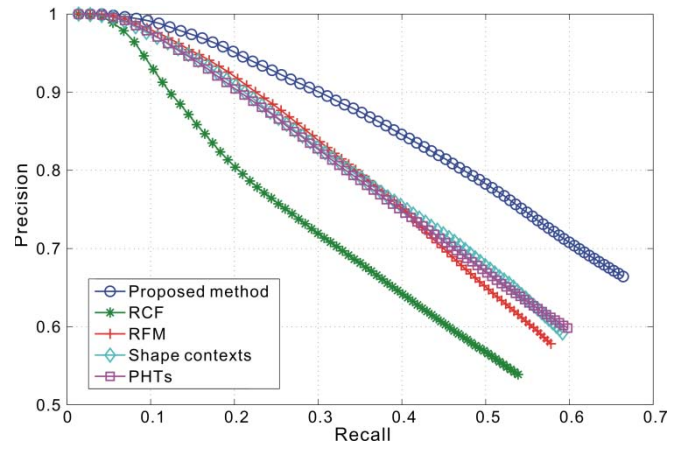


Fig. 15. The precision-recall curves obtained on the COIL-20 database.



Fig. 16. Example shapes of 70 categories from the MPEG-7 CE-1 Part B database.



Fig. 17. 20 shapes of the same category in MPEG-7 CE-1 Part B database.

TABLE V
THE BULLS-EYE TEST SCORES OBTAINED ON
THE MPEG-7 CE-1 PART B DATABASE

Algorithm	Bulls-eye test score (%)
Proposed method	80.1
Radon composite features (RCF) [1]	68.7
Radon-Fourier-Mellin transforms (RFM) [2]	68.0
Polar harmonic transforms (PHTs) [3]	67.1
Shape Contexts [11]	72.8

proposed SIT with the Radon methods and the region based benchmarks. All the shapes in the database are used, according to the bulls-eye test protocol, as queries in turn for retrieving

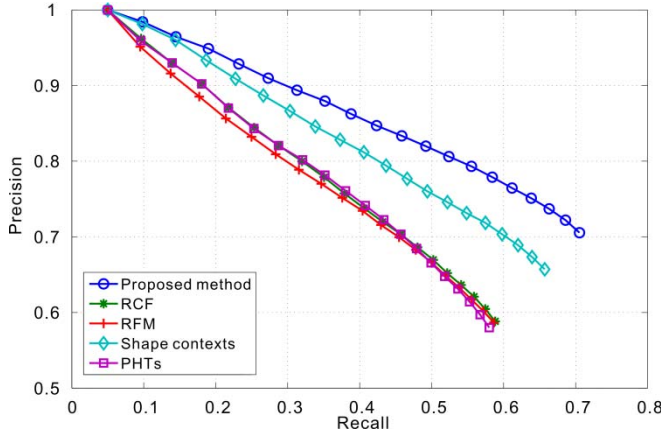


Fig. 18. The precision-recall curves obtained on the MPEG-7 CE-1 Part B database.

the similar shapes from the database. The bulls-eye test scores and the precision-recall curves for all the competing methods are given in Table V and Fig. 18 respectively. It can be seen from Table V that the proposed SIT achieves a retrieval score of at least 7.3% higher than the benchmarks. At each precision-recall curve point (see Fig. 18), the corresponding precision and recall values obtained by the SIT method are consistently higher than the Radon methods and the other methods. Note that the higher bulls-eye score of 76.51% reported in [11] for the shape contexts is obtained after combining a performance improvement technique of TPS [54], [55] with the shape contexts. In this paper, the results for all competing methods are obtained without using any performance improvement techniques or tricks for the purpose of direct and fair comparison.

V. CONCLUSION

We have presented a new tool for image transform and shape representation. Different from the Radon transform that is built on integrals along 1D straight lines, the proposed Structure Integral Transform (SIT) describes the orthogonal lengths of a 2D dissecting structure distributed in the shape region. It can embed the intrinsic structure relationship within the shape regions into complex valued integral functions when the 2D dissecting structure spans across all rotation angles. Thus it can be used to extract more discriminative features in shapes than existing line integral based transforms. Moreover, it is theoretically proved that the SIT of a shape image is invariant to translation and scaling, and rotation only results in a shift of the variable θ , which can be easily used to build a completely invariant shape descriptor. The performance of the proposed SIT approach has been evaluated using three standard MPEG-7 CE-2 test protocols and their datasets (i.e., MPEG-7 CE-2 perspective transform test, subject test, and overall robustness test), as well as the COIL-20 and MPEG-7 CE-1 databases, and compared against the benchmarks using Radon transform and two region based methods. The experimental results demonstrate the superiority and robustness of the proposed Structure Integral Transform over the well-known Radon Transform in describing and

recognizing region-based shapes. As a mathematical tool, the proposed SIT may find its usefulness in many other applications.

APPENDIX

Lemma 1: Let $g(x, y)$ denote the linear distorted version (translation, scaling and rotation) of a shape image $f(x, y)$.

- (1) For translation ($g(x, y) = f(x - x_0, y - y_0)$), we have

$$\lambda_g^i(\theta) = \lambda_f^i(\theta) + x_0 \cos \theta + y_0 \sin \theta, \quad (\text{A-1})$$

$$\begin{aligned} \mu_g^i(\theta) &= \mu_f^i(\theta) + x_0 \cos\left(\theta + \frac{\pi}{2}\right) \\ &\quad + y_0 \sin\left(\theta + \frac{\pi}{2}\right), \end{aligned} \quad (\text{A-2})$$

and

$$g_{\theta}^{(i)}(x, y) = f_{\theta}^{(i)}(x - x_0, y - y_0). \quad (\text{A-3})$$

- (2) For scaling ($g(x, y) = f(\alpha x, \alpha y)$ and $\alpha > 0$), we have

$$\lambda_g^i(\theta) = \frac{1}{\alpha} \lambda_f^i(\theta), \quad (\text{A-4})$$

$$\mu_g^i(\theta) = \frac{1}{\alpha} \mu_f^i(\theta), \quad (\text{A-5})$$

and

$$g_{\theta}^{(i)}(x, y) = f_{\theta}^{(i)}(\alpha x, \alpha y) \quad (\text{A-6})$$

- (3) For rotation ($g(x, y) = f(x \cos \theta_0 + y \sin \theta_0, -x \sin \theta_0 + y \cos \theta_0)$), we have

$$\lambda_g^i(\theta) = \lambda_f^i(\theta - \theta_0), \quad (\text{A-7})$$

$$\mu_g^i(\theta) = \mu_f^i(\theta - \theta_0), \quad (\text{A-8})$$

and

$$\begin{aligned} g_{\theta}^{(i)}(x, y) &= f_{\theta-\theta_0}^{(i)}(x \cos \theta_0 + y \sin \theta_0, -x \sin \theta_0 + y \cos \theta_0). \end{aligned} \quad (\text{A-9})$$

Proof:

We first prove that the above equations hold for $i = 1$.

- (1) For translation $g(x, y) = f(x - x_0, y - y_0)$:

Since $g_{\theta}^{(0)}(x, y) = g(x, y)$ and $f_{\theta}^{(0)}(x, y) = f(x, y)$, we have $g_{\theta}^{(0)}(x, y) = g(x, y) = f(x - x_0, y - y_0) = f_{\theta}^{(0)}(x - x_0, y - y_0)$. Thus, according to Eq. (8),

$$\begin{aligned} \iint_{-\infty}^{\infty} g_{\theta}^{(0)}(x, y) dx dy &= \iint_{-\infty}^{\infty} f_{\theta}^{(0)}(x - x_0, y - y_0) dx dy \\ &= \iint_{-\infty}^{\infty} f_{\theta}^{(0)}(x, y) dx dy = S. \end{aligned} \quad (\text{A-10})$$

From Eq. (4), we have

$$\begin{aligned}
 \lambda_g^1(\theta) &= \arg \min_{\hat{\lambda}} \left\{ \iint_{x \cos \theta + y \sin \theta \leq \hat{\lambda}} g_{\theta}^{(0)}(x, y) dx dy = \frac{S}{2} \right\} \\
 &= \arg \min_{\hat{\lambda}} \left\{ \iint_{x \cos \theta + y \sin \theta \leq \hat{\lambda}} f_{\theta}^{(0)}(x - x_0, y - y_0) dx dy = \frac{S}{2} \right\} \\
 &= \arg \min_{\hat{\lambda}} \left\{ \iint_{\substack{x \cos \theta + y \sin \theta \leq \\ \hat{\lambda} - x_0 \cos \theta - y_0 \sin \theta}} f_{\theta}^{(0)}(x, y) dx dy = \frac{S}{2} \right\}.
 \end{aligned} \tag{A-11}$$

Replacing $\hat{\lambda} - x_0 \cos \theta - y_0 \sin \theta$ with variable $\bar{\lambda}$, the above equation can be rewritten as

$$\lambda_g^1(\theta) = \arg \min_{\bar{\lambda} + x_0 \cos \theta + y_0 \sin \theta} \left\{ \iint_{x \cos \theta + y \sin \theta \leq \bar{\lambda}} f_{\theta}^{(0)}(x, y) dx dy = \frac{S}{2} \right\}. \tag{A-12}$$

So, we have

$$\begin{aligned}
 \lambda_g^1(\theta) &= \arg \min_{\bar{\lambda}} \left\{ \iint_{x \cos \theta + y \sin \theta \leq \bar{\lambda}} f_{\theta}^{(0)}(x, y) dx dy = \frac{S}{2} \right\} \\
 &\quad + x_0 \cos \theta + y_0 \sin \theta \\
 &= \lambda_f^1(\theta) + x_0 \cos \theta + y_0 \sin \theta.
 \end{aligned} \tag{A-13}$$

According to Eq. (5), by replacing $x \cos \theta + y \sin \theta$ and $x_0 \cos \theta + y_0 \sin \theta$ with $x \cos(\theta + \frac{\pi}{2}) + y \sin(\theta + \frac{\pi}{2})$ and $x_0 \cos(\theta + \frac{\pi}{2}) + y_0 \sin(\theta + \frac{\pi}{2})$, respectively in equations (A-11~13), we can conclude that

$$\mu_g^1(\theta) = \mu_f^1(\theta) + x_0 \cos(\theta + \frac{\pi}{2}) + y_0 \sin(\theta + \frac{\pi}{2}). \tag{A-14}$$

From Eq. (6), we have

$$g_{\theta}^{(1)}(x, y) = \begin{cases} g_{\theta}^{(0)}(x, y) & \text{if } x \cos \theta + y \sin \theta \leq \lambda_g^1(\theta) \\ 0 & \text{otherwise.} \end{cases} \tag{A-15}$$

According to (A-13), the inequality $x \cos \theta + y \sin \theta \leq \lambda_g^1(\theta)$ can be rewritten as

$$(x - x_0) \cos \theta + (y - y_0) \sin \theta \leq \lambda_f^1(\theta), \tag{A-16}$$

and Eq. (A-15) becomes

$$\begin{aligned}
 g_{\theta}^{(1)}(x, y) &= \begin{cases} f_{\theta}^{(0)}(x - x_0, y - y_0) & \text{if } (x - x_0) \cos \theta + (y - y_0) \sin \theta \\ & \leq \lambda_f^1(\theta) \\ 0 & \text{otherwise.} \end{cases}
 \end{aligned} \tag{A-17}$$

Thus we have

$$g_{\theta}^{(1)}(x, y) = f_{\theta}^{(1)}(x - x_0, y - y_0). \tag{A-18}$$

(2) For scaling $g(x, y) = f(ax, ay)$:

Since $g_{\theta}^{(0)}(x, y) = g(x, y) = f(ax, ay) = f_{\theta}^{(0)}(ax, ay)$, we have

$$\iint_{-\infty}^{\infty} g_{\theta}^{(0)}(x, y) dx dy = \iint_{-\infty}^{\infty} f_{\theta}^{(0)}(ax, ay) dx dy = \frac{S}{\alpha^2}. \tag{A-19}$$

From Eqs. (4) and (A-19), we have

$$\begin{aligned}
 \lambda_g^1(\theta) &= \arg \min_{\hat{\lambda}} \left\{ \iint_{x \cos \theta + y \sin \theta \leq \hat{\lambda}} g_{\theta}^{(0)}(x, y) dx dy = \frac{S}{2\alpha^2} \right\} \\
 &= \arg \min_{\hat{\lambda}} \left\{ \iint_{x \cos \theta + y \sin \theta \leq \hat{\lambda}} f_{\theta}^{(0)}(ax, ay) dx dy = \frac{S}{2\alpha^2} \right\} \\
 &= \arg \min_{\hat{\lambda}} \left\{ \iint_{x \cos \theta + y \sin \theta \leq \alpha \hat{\lambda}} f_{\theta}^{(0)}(x, y) dx dy = \frac{S}{2} \right\}.
 \end{aligned} \tag{A-20}$$

Replacing $\alpha \hat{\lambda}$ with variable $\bar{\lambda}$, the above equation can be rewritten as

$$\begin{aligned}
 \lambda_g^1(\theta) &= \arg \min_{\bar{\lambda}/\alpha} \left\{ \iint_{x \cos \theta + y \sin \theta \leq \bar{\lambda}} f_{\theta}^{(0)}(x, y) dx dy = \frac{S}{2} \right\} \\
 &= \frac{1}{\alpha} \arg \min_{\bar{\lambda}} \left\{ \iint_{x \cos \theta + y \sin \theta \leq \bar{\lambda}} f_{\theta}^{(0)}(x, y) dx dy = \frac{S}{2} \right\}.
 \end{aligned} \tag{A-21}$$

Thus we have

$$\lambda_g^1(\theta) = \frac{1}{\alpha} \lambda_f^1(\theta). \tag{A-22}$$

According to Eq. (5), by replacing $x \cos \theta + y \sin \theta$ with $x \cos(\theta + \frac{\pi}{2}) + y \sin(\theta + \frac{\pi}{2})$ in Eqs. (A-20) and (A-21), we have

$$\mu_g^1(\theta) = \frac{1}{\alpha} \mu_f^1(\theta). \tag{A-23}$$

From Eq. (6), we have

$$g_{\theta}^{(1)}(x, y) = \begin{cases} g_{\theta}^{(0)}(x, y) & \text{if } x \cos \theta + y \sin \theta \leq \lambda_g^1(\theta) \\ 0 & \text{otherwise.} \end{cases} \tag{A-24}$$

According to (A-22), the inequality $x \cos \theta + y \sin \theta \leq \lambda_g^1(\theta)$ can be rewritten as

$$x \cos \theta + y \sin \theta \leq \frac{1}{\alpha} \lambda_f^1(\theta), \tag{A-25}$$

and Eq. (A-24) becomes

$$g_{\theta}^{(1)}(x, y) = \begin{cases} f_{\theta}^{(0)}(ax, ay) & \text{if } a \cos \theta + ay \sin \theta \leq \lambda_f^1(\theta) \\ 0 & \text{otherwise.} \end{cases} \quad (\text{A-26})$$

Thus, we have

$$g_{\theta}^{(1)}(x, y) = f_{\theta}^{(1)}(ax, ay) \quad (\text{A-27})$$

(3) For rotation of θ_0 degree:

Let $u = x \cos \theta_0 + y \sin \theta_0$ and $v = -x \sin \theta_0 + y \cos \theta_0$, we have

$$g_{\theta}^{(0)}(x, y) = g(x, y) = f(u, v) = f_{\theta-\theta_0}^{(0)}(u, v). \quad (\text{A-28})$$

Since $u^2 + v^2 = x^2 + y^2$, we have $dudv = dxdy$ and

$$\iint_{-\infty}^{\infty} g_{\theta}^{(0)}(x, y) dxdy = \iint_{-\infty}^{\infty} f_{\theta-\theta_0}^{(0)}(u, v) dudv = S. \quad (\text{A-29})$$

From Eq. (4), we have

$$\lambda_g^1(\theta) = \arg \min_{\hat{\lambda}} \left\{ \iint_{x \cos \theta + y \sin \theta \leq \hat{\lambda}} g_{\theta}^{(0)}(x, y) dxdy = \frac{S}{2} \right\}. \quad (\text{A-30})$$

Since the inequality $x \cos \theta + y \sin \theta \leq \hat{\lambda}$ can be rewritten as $u \cos(\theta - \theta_0) + v \sin(\theta - \theta_0) \leq \hat{\lambda}$, (A-30) becomes

$$\lambda_g^1(\theta) = \arg \min_{\hat{\lambda}} \left\{ \iint_{u \cos(\theta - \theta_0) + v \sin(\theta - \theta_0) \leq \hat{\lambda}} f_{\theta-\theta_0}^{(0)}(u, v) dudv = \frac{S}{2} \right\}.$$

Thus, we have

$$\lambda_g^1(\theta) = \lambda_f^1(\theta - \theta_0). \quad (\text{A-31})$$

Similarly, we can obtain

$$\mu_g^1(\theta) = \mu_f^1(\theta - \theta_0). \quad (\text{A-32})$$

From Eq. (6), we have

$$g_{\theta}^{(1)}(x, y) = \begin{cases} g_{\theta}^{(0)}(x, y) & \text{if } x \cos \theta + y \sin \theta \leq \lambda_g^1(\theta) \\ 0 & \text{otherwise.} \end{cases} \quad (\text{A-33})$$

According to (A-31), the inequality $x \cos \theta + y \sin \theta \leq \lambda_g^1(\theta)$ can be rewritten as $u \cos(\theta - \theta_0) + v \sin(\theta - \theta_0) \leq \lambda_f^1(\theta - \theta_0)$, and (A-33) becomes

$$g_{\theta}^{(1)}(x, y) = \begin{cases} f_{\theta-\theta_0}^{(0)}(u, v) & \text{if } u \cos(\theta - \theta_0) + v \sin(\theta - \theta_0) \leq \lambda_f^1(\theta - \theta_0) \\ 0 & \text{otherwise.} \end{cases} \quad (\text{A-34})$$

From Eqs. (6) and (A-34), we have

$$\begin{aligned} g_{\theta}^{(1)}(x, y) &= f_{\theta-\theta_0}^{(1)}(u, v) \\ &= f_{\theta-\theta_0}^{(1)}(x \cos \theta_0 + y \sin \theta_0, -x \sin \theta_0 + y \cos \theta_0). \end{aligned} \quad (\text{A-35})$$

Till now, we have proved that the equations (A-1)~(A-9) hold for $i = 1$. Repeat the above process for $i = 2, 3, \dots$, iteratively, we can conclude that the equations (A-1)~(A-9) hold for all i . ■

Lemma 2: Let $g(x, y)$ be the linear distorted version (translation, scaling and rotation) of a shape image function $f(x, y)$. Their Structure Integral Transforms $\Psi_f(i, \theta)$ and $\Psi_g(i, \theta)$ have the following relationship.

(4) For translation:

$$\Psi_g(i, \theta) = \Psi_f(i, \theta), \quad (\text{A-36})$$

where $g(x, y) = f(x - x_0, y - y_0)$.

(5) For scaling:

$$\Psi_g(i, \theta) = \Psi_f(i, \theta), \quad (\text{A-37})$$

where $g(x, y) = f(ax, ay)$ and $a > 0$.

(6) For rotation:

$$\Psi_g(i, \theta) = \Psi_f(i, \theta - \theta_0), \quad (\text{A-38})$$

where $g(x, y) = f(x \cos \theta_0 + y \sin \theta_0, -x \sin \theta_0 + y \cos \theta_0)$.

Proof:

According to the definition of SIT, we have

$$\begin{aligned} \Psi_g(i, \theta) &= \frac{1}{\sqrt{S}} \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_{\theta}^{(i-1)}(x, y) \delta(x \cos \theta \right. \\ &\quad \left. + y \sin \theta - \lambda_g^i(\theta)) dxdy \right. \\ &\quad \left. + j \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_{\theta}^{(i-1)}(x, y) \delta(x \cos(\theta + \pi/2) \right. \\ &\quad \left. + y \sin(\theta + \pi/2) - \mu_g^i(\theta)) dxdy \right) \end{aligned} \quad (\text{A-39})$$

where

$$\iint_{-\infty}^{\infty} g(x, y) dxdy = S. \quad (\text{A-40})$$

(4) For translation:

From Lemma 1, we have

$$g_{\theta}^{(i-1)}(x, y) = f_{\theta}^{(i-1)}(x - x_0, y - y_0), \quad (\text{A-41})$$

$$\lambda_g^i(\theta) = \lambda_f^i(\theta) + x_0 \cos \theta + y_0 \sin \theta, \quad (\text{A-42})$$

and

$$\begin{aligned} \mu_g^i(\theta) &= \mu_f^i(\theta) + x_0 \cos(\theta + \frac{\pi}{2}) \\ &\quad + y_0 \sin(\theta + \frac{\pi}{2}). \end{aligned} \quad (\text{A-43})$$

Therefore Eq. (A-39) can be rewritten as

$$\begin{aligned} \Psi_g(i, \theta) &= \frac{1}{\sqrt{S}} \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\theta}^{(i-1)}(x - x_0, y - y_0) \delta((x - x_0) \cos \theta \right. \\ &\quad \left. + (y - y_0) \sin \theta - \lambda_f^i(\theta)) dx dy \right. \\ &\quad \left. + j \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\theta}^{(i-1)}(x - x_0, y - y_0) \delta((x - x_0) \cos(\theta + \pi/2) \right. \\ &\quad \left. + (y - y_0) \sin(\theta + \pi/2) - \mu_f^i(\theta)) dx dy \right), \end{aligned}$$

which can be simplified as

$$\begin{aligned} \Psi_g(i, \theta) &= \frac{1}{\sqrt{S}} \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\theta}^{(i-1)}(x, y) \delta(x \cos \theta \right. \\ &\quad \left. + y \sin \theta - \lambda_f^i(\theta)) dx dy \right. \\ &\quad \left. + j \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\theta}^{(i-1)}(x, y) \delta(x \cos(\theta + \pi/2) \right. \\ &\quad \left. + y \sin(\theta + \pi/2) - \mu_f^i(\theta)) dx dy \right). \quad (\text{A-44}) \end{aligned}$$

Since $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x - x_0, y - y_0) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = S$, we have

$$\Psi_g(i, \theta) = \Psi_f(i, \theta). \quad (\text{A-45})$$

(5) For scaling:

From Lemma 1, we have

$$g_{\theta}^{(i-1)}(x, y) = f_{\theta}^{(i-1)}(\alpha x, \alpha y), \quad (\text{A-46})$$

$$\lambda_g^i(\theta) = \frac{1}{\alpha} \lambda_f^i(\theta), \quad (\text{A-47})$$

and

$$\mu_g^i(\theta) = \frac{1}{\alpha} \mu_f^i(\theta). \quad (\text{A-48})$$

Therefore Eq. (A-39) can be rewritten as

$$\begin{aligned} \Psi_g(i, \theta) &= \frac{1}{\sqrt{S}} \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\theta}^{(i-1)}(\alpha x, \alpha y) \delta(x \cos \theta \right. \\ &\quad \left. + y \sin \theta - \frac{1}{\alpha} \lambda_f^i(\theta)) dx dy \right. \\ &\quad \left. + j \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\theta}^{(i-1)}(\alpha x, \alpha y) \delta(x \cos(\theta + \pi/2) \right. \\ &\quad \left. + y \sin(\theta + \pi/2) - \frac{1}{\alpha} \mu_f^i(\theta)) dx dy \right), \end{aligned}$$

which can be simplified as

$$\begin{aligned} \Psi_g(i, \theta) &= \frac{1}{\alpha \sqrt{S}} \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\theta}^{(i-1)}(x, y) \delta(x \cos \theta \right. \\ &\quad \left. + y \sin \theta - \lambda_f^i(\theta)) dx dy \right. \\ &\quad \left. + j \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\theta}^{(i-1)}(x, y) \delta(x \cos(\theta + \pi/2) \right. \\ &\quad \left. + y \sin(\theta + \pi/2) - \mu_f^i(\theta)) dx dy \right). \quad (\text{A-49}) \end{aligned}$$

Since

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) dx dy &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\alpha x, \alpha y) dx dy \\ &= \frac{1}{\alpha^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = S \end{aligned}$$

we have

$$\Psi_g(i, \theta) = \Psi_f(i, \theta). \quad (\text{A-50})$$

(6) For rotation:

Let $u = x \cos \theta_0 + y \sin \theta_0$ and $v = -x \sin \theta_0 + y \cos \theta_0$. Since $u^2 + v^2 = x^2 + y^2$, we have $du dv = dx dy$. According to Lemma 1, we have

$$g_{\theta}^{(i-1)}(x, y) = f_{\theta - \theta_0}^{(i-1)}(u, v), \quad (\text{A-51})$$

$$\lambda_g^i(\theta) = \lambda_f^i(\theta - \theta_0), \quad (\text{A-52})$$

and

$$\mu_g^i(\theta) = \mu_f^i(\theta - \theta_0). \quad (\text{A-53})$$

Therefore Eq. (A-39) can be rewritten as

$$\begin{aligned} \Psi_g(i, \theta) &= \frac{1}{\sqrt{S}} \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\theta - \theta_0}^{(i-1)}(u, v) \delta(u \cos(\theta - \theta_0) \right. \\ &\quad \left. + v \sin(\theta - \theta_0) - \lambda_f^i(\theta - \theta_0)) du dv \right. \\ &\quad \left. + j \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\theta - \theta_0}^{(i-1)}(u, v) \delta(u \cos((\theta - \theta_0) + \pi/2) \right. \\ &\quad \left. + v \sin((\theta - \theta_0) + \pi/2) - \mu_f^i(\theta - \theta_0)) du dv \right). \quad (\text{A-54}) \end{aligned}$$

Since

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u, v) du dv = S,$$

we have

$$\Psi_g(i, \theta) = \Psi_f(i, \theta - \theta_0).$$

■

REFERENCES

- [1] Y. W. Chen and Y. Q. Chen, "Invariant description and retrieval of planar shapes using radon composite features," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4762–4771, Oct. 2008.
- [2] T. V. Hoang and S. Tabbone, "Invariant pattern recognition using the RFM descriptor," *Pattern Recognit.*, vol. 45, no. 1, pp. 271–284, Jan. 2012.
- [3] P.-T. Yap, X. Jiang, and A. C. Kot, "Two-dimensional polar harmonic transforms for invariant image representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 7, pp. 1259–1270, Jul. 2010.
- [4] X. Wang, B. Xiao, J.-F. Ma, and X.-L. Bi, "Scaling and rotation invariant analysis approach to object recognition based on radon and Fourier—Mellin transforms," *Pattern Recognit.*, vol. 40, no. 12, pp. 3503–3508, Dec. 2007.
- [5] S. R. Deans, "Hough transform from the radon transform," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 3, no. 2, pp. 185–188, Mar. 1981.
- [6] L. J. Latecki, R. Lakamper, and T. Eckhardt, "Shape descriptors for non-rigid shapes with a single closed contour," in *Proc. CVPR*, vol. 1, 2000, pp. 424–429.
- [7] E. Milios and E. G. M. Petrakis, "Shape retrieval based on dynamic programming," *IEEE Trans. Image Process.*, vol. 9, no. 1, pp. 141–147, Jan. 2000.

- [8] E. G. M. Petrakis, A. Diplaros, and E. Milios, "Matching and retrieval of distorted and occluded shapes using dynamic programming," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 11, pp. 1501–1516, Nov. 2002.
- [9] D. Zhang and G. Lu, "Review of shape representation and description techniques," *Pattern Recognit.*, vol. 37, no. 1, pp. 1–19, Jan. 2004.
- [10] H. Ling and D. W. Jacobs, "Shape classification using the inner-distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 2, pp. 286–299, Feb. 2007.
- [11] 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, Apr. 2002.
- [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] P. Sidiropoulos, S. Vrochidis, and I. Kompatsiaris, "Content-based binary image retrieval using the adaptive hierarchical density histogram," *Pattern Recognit.*, vol. 44, no. 4, pp. 739–750, Apr. 2011.
- [14] S. R. Deans, *The Radon Transform and Some of its Applications*. New York, NY, USA: Wiley, 1983.
- [15] A. Srivastava, S. H. Joshi, W. Mio, and X. Liu, "Statistical shape analysis: Clustering, learning, and testing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 4, pp. 590–602, Apr. 2005.
- [16] R. Gal, A. Shamir, and D. Cohen-Or, "Pose-oblivious shape signature," *IEEE Trans. Vis. Comput. Graph.*, vol. 13, no. 2, pp. 261–271, Mar./Apr. 2007.
- [17] Y.-S. Liu, K. Ramani, and M. Liu, "Computing the inner distances of volumetric models for articulated shape description with a visibility graph," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2538–2544, Dec. 2011.
- [18] J. Neumann, H. Samet, and A. Soffer, "Integration of local and global shape analysis for logo classification," *Pattern Recognit. Lett.*, vol. 23, no. 12, pp. 1449–1457, Oct. 2002.
- [19] P. Korn, N. Sidiropoulos, C. Faloutsos, E. Siegel, and Z. Protopoulos, "Fast and effective retrieval of medical tumor shapes," *IEEE Trans. Knowl. Data Eng.*, vol. 10, no. 6, pp. 889–904, Nov. 1998.
- [20] Y. Wang, K. Huang, and T. Tan, "Human activity recognition based on R transform," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.
- [21] D. D. Hoffman and W. A. Richards, "Parts of recognition," *Cognition*, vol. 18, nos. 1–3, pp. 65–96, 1984.
- [22] S. Vrochidis, S. Papadopoulos, A. Moutzidou, P. Sidiropoulos, E. Pianta, and I. Kompatsiaris, "Towards content-based patent image retrieval: A framework perspective," *World Patent Inf. J.*, vol. 32, no. 2, pp. 94–106, Jun. 2010.
- [23] C.-H. Wei, Y. Li, W.-Y. Chau, and C.-T. Li, "Trademark image retrieval using synthetic features for describing global shape and interior structure," *Pattern Recognit.*, vol. 42, no. 3, pp. 386–394, Mar. 2009.
- [24] H.-K. Kim and J.-D. Kim, "Region-based shape descriptor invariant to rotation, scale and translation," *Signal Process. Image Commun.*, vol. 16, nos. 1–2, pp. 87–93, Sep. 2000.
- [25] K. Jafari-Khouzani and H. Soltanian-Zadeh, "Radon transform orientation estimation for rotation invariant texture analysis," *IEEE Trans. Vis. Comput. Graph.*, vol. 27, no. 6, pp. 1004–1008, Jun. 2005.
- [26] J. S. Seo, J. Haitisma, T. Kalker, and C. D. Yoo, "A robust image fingerprinting system using the Radon transform," *Signal Process., Image Commun.*, vol. 19, no. 4, pp. 325–339, 2004.
- [27] Y. Lei, Y. Wang, and J. Huang, "Robust image hash in Radon transform domain for authentication," *Signal Process., Image Commun.*, vol. 26, no. 6, pp. 280–288, 2011.
- [28] F. Hjouj and D. W. Kammiller, "Identification of reflected, scaled, translated, and rotated objects from their radon projections," *IEEE Trans. Image Process.*, vol. 17, no. 3, pp. 301–310, Mar. 2008.
- [29] S. Loncaric, "A survey of shape analysis techniques," *Pattern Recognit.*, vol. 31, no. 8, pp. 983–1001, Aug. 1998.
- [30] R. C. Velkamp and M. Hagedoorn, "State of the Art in Shape Matching," Dept. Inf. Comput. Sci., Utrecht Uni., Utrecht, The Netherlands, Tech. Rep. UU-CS-1999-27, 1999.
- [31] C. Direkoglou and M. S. Nixon, "Shape classification via image-based multiscale description," *Pattern Recognit.*, vol. 44, no. 9, pp. 2134–2146, Sep. 2011.
- [32] C. T. Zahn and R. Z. Roskies, "Fourier descriptors for plane closed curves," *IEEE Trans. Comput.*, vol. 21, no. 3, pp. 269–281, Mar. 1972.
- [33] E. Persoon and K.-S. Fu, "Shape discrimination using Fourier descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 8, no. 3, pp. 388–397, May 1986.
- [34] G.-C. H. Chuang and C.-C. J. Kuo, "Wavelet descriptor of planar curves: Theory and applications," *IEEE Trans. Image Process.*, vol. 5, no. 1, pp. 56–70, Jan. 1996.
- [35] Q. M. Tieng and W. W. Boles, "Recognition of 2D object contours using the wavelet transform zero-crossing representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 8, pp. 910–916, Aug. 1997.
- [36] H. S. Yang, S. U. Lee, and K. M. Lee, "Recognition of 2D object contours using starting-point-independent wavelet coefficient matching," *J. Vis. Commun. Image Represent.*, vol. 9, no. 2, pp. 171–181, Jun. 1998.
- [37] M. Pelillo, K. Siddiqi, and S. W. Zucker, "Matching hierarchical structures using association graphs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 11, pp. 1105–1120, Nov. 1999.
- [38] K. Siddiqi, A. Shokoufandeh, S. J. Dickinson, and S. W. Zucker, "Shock graphs and shape matching," *Int. J. Comput. Vis.*, vol. 35, no. 1, pp. 13–32, Nov. 1999.
- [39] T.-L. Liu and D. Geiger, "Approximate tree matching and shape similarity," in *Proc. Int. Conf. Comput. Vis.*, Sep. 1999, pp. 456–462.
- [40] R.-X. Hu, W. Jia, Y. Zhao, and J. Gui, "Perceptually motivated morphological strategies for shape retrieval," *Pattern Recognit.*, vol. 45, no. 9, pp. 3222–3230, Sep. 2012.
- [41] P. F. Felzenszwalb and J. D. Schwartz, "Hierarchical matching of deformable shapes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.
- [42] N. Alajlan, M. S. Kamel, and G. H. Freeman, "Geometry-based image retrieval in binary image databases," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 6, pp. 1003–1013, Jun. 2008.
- [43] T. Adamek and N. E. O'Connor, "A multiscale representation method for nonrigid shapes with a single closed contour," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 5, pp. 742–753, May 2004.
- [44] J. Wang, X. Bai, X. You, W. Liu, and L. J. Latecki, "Shape matching and classification using height functions," *Pattern Recognit. Lett.*, vol. 33, no. 2, pp. 134–143, Jan. 2012.
- [45] G. McNeill and S. Vijayakumar, "Hierarchical procrustes matching for shape retrieval," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2006, pp. 885–894.
- [46] A. Goshtasby, "Description and discrimination of planar shapes using shape matrices," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 7, no. 6, pp. 738–743, Nov. 1985.
- [47] G. Lu and A. Sajjanhar, "Region-based shape representation and similarity measure suitable for content-based image retrieval," *Multimedia Syst.*, vol. 7, no. 2, pp. 165–174, Mar. 1999.
- [48] M. Yang, G. Qiu, J. Huang, and D. Elliman, "Near-duplicate image recognition and content-based image retrieval using adaptive hierarchical geometric centroids," in *Proc. 18th Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2006, pp. 958–961.
- [49] C. Bauckhage, "Tree-based signatures for shape classification," in *Proc. Int. Conf. Image Process. (ICIP)*, Oct. 2006, pp. 2105–2108.
- [50] S. Tabbone, L. Wendling, and J.-P. Salmon, "A new shape descriptor defined on the radon transform," *Comput. Vis. Image Understand.*, vol. 102, no. 1, pp. 42–51, Apr. 2006.
- [51] T. V. Hoang and S. Tabbone, "The generalization of the R-transform for invariant pattern representation," *Pattern Recognit.*, vol. 45, no. 6, pp. 2145–2163, Jun. 2012.
- [52] Z. A. Khan and W. Sohn, "Abnormal human activity recognition system based on r-transform and kernel discriminant technique for elderly home care," *IEEE Trans. Consum. Electron.*, vol. 57, no. 4, pp. 1843–1850, Nov. 2011.
- [53] S. A. Nene, S. K. Nayar, and H. Murase, "Columbia object image library (COIL-20)," Dept. Comput. Sci., Columbia Univ., New York, NY, USA, Tech. Rep. CUCS-005-96, 1996.
- [54] J. Duchon, "Splines minimizing rotation-invariant semi-norms in Sobolev space," *Constructive Theory Functions Several Variables*, W. Schempp K. Zeller, Eds, Berlin, Germany: Springer-Verlag, 1977, pp. 85–100.
- [55] J. Meinguet, "Multivariate interpolation at arbitrary points made simple," *J. Appl. Math. Phys.*, vol. 30, no. 2, pp. 292–304, 1979.



Bin Wang received the Ph.D. degree in computer science from Fudan University, Shanghai, China, in 2007. Since 2007, he had been with the Faculty of Nanjing University of Finance and Economics, Nanjing, China, where he became a Professor in 2015. From 2007 to 2011, he was a Post-Doctoral Research Fellow (part time) with Southeast University, Nanjing. From 2011 to 2012, he was a Visiting Scholar with the School of Engineering, Griffith University, Australia. In 2015, he joined the School of Engineering, Griffith University, where he is currently a Research Fellow. His main research interests include computer vision, image processing, and pattern recognition.



Yongsheng Gao received the B.Sc. and M.Sc. degrees in electronic engineering from Zhejiang University, Hangzhou, China, in 1985 and 1988, respectively, and the Ph.D. degree in computer engineering from Nanyang Technological University, Singapore. He is currently a Professor with the School of Engineering, Griffith University, Brisbane, QLD, Australia. He had been the Leader of Biosecurity Group, Queensland Research Laboratory, National ICT Australia (ARC Centre of Excellence), a Consultant of Panasonic Singapore Laboratories, and an Assistant Professor with the School of Computer Engineering, Nanyang Technological University. His research interests include face recognition, biometrics, biosecurity, image retrieval, computer vision, pattern recognition, environmental informatics, and medical imaging.