A Novel Line Integral Transform for 2D Affine-Invariant Shape Retrieval

Bin Wang[1,2(✉)] and Yongsheng Gao[1]

[1] Griffith University, Nathan, QLD 4111, Australia
{bin.wang,yongsheng.gao}@griffith.edu.au
[2] Nanjing University of Finance & Economics, Nanjing 210023, China

Abstract. Radon transform is a popular mathematical tool for shape analysis. However, it cannot handle affine deformation. Although its extended version, trace transform, allow us to construct affine invariants, they are less informative and computational expensive due to the loss of spatial relationship between trace lines and the extensive repeated calculation of transform. To address this issue, a novel line integral transform is proposed. We first use binding line pairs that have the desirable property of affine preserving as a reference frame to rewrite the diametrical dimension parameters of the lines in a relative manner which make them independent on affine transform. Along polar angle dimension of the line parameters, a moment-based normalization is then conducted to degrade the affine transform to similarity transform which can be easily normalized by Fourier transform. The proposed transform is not only invariant to affine transform, but also preserves the spatial relationship between line integrals which make it very informative. Another advantage of the proposed transform is that it is more efficient than the trace transform. Conducting it one time can allow us to achieve a 2D matrix of affine invariants. While conducting the trace transform once only generates a single feature and multiple trace transforms of different functionals are needed to derive more to make the descriptors informative. The effectiveness of the proposed transform has been validated on two types of standard shape test cases, affinely distorted contour shape dataset and region shape dataset, respectively.

Keywords: Shape analysis · Affine distortions · Affine invariants · Radon transform · Trace transform

1 Introduction

Shape analysis is an active research area in the computer vision and pattern recognition community and has a large body of potential applications such as human activity recognition [30], target tracking [31], medical image retrieval [32], etc. The object images captured by the camera are generally subject to various deformations. One of the typical distortions is perspective transform which occurs on the situation of the pictures of the objects are taken under arbitrary

orientations. The computer vision systems are also expected to effectively handle the recognition of perspective distorted shapes. Under certain circumstance like the distance between the camera and the object being far enough, or the thick of the object being very small that can be approximated a plane, the perspective transform can be well approximated by affine transformation [22][33]. In this paper, we focus on recognizing shapes distorted by affine transform.

Conducting line integrals over the image plane and transform the image function to another 2D function of line parameters (θ, λ) is an effective way for shape analysis. This is one of the significant applications of the well-known Radon transform [1]. The main appealing characteristics of the Radon transform is that it benefits extracting feature of the inner structure of the object and the image function of the object can be fully reconstructed from its Radon transform.

Numerous efforts have been made on extracting invariant shape features from Radon transform for object recognition [34][35][36][37][38]. However, these methods can only handle similar transform (translation, rotation and scaling) which is a kind of shape-preserving transformation and is a subset of larger affine group. Although affine transform is line-preserving, the complex behaviours of the line integrals of the image function make the affine transform challenging to extract affine invariants from the domain of the line integral transform.

Trace transform [39] is a generalized version of the Radon transform. It extends the line integral to any 1D functional along lines (termed trace functional). So, various functionals used will derive different trace transform. However, trace transform is still sensitive to affine transform because the functionals used along the line can not yield invariants. To achieve affine invariants, Petrou and Kadyrov [5] propose to further conduct diametrical functionals along the dimension of λ and circus functionals along the dimension of θ. It is worth noting that in this method, a set of functionals including a trace functional, a diametrical functionals and a circus functionals can only generate one invariant feature. So, to make the shape descriptors informative, many triples of functionals have to be developed for yielding more affine invariants. However, it is very difficult to find appropriate triple of functionals to constructing the desirable invariants and the expensive computational cost also make it not suitable for real applications.

In this paper, we propose a novel integral transform that can effectively and efficiently handle affine-distorted shapes. The proposed transform has the following advantages over the trace transform: (1) It is much more discriminative than the trace-transform based method [5]. The later uses diametrical functionals and circus functionals to achieve affine invariants which cannot preserve the information of diametrical dimension and circus dimension of the trace lines. While the proposed transform perfectly preserves it; (2) It has the higher efficiency to yield affine invariants than the trace transform. Conducting it one time can allow us to achieve a group of shape invariants. While conducting the trace transform once only generates a single feature and multiple trace transforms of different functionals are needed to derive more to make the descriptors informative; (3) The features obtained by the proposed transform has a physical interpretation,

while the features obtained from the trace transform generally has no clearly physical meaning.

2   Related Work

The existing affine shape analysis methods can be categorized into two groups. One is contour based methods which represent the object boundary as ordered points or parameterized curves. The other is region based methods which attempt to generate affine invariants from the whole shape region.

A large body of methods have made attempts to model the silhouette of the shape for characterizing the behavior of the shapes that are subject to affine transform. The popular ideas of them model the silhouette as a parameterized curve. Then some mathematical tools can be used for analysis the geometric properties. The earlier work [22] applied Fourier transform to the affine-length parameterized boundary description for eliminating its dependency on the affine transformation. Various wavelet transforms with different wavelet basis functions are also utilized to generate affine shape invariants [24][25].

B-spline is a kind of continuous curve representation which make it very suitable for shape analysis. Wang and Teoh [29] modeled the shape contour using B-spline to construct the Curvature Scale Space (CSS) image for affine invariant shape matching. Huang and Cohen [19] proposed a fast algorithm for estimating the B-spline control points and used a new class of weighted B-spline curve moments to handle the affine transformation between curves. Various algebraic curve models such as quartic implicit polynomials [16] and conic curve [21] were proposed for extracting affine or perspective invariants. Zuliani et al. [10] used the area, centroid, and covariance of the domain enclosed by the curve to normalize the shape for removing the effect of affine transform.

There are also many methods treating the silhouette of the shape as a sequence of order points to construct shape descriptors or building corresponding to matching shapes. Mai et al. [13] represented shape as a sequence of chained edge points and proposed to project one shape onto the subspace spanned by the other. The two shapes are then matched by minimizing a subspace projection error. This method has a clear physical interpretation and works very fast for estimating the affine transform. Jia et al. [14] developed a new projective invariant, the characteristic number (CN) whose values is calculated on a series of five sample points along the shape contour. With the sample intervals varying, a coarse to fine strategy is developed for capturing both the global geometry described by projective invariants and the local contextual information. Xue et al. [27] proposed a fuzzy algorithm for aligning shapes under affine transform. This algorithm can efficiently estimate the point correspondence and the relevant affine parameters. Recently, Bryner et al. [7][8] presented a contour-based shape analysis framework based on Riemannian geometry that is invariant to affine transform and re-parameterization of contours. Three shape space, landmark-affine, curve-affine and landmark projective are studied in their work.

Different from the contour based methods, the region based shape analysis approaches characterize shape considering all the pixels over the shape domain. Moment invariants are the popular region based shape descriptors. In the earlier work, Flusser and Suk [12] introduced affine moment invariants for object recognition. Heikkilä [23] used the second and higher order moments of the image points for both recognition and alignment under affine transform. The advantage of this method is that it does not require to know any point correspondences in advance. Yang and Cohen [28] proposed a framework for deriving a class of new global affine invariants based on a novel concept of cross-weighted moments with fractional weights.

Besides the moment descriptors, many recent works developed various strategies for region based shape recognition and matching. Domokos et al. [15] proposed a novel parametric estimation of affine deformations for planar shapes. Instead of finding the correspondences between the landmarks of the template shape and target shape for computing the affine parametric, they treat the affine parametric estimation as a solution of polynomial equations in which all the information available in the input image is used. More recently, Ruiz et al. [26] proposed a fast and accurate affine canonicalization method for recognizing and matching planar shapes. Different from many shape analysis methods which extract invariant features for recognition, this work attempts to produce multiple canonical versions of the shape for provides a common reference frame for accurate shape comparison.

The works that are most relevant to our research in this paper are various transform based methods. Ruiz et al. [11] proposed a multiscale autoconvolution (MSA) transform based on a probabilistic interpretation of the image function. The MSA transform is a 2D function derived from the shape image function which can present infinitely many affine invariant features by varying its two variable. So, it can be directly applied for shape recognition. The trace transform is a generalization of the Radon transform. Petrou and Kadyrov [5] proposed to conduct several trace functionals on the original image function firstly to obtain the same number of trace transforms of the image function, then for each available trace transform, several diametrical functionals are used to transform them to the same number of 1D functions of directional angle of line. The affine invariants are finally achieved by further performing circus functionals to the available 1D functions. The number of the available affine invariant features is the number of the combinations of the used three types of functionals. It can be directly used for shape recognition and its theory can be extended for affine parametric estimation [29]. Recently, Zhang and Chu [30] developed a ray projection transform and apply it for recovering a geometric transformation and an affine lighting change between two objects.

3   Affine Theory of Line Integral

Given a 2D function $f(\mathrm{x})$ and a straight line ℓ with equation $\lambda - \overrightarrow{\theta}^T\mathrm{x}=0$, where $\overrightarrow{\theta} = (cos\theta, sin\theta)^T$ is a unit vector in the direction of the normal to the line ℓ,

and x$=(x,y)^T \in \mathbb{R}^2$ denote the coordinates of a point of the line ℓ. The integral of the function $f(\mathrm{x})$ over the line ℓ can be mathematically expressed as

$$\Re_f(\lambda, \theta) = \int f(\mathrm{x})\delta(\lambda - \overrightarrow{\theta}^T\mathrm{x})d\mathrm{x}, \tag{1}$$

where $\delta(\cdot)$ denotes a Dirac delta function. Through the line integral, the function $f(\mathrm{x})$ is transformed into another 2D function $\Re_f(\lambda, \theta)$ of line parameters (λ, θ) (in the later sections, we directly use the parameters (λ, θ) to denote a line). This is the well-known Radon transform [1] that has been widely applied for image analysis. In this section, we study the behavior of the line integral of the function $f(\mathrm{x})$ under affine transform.

An affine transform $\mathcal{H} = \mathcal{H}(A, s)$ can be defined as $\mathrm{x}' = A\mathrm{x} + s$, where A is a 2×2 nonsingular real matrix and $s \in \mathbb{R}^2$.Using it, a function $f(\mathrm{x})$ can be transformed to another function $g(\mathrm{x}) = f(A^{-1}\mathrm{x} - A^{-1}s)$.The relationship between their Radon transformed versions,$\Re_f(\lambda, \theta)$ and $\Re_g(\lambda, \theta)$,can be deduced as

$$\Re_g(\lambda, \theta) = \int g(\mathrm{x})\delta(\lambda - \overrightarrow{\theta}^T\mathrm{x})d\mathrm{x}$$
$$= \int f(A^{-1}\mathrm{x} - A^{-1}s)\delta(\lambda - \overrightarrow{\theta}^T\mathrm{x})d\mathrm{x}. \tag{2}$$

Let y $= A^{-1}\mathrm{x} - A^{-1}s$, Eq.(2) can be rewritten as

$$\Re_g(\lambda, \theta) = |det(A)| \int f(\mathrm{y})\delta(\lambda - \overrightarrow{\theta}^T s - (A^T\overrightarrow{\theta})^T\mathrm{y})d\mathrm{y}. \tag{3}$$

Note that the vector $A^T\overrightarrow{\theta}$ is a transformed version of the vector $\overrightarrow{\theta}$ and may not be a unit vector. Here, we use the scaling property [2] of the Dirac delta function, that is $\delta(\alpha x) = |\alpha|^{-1}\delta(x)$ for $\alpha \neq 0$, to normalize it and Eq. (3) can be rewritten as

$$\Re_g(\lambda, \theta) = \frac{|det(A)|}{\left\|A^T\overrightarrow{\theta}\right\|} \int f(\mathrm{y})\delta\left(\frac{\lambda}{\left\|A^T\overrightarrow{\theta}\right\|} - \frac{\overrightarrow{\theta}^T s}{\left\|A^T\overrightarrow{\theta}\right\|} - \frac{(A^T\overrightarrow{\theta})^T}{\left\|A^T\overrightarrow{\theta}\right\|}\mathrm{y}\right) d\mathrm{y}, \tag{4}$$

where $\|\cdot\|$ denotes the length of the vector. Defining following functions of variable θ:

$$\epsilon(\theta) = \frac{|det(A)|}{\left\|A^T\overrightarrow{\theta}\right\|}, \zeta(\theta) = \frac{1}{\left\|A^T\overrightarrow{\theta}\right\|}, \eta(\theta) = -\frac{\overrightarrow{\theta}^T s}{\left\|A^T\overrightarrow{\theta}\right\|},$$

and

$$\rho(\theta) = \left\langle A^T\overrightarrow{\theta}\right\rangle, \tag{5}$$

where $\langle\cdot\rangle$ denotes the direction angle of the vector, we then have

$$\Re_g(\lambda, \theta) = \epsilon(\theta) \int f(\mathbf{y}) \, \delta\left(\zeta(\theta) \cdot \lambda + \eta(\theta) - \overrightarrow{\rho(\theta)}^T \mathbf{y}\right) d\mathbf{y}$$
$$= \epsilon(\theta) \cdot \Re_f(\zeta(\theta) \cdot \lambda + \eta(\theta), \rho(\theta)). \tag{6}$$

The above equation indicates the following effects of the affine transform $\mathcal{H}(A, s)$ on the line integral of the function $f(\mathbf{x})$: (1)its amplitude is scaled by $\epsilon(\theta)$; (2)its parameter λ is scaled by $\zeta(\theta)$ and shifted by $\eta(\theta)$, and (3) its parameter θ is transformed to be $\rho(\theta)$.

   Let the inverse transform of $\mathcal{H}(A, s)$ be $\mathcal{H}^{-1}(A^{-1}, -A^{-1}s)$. The inverse transformed versions $\epsilon^{-1}(\theta)$, $\zeta^{-1}(\theta)$, $\eta^{-1}(\theta)$, and $\rho^{-1}(\theta)$ of the functions $\epsilon(\theta)$, $\zeta(\theta)$, $\eta(\theta)$, and $\rho(\theta)$ can be defined by replacing A and s appeared in the Eq. (5) with A^{-1} and $-A^{-1}s$, respectively. From Eq. (6), we can also conclude that when a line (λ, θ) is subject to the affine transform $\mathcal{H}(A, s)$, it will become the line $(\zeta^{-1}(\theta) \cdot \lambda + \eta^{-1}(\theta), \rho^{-1}(\theta))$. In the next section, we will use the affine theories of the line and line integral under affine transform to construct a novel line integral transform for affine shape analysis.

4   The Proposed line Integral Transform

A shape is defined as region D that is a subset of pixels in the image plane $\mathbb{R}^2$ [3]. The shape image function $f(\mathbf{x})$ can then be defined an indicator function $f(\mathbf{x})=1$ if $x \in D$ and $f(\mathbf{x})=0$ otherwise.

4.1   Binding Line Pair and Its Affine Property

Using the line integral $\Re_f(\lambda, \theta)$ of the function $f(\mathbf{x})$, we define a 1D function of the variable θ as

$$\sigma_f(\theta) = arg\,\min_{\lambda}\left\{\Re_f(\lambda, \theta) > 0\right\}. \tag{7}$$

Then for an angle $\theta \in [0, 2\pi]$,we can uniquely derive a line pair $(\sigma_f(\theta), \theta)$ and $(-\sigma_f(\theta + \pi), \theta)$.It can be easily concluded that they have the following relationships with the shape region D:

$$\exists\, x \in D, \sigma_f(\theta) - \overrightarrow{\theta}^T \mathbf{x} = 0. \tag{8}$$

and

$$\forall\, x \in D, \sigma_f(\theta) \leq \overrightarrow{\theta}^T \mathbf{x} \leq -\sigma_f(\theta + \pi). \tag{9}$$

The Eq. (8) and Eq. (9) indicate that the shape region D is located between the line pair $(\sigma_f(\theta), \theta)$ and $(-\sigma_f(\theta + \pi), \theta)$ and has at least one intersection point with them which also means that the shape region D is bound by the line pair. We term the line $(\sigma_f(\theta), \theta)$ and $(-\sigma_f(\theta+\pi), \theta)$ as binding line pair. An example of binding line pair is presented in Fig. 1.
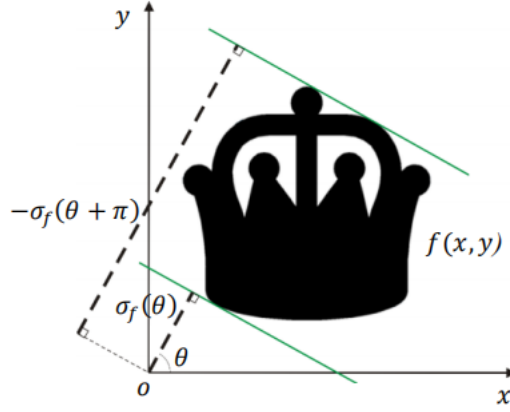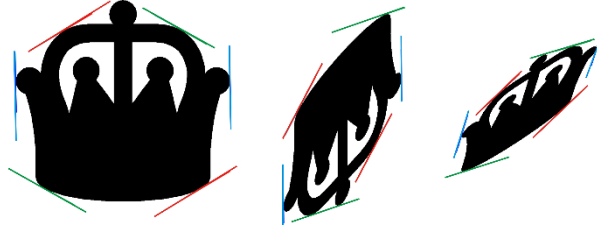
Fig. 1. An example of a binding line pair for a shape (marked in green color).

Now, we analysis the property of the binding line pair under the affine transform $\mathcal{H}(A, s)$.According to Eq. (7), we can conclude that the function $\sigma_f(\theta)$ has the following relationship with the function $\sigma_g(\theta)$:

$$\sigma_f(\theta) = \frac{\sigma_g(\rho^{-1}(\theta)) - \eta^{-1}(\theta)}{\zeta^{-1}(\theta)}.$$

(10)

Under the affine transform $\mathcal{H}(A, s)$, the line $(\sigma_f(\theta), \theta)$ is transformed to be the line $(\zeta^{-1}(\theta) \cdot \sigma_f(\theta) + \eta^{-1}(\theta), \rho^{-1}(\theta))$ which can be rewritten as $(\sigma_g(\rho^{-1}(\theta)), \rho^{-1}(\theta))$ in terms of Eq. (10). Therefore, we can conclude that the affine transformed version of a binding line of the shape is the binding line of the affine transformed version of the shape. Note that the parameters $(-\sigma_f(\theta + \pi), \theta)$ and the parameters $(\sigma_f(\theta + \pi), \theta + \pi)$ represent the same line. Thus, another binding line $(-\sigma_f(\theta + \pi), \theta)$ also keeps its binding property under the affine transform. A graphical illustration of this property is shown in Fig. 2.

Fig. 2. An example to indicate the binding line pairs (marked by the same color for each) having the property of being affine transform preserved, i.e. the affine transformed version of a binding line pair of the shape is the binding pair of the affine transformed version of the shape.

4.2   The Proposed Transform

Given an image function $f(\mathrm{x})$ and a direction angle $\theta \in [0, 2\pi)$, a binding line pair $(\sigma_f(\theta), \theta)$ and $(-\sigma_f(\theta+\pi), \theta)$ can be calculated. The distance between them is $-\sigma_f(\theta + \pi) - \sigma_f(\theta)$. The line that has equal distance to the binding line pair, i.e., the center line, can be represented as $((\sigma_f(\theta) - \sigma_f(\theta + \pi))/2, \theta)$. Taking the center line and the binding line $(\sigma_f(\theta), \theta)$ as the reference lines, each line that is located between them can be then uniquely represented as the parameter form of $(\Gamma_f(\mu, \theta), \theta)$. Where $\theta \in [0, 2\pi]$ is the direction angle of the line which is also the direction angle of the reference lines and the parameter $\Gamma_f(\mu, \theta)$ is a 2D function of the variables (μ, θ) defined by

$$\Gamma_f(\mu, \theta) = \frac{\sigma_f(\theta) - \sigma_f(\theta + \pi)}{2} - \mu \cdot \frac{-\sigma_f(\theta + \pi) - \sigma_f(\theta)}{2}, \qquad (11)$$

where $\mu \in [0, 1]$ is the ratio of the distance between the line with the reference line $(\sigma_f(\theta), \theta)$ to the distance between the two reference lines, $(\sigma_f(\theta), \theta)$ and $((\sigma_f(\theta) - \sigma_f(\theta + \pi))/2, \theta)$.

Let the parameter θ vary from 0 to 2π and the parameter μ vary from 0 to 1, all the lines that go through the shape region D can then be available to form a line set denoted by

$$\{(\Gamma_f(\mu, \theta), \theta) : \mu \in [0, 1], \theta \in [0, 2\pi)\}. \qquad (12)$$

We integral the shape image function $f(\mathrm{x})$ over each line in the above set and obtain a novel integral transform for the function $f(\mathrm{x})$ as follows:

$$\Psi_f(\mu, \theta) = \Re_f(\Gamma_f(\mu, \theta), \theta). \qquad (13)$$

which is a 2D function of the parameters $\mu \in [0, 1]$ and $\theta \in [0, 2\pi]$.

The proposed transform $\Psi_f(\mu, \theta)$ has the following property under the affine transform $\mathcal{H} = \mathcal{H}(A, s)$ of the shape image function $f(\mathrm{x})$:

$$\Psi_g(\mu, \theta) = \epsilon(\theta) \cdot \Psi_f(\mu, \rho(\theta)). \qquad (14)$$

Comparing the above equation with Eq. (6), we can find the difference between the proposed transform with the Radon transform under affine transform as follows: the first parameter λ for the Radon transform embeds the parameters of affine transform, while the first parameter μ for the proposed transform is independent of the affine transform. However, Eq. (14) also indicates that the second parameter θ and the amplitude of the proposed transform still encode the affine transform parameters. Since the parameter μ of the proposed transform is independent of the affine transform, we fix it and rewrite the function $\Psi_f(\mu, \theta)$ as the function $\Psi_f^\mu(\theta)$ which has only one variable θ. Delimited by the function $\Psi_f^\mu(\theta)$, a region W_f^μ can be derived and mathematically defined as

$$W_f^\mu = \left\{ w \in \mathbb{R}^2 : \exists \theta \in [0, 2\pi] \wedge \exists r \in [0, 1] such\ that\ w = (r \cdot \Psi_f^\mu(\theta)) \overrightarrow{\theta} \right\}. \quad (15)$$

Similarly, a region W_g^μ that is delimited by the function $\Psi_g^\mu(\theta)$ can be defined by replacing f with g in Eq. (15). We are interested in the relationship between the available regions W_f^μ and W_g^μ. For any $w' \in W_g^\mu$, there exists $w \in W_f^\mu$ such that

$$w' = |det(A)|A^{-T}w. \tag{16}$$

and on the other hand, for any $w \in W_f^\mu$, there exists $w' \in W_g^\mu$ which also makes Eq. (16) hold, where A^{-T} denotes the transpose matrix of A^{-1}. Therefore, we can deduce that the region W_f^μ and the region W_g^μ are correlated by the affine transform $\mathcal{H}' = \mathcal{H}'(|det(A)|A^{-T}, 0)$. This is a desirable property which allows us to use their respective second-order geometric moment to derive a transformation for reducing the relation between them from an affinity into a similarity. Similar ideas can be found in [4][5].

For the region W_f^μ, we construct its second-order geometric moment matrix as

$$M_f^\mu = \int_{W_f^\mu} ww^T dw. \tag{17}$$

Similarly, the second-order geometric moment matrix M_g^μ for the region W_g^μ can be available. For the integration expression of calculating the moment matrix M_g^μ, by making following changes: the variable $w' = |det(A)|A^{-T}w$, the derivative $dw' = |det(A)|dw$, and the integration region $W_g^\mu \rightarrow M_f^\mu$, we can easily deduce that

$$M_g^\mu = |det(A)|^3 A^{-T} M_f^\mu A^{-1}. \tag{18}$$

which indicates the relationship between the moment matrix M_f^μ and the moment matrix M_g^μ. Define $M^{1/2}$ as any matrix that satisfies $(M^{1/2})(M^{1/2})^T = M$. Since $det(M) > 0$, the $M^{1/2}$ can be always achieved. It can be calculated by using an eigenvalue method [5][6]. Let $\Delta_f = (M_f^\mu)^{1/2}$ and $\Delta_g = (M_g^\mu)^{1/2}$, we accordingly achieve $M_f^\mu = \Delta_f \Delta_f^T$ and $M_g^\mu = \Delta_g \Delta_g^T$. Using the matrices Δ_f and Δ_g, we can produce a transformation matrix as

$$E = \Delta_g^{-1}(|det(A)|A^{-T})\Delta_f. \tag{19}$$

Then we have

$$EE^T = \Delta_g^{-1}(|det(A)|A^{-T})\Delta_f \Delta_f^T (|det(A)|A^{-1})\Delta_g^{-T}. \tag{20}$$

Using Eq. 18, the above equation can be rewritten as

$$\begin{aligned} EE^T &= |det(A)|^{-1}\Delta_g^{-1} M_g^\mu \Delta_g^{-T} \\ &= |det(A)|^{-1}\Delta_g^{-1}\Delta_g \Delta_g^T \Delta_g^{-T} \\ &= |det(A)|^{-1}I. \end{aligned} \tag{21}$$

where I is an identity matrix. The above equation indicates that E is a similarity matrix which encodes the relation of the matrices A, Δ_f and Δ_g. According to Eq. (21), the transform matrix E can be denoted by $E = \alpha R_{\theta_0}$ for $det(E) > 0$

and $E = \alpha \tilde{R}_{\theta_0}$ for $det(E) < 0$, where $\alpha = |det(A)|^{-1/2}$ play the role of scale factor, R_{θ_0} and $\tilde{R}_{\theta_0}$ are

$$R_{\theta_0} = \begin{bmatrix} cos\theta_0 & -sin\theta_0 \\ sin\theta_0 & cos\theta_0 \end{bmatrix} \tag{22}$$

and

$$\tilde{R}_{\theta_0} = \begin{bmatrix} cos\theta_0 & sin\theta_0 \\ sin\theta_0 & -cos\theta_0 \end{bmatrix} \tag{23}$$

which can transform a vector $\overrightarrow{\theta}$ to be $\overrightarrow{\theta + \theta_0}$ and $\overrightarrow{\theta_0 - \theta}$ respectively. It is worth noting that for Eq. (19), since $det(\Delta_f) > 0$ and $det(\Delta_g) > 0$ [5][6], $det(E)$ takes the same sign as $det(A)$. While $det(A) < 0$ indicates that besides the scaling, rotation and shearing, the affine transform with the affine matrix A also includes a mirror transform.

We have now obtained the matrices Δ_f and Δ_g, they are then be used to normalize the function $\Psi_f^\mu(\theta)$ and its affine-transform version $\Psi_g^\mu(\theta)$ respectively. We first normalize the function $\Psi_f^\mu(\theta)$ as

$$\dot{\Psi}_f^\mu(\theta) = \frac{1}{\left\| \Delta_f \overrightarrow{\theta} \right\|} \Psi_f^\mu \left(\left\langle \Delta_f \overrightarrow{\theta} \right\rangle \right). \tag{24}$$

Similarly, the normalized version of $\Psi_g^\mu(\theta)$ is defined as $\dot{\Psi}_g^\mu(\theta)$. The relationship between them can be achieved as

$$\dot{\Psi}_g^\mu(\theta) = \alpha \cdot \dot{\Psi}_f^\mu(\theta - \theta_0) \quad for \quad det(A) > 0, \tag{25}$$

and

$$\dot{\Psi}_g^\mu(\theta) = \alpha \cdot \dot{\Psi}_f^\mu(\theta_0 - \theta) \quad for \quad det(A) < 0. \tag{26}$$

which indicate that the normalized versions $\dot{\Psi}_f^\mu(\theta)$ and $\dot{\Psi}_g^\mu(\theta)$ are only correlated by a similar transform. The Eq. (26) also indicates that when the affine transform include a mirror transform, the original function $\dot{\Psi}_f^\mu(\theta)$ is also subject to an additional mirror transform.

Also, we can see that for any $\mu \in [0, 1]$, its corresponding 1D function $\dot{\Psi}_f^\mu(\theta)$ for the shape f only suffers from scaling, translation and mirror distortions when the shape f is subject to affine transform. In the former section, we fix the variable μ for the convenience of presenting the details of the proposed transform. Now, we set it free and rewrite the 1D function $\dot{\Psi}_f^\mu(\theta)$ to a 2D function $\dot{\Psi}_f(\mu, \theta)$. Obviously, the affine transform makes the 2D function $\dot{\Psi}_f(\mu, \theta)$ occur only shifting and mirror effects on the dimension of θ and a scaling of the amplitude of the function.

4.3   Affine Invariants

Here we apply the proposed transform $\dot{\Psi}_f(\mu, \theta)$ to affine invariant shape recognition. As discussed in the former section, the transform $\dot{\Psi}_f(\mu, \theta)$ is only subject to a shifting and a mirror transform on the dimension of θ and a scaling on

the amplitude of the function. So, it is very easy to remove these effects from the proposed transform. To do so, we apply the 1D Fourier transform to the dimension of θ against the proposed transform $\dot{\Psi}_f(\mu, \theta)$. Assume that $\dot{\Psi}_f(\mu_i, \theta_j)$, $i = 1, ..., k$ and $j = 1, ..., N$ is the digital form of the proposed transform $\dot{\Psi}_f(\mu, \theta)$, where K and N are the number of the points uniformly sampled from the range $[0, 1]$ and the range $[0, 2\pi]$ respectively. Then we obtain a matrix of size $K \times N$. For each row of the matrix, we perform discrete 1D Fourier transform against it and keep the magnitudes of Fourier coefficients. According to the theory of Fourier transform, the shifting and mirror effects are removed from the transform $\dot{\Psi}_f(\mu_i, \theta_j)$. As for the scaling effect, we normalize each row of the matrix using its respective 0th-order Fourier coefficient. Then we obtain a completely affine invariant version of the transform $\dot{\Psi}_f(\mu_i, \theta_j)$ which can be directly utilized for shape recognition. The dissimilarity between two shapes can be measured by calculating the L1 norm between their transforms $\dot{\Psi}(\mu_i, \theta_j)$.

5 Experimental Results and Discussions

To examine the feasibility and effectiveness of the proposed method on shape retrieval, we perform the proposed method on two groups of shape image datasets: (1) Contour shape dataset in which each shape is enclosed by a single silhouette and no content is contained inside, and (2) Region shape dataset in which each sample has several separated regions or its whole region is though enclosed by a single silhouette, it also contain some contents inside which usually have complex structure. For all the experiments, we uniformly sample 18 values from the range $[0,1]$ and 180 values from the range $[0, 2\pi]$ for conducting the proposed transform.

To quantify the retrieval performance of the algorithms, the standard metric for information retrieval, knee-point score [7][8], is used in our experiments. For any query shape Q_i, calculate the distances of all the dataset samples to it and rank them to a sequence in ascending order. Let H be the number of all the dataset samples and V_i be the number of the relevant ones to the query shape Q_i in the dataset. For each integer number $1 \leq h \leq H$, count the number $v_{i,h}$ of the relevant samples of the top k best matches in the sequence. Then for the given query Q_i, its precision and recall at the top h best matches are defined as

$$p_{i,h} = \frac{v_{i,h}}{h} \quad and \quad r_{i,h} = \frac{v_{i,h}}{V_i}, \tag{27}$$

respectively. We calculate their average values over all the queries Q_i. In our experiments, each sample from the dataset is taken in turn as a query. So, there are a total of H queries. When $V_i = h$, precision and recall will take the same value which is termed knee-point score as measurement [7][8].

Multiview Curve Database (MCD): To evaluate the performance of the algorithms on recognizing the curved shapes in presence of affine distortion, Zuliani et al. [10] chose 40 samples from the MPEG-7 CE-1 contour shape dataset

[9] with each selected from one shape class. Each of them is then printed on a white paper and 7 pictures are taken to it from different view angles using a digital camera. Another 7 images are achieved by randomly rotating and reflecting the available seven samples. So, there are a total of $40 \times 14 = 560$ samples in the MCD which consists of MPEG-7 shapes that are affected by natural perspective skew due to the manner of extracting from real images. Some typical samples from the MCD are shown in Fig. 3. This dataset is publicly available and has been used as test case in many works [7][10][14].
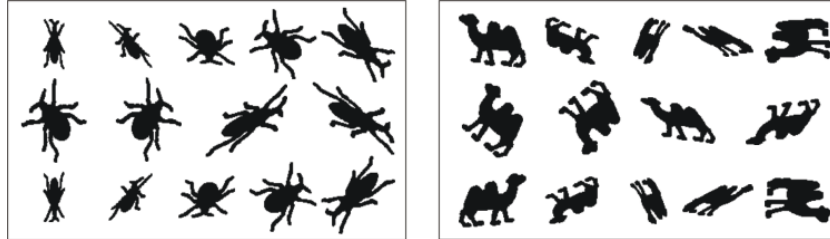
Fig. 3. Part of typical samples in the dataset MCD. Left: the 14 affine-distorted insect shapes, right: the 14 affine-distorted camel shapes.

To make a fair comparison, we choose those approaches which are particularly designed for affine shape recognition. Three region based methods, including Trace transform based method [5], Multiscale autoconvolution (MSA) [11] and Affine moment invariants [12] which are state-of-the-arts descriptors for affine shape recognition, are used as benchmarks in our experiments. Since the template sample of the shapes in MCD database are from the MPEG-7 CE-1 dataset which is a contour based test case, four recently published contour based methods including Affine-invariant elastic metric [7], Hierarchical projective invariant contexts [14], subspace approach [13] and Affine-invariant curve matching [10] are selected as benchmarks for a wide comparison. All of them take the MCD dataset as the test case in their experiments. The Knee-point scores of the proposed method together with the benchmark methods are summarized in Table 1. It can be seen that the proposed method achieves 97.07% of retrieval accuracy which is about 17% higher than the Trace transform based method [5] and much more than the other two region based methods [12] and [11]. While compare with the other four contour based methods which generally perform better on the contour based test case than those region based methods, the proposed method still achieves more than about 7% of retrieval accuracy than them.

MPEG-7 CE-2 Perspective Transform Test: MPEG-7 CE-2 database is developed for evaluating the performance of those region based shape analysis methods. Here, we perform the MPEG-7 CE-2 perspective test to validate the effectiveness of the proposed on the retrieval of region shapes in presence of perspective transform. In this test protocol, all the 3101 region-based shape

Table 1. The retrieval accuracy on the MCD dataset

Algorithm	Knee-point score (%)
Affine-invariant elastic metric [7]	90.00*
Hierarchical projective invariant contexts [14]	88.14*
Subspace Approach [13]	90.15
Affine-invariant curve matching [10]	80.00*
Affine moment invariants [12]	66.01
Mutiscale Autoconvolution (MSA) [11]	71.84
Affine invariant features from the Trace transform [5]	80.08
Proposed method	**97.07**

* *The results from the original papers.*
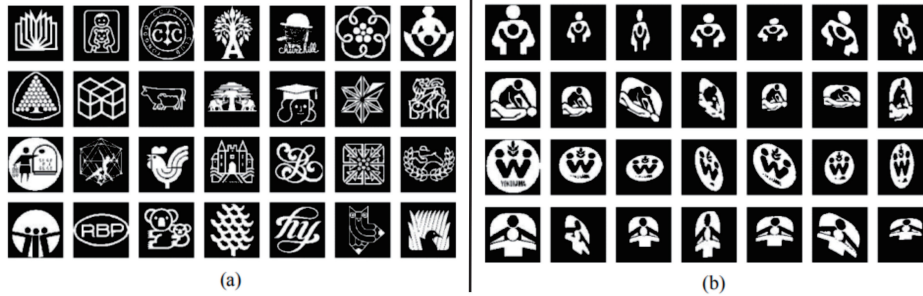
Fig. 4. Example samples from the MPEG CE-2 database which consists of 3101 region shape images. (a) Some typical samples that are used as gallery images. (b) Some typical images that are used as queries, where for each row, the top left one is template shape and the remaining ones are part of its various perspective transformed versions.

images in the database are used as gallery images (Some typical samples are shown in Fig. 4(a)). 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 (Example images are shown in Fig. 4(b)). As can be seen that different from the CE-1 shapes, the CE-2 samples usually have complex interior structures and some samples have even separate shape regions.

In our experiments, we follow the protocol of the CE-2 perspective transform test. Since the contour based shape recognition methods used in the former experiments can not handle region shapes, we only compare the proposed method with the other three region based methods. The knee-point scores for all the compared methods are summarized in Table 2. It can be seen that the proposed method achieved an accuracy of 97.49%, much higher than those of all the benchmarks (2.48% higher than the second best approach), on retrieving shapes with various perspective transformations. As can be seen that on the CE-2 perspective test, the proposed method keeps the best retrieval performance

over the benchmark methods. The encouraging experimental results demonstrate the effectiveness of the proposed transform in describing shapes in presence of perspective transform and its superior discriminability over the existing region based methods on handling the shapes with complex interior structures.

Table 2. The retrieval accuracy for the MPEG-7 CE-2 perspective test

Algorithm	Knee-point score (%)
Affine moment invariants [12]	42.53
Mutiscale Autoconvolution (MSA) [11]	62.09
Affine invariant features from the Trace transform [5]	95.01
Proposed method	**97.49**

6   Conclusions

A novel line integral transform has been presented for affine-invariant shape recognition. It is a 2D function which is not only invariant to affine transform, but also preserves the spatial relationship between the line integrals which makes it more discriminative than those shape descriptors from the trace transform. In additional, the proposed method is more efficient than the trace transform. In the proposed method, a 2D matrix of affine invariants can be generated by conducting the proposed transform once. While conducting the trace transform once can only generate a single feature and multiple times of transforms should be performed make shape descriptors discriminative. The proposed transform has been tested on the standard affinely distorted contour shape database and region shape database and compared with the state-of-the-art shape descriptors that are designed for affine shape analysis. The encouraging experimental results showed that the proposed method is effective for affine shape recognition.

References

1. Deans, S.R.: The Radon Transform and Some of Its Applications. New York: Wiley (1983)
2. Strichartz, R.: A Guide to Distribution Theory and Fourier Transforms. CRC Press.(1994)
3. Hong, B., Soatto, S.: Shape Matching Using Multiscale Integral Invariants. IEEE Transactions on Pattern Analysis and Machine Intelligence **37**(1), 151-160 (2015)

4. Baumberg, A.: Reliable Feature Matching across Widely Separated Views. Proc. IEEE Conf. Computer Vision and Pattern Recognition, vol. 1, pp. 774-781 (2000)
5. Petrou, M., Kadyrov, A.: Affine Invariant Features from the Trace Transform. IEEE Transactions on Pattern Analysis and Machine Intelligence **26**(1), 30-44 (2004)
6. Golub, G.H., Van Loan, C.F.: Matrix computations, 3rd edn. Johns Hopkins University Press, Baltimore(1996)
7. Bryner, D., Srivastava, A., Klassen, E.: Affine-invariant, elastic shape analysis of planar contours. in Proc. IEEE Conf. Computer Vision Pattern Recognition, pp. 390-397 (2012)
8. Bryner, D., Klassen, E., Le, H., Srivastava, A.: 2D affine and projective shape analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence **36**(5), 998-1011 (2014)
9. Latecki, L.J., Lakamper, R., Eckhardt, T.: Shape descriptors for no-rigid shapes with a single closed contour. in Proc. IEEE Conf. Computer Vision Pattern Recognition, vol. 1, pp. 424-429 (2000)
10. Zuliani, M., Bhagavathy, S., Manjunath, B.S., Kenney, C.: Affine-invariant curve matching. in: Proc. IEEE Int. Conf. Image Processing (2004)
11. Rahtu, E., Salo, M., Heikkila, J.: Affine invariant pattern recognition using multi-scale autoconvolution. IEEE Transactions on Pattern Analysis and Machine Intelligence **27**(6), 908-918 (2005)
12. Flusser, J., Suk, T.: Pattern recognition by affine moment invariants. Pattern Recognition **26**(1), 167-174 (1993)
13. Mai, F., Chang, C., Hung, Y.S.: A subspace approach for matching 2D shapes under affine distortions. Pattern Recognition **44**(2), 210-221 (2011)
14. Jia, Q., Fan, X., Liu, Y., Li, H., Luo, Z., Guo, H.: Hierarchical projective invariant contexts for shape recognition. Pattern Recognition **52**(52), 358-374 (2016)
15. Domokos, C., Kato, Z.: Parametric estimation of affine deformations of planar shapes. Pattern Recognition **43**(3), 569-578 (2010)
16. Tarel, J., Wolovich, W., Cooper, D.B.: Covariant-conics decomposition of quartics for 2d shape recognition and alignment. Journal of Mathematical Imaging and Vision **19**(3), 255-273 (2003)
17. Yap, P., Jiang, X., Kot, A.C.: Two-Dimensional Polar Harmonic Transforms for Invariant Image Representation. IEEE Transactions on Pattern Analysis and Machine Intelligence **32**(7), 1259-1270 (2010)
18. Wang, B., Gao, Y.: Structure integral transform versus Radon transform: A 2D mathematical tool for invariant shape recognition. IEEE Transcations on Image Processing **25**(12), 5635-5648 (2016)
19. Huang, Z., Cohen, F.S.: Affine-invariant B-spline moments for curve matching. IEEE Transcations on Image Processing **5**(10), 1473-1480 (1996)
20. Wang, Y., Teoh, E.K.: 2D affine-invariant contour matching using B-spline model. IEEE Transcations Pattern Analysis and Machine Intelligence **29**(10), 1853-1858 (2007)
21. Srestasathiern, P., Yilmaz, A.: Planar shape representation and matching under projective transformation. Computer Vision and Image Understanding **115**(11), 1525-1535 (2011)
22. Arbter, K., Snyder, W.E., Burkhardt, H., Hirzinger, G.: Application of affine-invariant Fourier descriptors to recognition of 3-D objects.IEEE Trans. Pattern Analysis and Machine Intelligence **12**(7), 640-647 (1990)
23. Heikkil, J.: Pattern matching with affine moment descriptors. Pattern Recognition **37**(9), 1825-1834 (2004)

24. Khalil, M.I., Bayoumi, M.M.: A dyadic wavelet affine invariant function for 2D shape recognition. IEEE Trans. Pattern Analysis and Machine Intelligence **23**(10), 1152-1164 (2001)
25. Rube, I.E., Ahmed, M., Kamel, M.S.: Wavelet approximation-based affine invariant shape representation functions. IEEE Trans. Pattern Analysis and Machine Intelligence **28**(2), 323-327 (2006)
26. Ruiz, A., Lopez-de-Teruel, P.E., Fernandez-Maimo, L.: Efficient planar affine canonicalization. Pattern Recognition, vol. 72, pp. 236-253 (2017)
27. Xue, Z., Shen D., Teoh, E.K.: An efficient fuzzy algorithm for aligning shapes under affine transformations. Pattern Recognition **34**(6), 1171-1180 (2001)
28. Yang, Z., Cohen, F.S.: Cross-weighted moments and affine invariants for image registration and matching. IEEE Trans. Pattern Analysis and Machine Intelligence **21**(8), 804-814 (1999)
29. Kadyrov, A., Petrou, M.: Affine parameter estimation from the trace transform. IEEE Trans. Pattern Analysis and Machine Intelligence **28**(10), 1631-1645 (2006)
30. Wang, Y., Huang, K., Tan, T.: Human Activity Recognition Based on R Transform. In: CVPR, pp. 1-8 (2007)
31. Andriluka, M., Roth, S., Schiele, B.: People-tracking-by-detection and people-detection-by-tracking. in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR) (2008)
32. Korn, P., Sidiropoulos, N.D., Faloutsos, C., Siegel, E.L., Protopapas, Z.: Fast and Effective Retrieval of Medical Tumor Shapes. IEEE Transactions on Knowledge and Data Engineering **10**(6), 889-904 (1998)
33. Suk, T., Flusser, J.: Affine moment invariants generated by graph method. Pattern Recognition **44**(9), 2047-2056 (2011)
34. Chen, Y.W., Chen, Y.Q.: Invariant Description and Retrieval of Planar Shapes Using Radon Composite Features. IEEE Transactions on Signal Processing **56**(10), 4762-4771 (2008)
35. Hjouj, F., Kammler, D.W.: Identification of Reflected, Scaled, Translated, and Rotated Objects from Their Radon Projections. IEEE Transactions on Image Processing **17**(3), 301-310 (2008)
36. Tabbone, S., Wendling, L., Salmon, J.: A New Shape Descriptor Defined on the Radon Transform. Computer Vision and Image Understanding **102**(1), 42-51 (2006)
37. Hoang, T.V., Tabbone, S.: The Generalization of the R-transform for Invariant Pattern Representation. Pattern Recognition **45**(6), 2145-2163 (2012)
38. Hasegawa, M., Tabbone, S.: Amplitude-only log Radon transform for geometric invariant shape descriptor. Pattern Recognition **47**(2), 643-658 (2014)
39. Kadyrov, A., Petrou, M.: The trace transform and its applications. IEEE Transactions on Pattern Analysis and Machine Intelligence **23**(8), 811-828 (2001)