

## ORIGINAL RESEARCH

# A matrix projection approach for matching shapes under affine distortions

Zhaoyang Zhang<sup>\*1</sup>, Zheng Tian<sup>2</sup>, Weiwei Cui<sup>2</sup>

<sup>1</sup>*School of Information Engineering, Chang'an University, Xi'an, China*

<sup>2</sup>*School of Science, Northwestern Polytechnical University, Xi'an, China*

**Received:** November 16, 2014

**Accepted:** December 17, 2014

**Online Published:** January 20, 2015

**DOI:** 10.5430/air.v4n1p45

**URL:** <http://dx.doi.org/10.5430/air.v4n1p45>

## Abstract

This paper presents a matrix projection method for matching a pair of shapes, and estimating the affine transformation that aligns the two shapes. First, shapes are considered as 2D disordered point sets, and their normalized forms are given. Then, we prove that only a rotational transform exists between the normalized forms of the two shapes under affine distortions. Second, correspondences are found by minimizing the inner product between one matrix and projection of the other. Finally, the affine transformation for shape registration is estimated by the correspondences. Experimental results show that our approach compares favorably to other methods under affine distortions.

**Key Words:** Shape matching, Affine distortions, Matrix projection, Shape registration

## 1 Introduction

Shape matching, to find a transformation aligning two shapes, is a crucial problem in computer vision, where images of different views of an object need to be recognized, compared, and registered. The alignment has extensive use in recognition, indexing and retrieval, and tracking. However, the images of an object taken from different viewpoints suffer from perspective distortions, which make shape matching a difficult task. In general, a good shape matching algorithm needs to be invariant to rotation, translation, scaling and shearing. Under the assumption of weak perspective, which means two images are taken from a distance that is an order of magnitude or more greater than the maximum object diameter along the direction of the view, two planar views of the same object are related through an affine transformation.<sup>[1,2]</sup> This explains why there are a large number of methodologies developed to

compare shapes under affine distortions. Shape registration algorithms can be classified into two main categories: contour-based methods and region-based methods.

Contour-based methods try to establish correspondences between points on contours of two shapes. A large number of contour-based methods are based on affine-invariant Fourier descriptor (FD),<sup>[3,4]</sup> due to the simplicity and computational efficiency of the Fast Fourier Transform (FFT). These methods use a set of Fourier coefficients to represent contours to compute affine invariants. However, a descriptor with a small number of coefficients may not be sufficient to represent the shape, while one with a large number of coefficients may be sensitive to noise given the involvement of high frequency Fourier coefficients; which makes the selection a suitable set of coefficients a crucial issue. Another affine invariant descriptors for contour-based shape matching is wavelet descriptors (WD),<sup>[5-7]</sup> which analyzes shapes

<sup>\*</sup>**Correspondence:** Zhaoyang Zhang; Email: zhzi0218@yahoo.com.cn; Address: School of Information Engineering, Chang'an University, Xi'an 710064, China

in the spectral domain as FD. Most wavelet descriptors are highly sensitive to noise and changes to the wavelet basis because they are constructed by the detail-based representation of the shapes. Therefore, an affine invariant shape representation is proposed by applying two different wavelet transforms with two different wavelet basis functions based on approximation principles.

Region-based methods treat the matching problem without detecting contours of objects. They consider both the boundary of a shape and the pixels within a shape region to obtain the shape information. In this category, moments are commonly used. An affine invariant descriptor<sup>[8,9]</sup> based on the second and higher order moments of image points was proposed, it is affine invariant but sensitive to observation noise. Recently, a novel moment based method<sup>[10]</sup> was proposed. For this method, a polynomial system of equations is generated by image moments, and its solution is obtained as the transformation parameter for registration. Noise in the image coordinates may make the alignment precision less satisfactory in some cases, or even make it hard for the method to find a solution for the polynomial system. Region-based methods need to store and handle much more data than contour-based ones, which adds a burden on storage and computation time.

Unlike previously mentioned approaches, three state-of-art methods are proposed. Independent component analysis (ICA) is used to construct affine-invariant descriptors,<sup>[11-13]</sup> where an observed curve is represented by two variables. If a pair of observed curves is related by an affine transformation, their independent components are related only by a reflection. However, this kind of method is restricted by independency assumption. The other two methods are the “shape context” method<sup>[14]</sup> and the subspace-based method,<sup>[15]</sup> which treat curves as 2D signals. The “shape context” method is used as a vector-valued attribute in a bipartite graph matching framework, and can be regarded as a graph-based method. The solution to graph matching can provide the correspondences for registration. The method is robust to deformation and noise, but it is time consuming due to an iterative minimization process for establishing point correspondences. In the subspace-based method,<sup>[15]</sup> one shape is projected onto the subspace spanned by the other, and the affine transformation is estimated by minimizing the projection error in this subspace. It is easy to implement, but has problems dealing with occlusions.

In this paper a novel matrix projection method for affine shape registration is proposed, which aims to handle some of the existing problems in shape matching. Given a pair of curves to be aligned, the idea is to use the matrix projection space to find correspondences between their canonical forms with different size, and the affine transformation is estimated by minimizing the inner product between one matrix and a projection of the other. The proposed approach is easy to implement, and can be clearly defined in terms of

mathematical and physical interpretation. Furthermore, it is more robust to noise compared to other state-of-art methods. The organization of the paper is as follows. The problem of affine invariant shape registration is formulated in Section 2 and details of the proposed method are presented in Section 3. Experimental results are described in Section 4 and concluding remarks are given in Section 5.

## 2 Problem formulation

We are given a pair of 2D shapes, which are images of the same object, the template and the target. Assuming that perspective transformation between a pair of shapes can be approximated by an affine transformation, the goal of this work is to estimate the affine transformation aligning them. Each shape in the image is originally represented by a set of unlabelled edge points, which are image pixels belonging to its contour. The center of the curve is set as the origin of the coordinate system describing the position of the points.

Let the two given 2D shapes, template and target, be represented by  $X = (x_1, x_2, \dots, x_{n_x})$  and  $Y = (y_1, y_2, \dots, y_{n_y})$ , where  $x_i = (x_{1i}, x_{2i})^T$  and  $y_i = (y_{1i}, y_{2i})^T$  are points belonging to the contours of the shapes. Normally, estimating the affine transformation aligning the two shapes through the two given matrices leads to two problems: (1) the number of points  $n_x = n_y = n$  or  $n_x \neq n_y$ ; (2) the  $i$ -th point in  $X$  does not correspond to the  $i$ -th point in  $Y$  or for points in  $X$  there do not exist corresponding points in  $Y$ . In the following sections, we will describe the two problems in detail and show relevant solutions for each.

In the ideal case when there is no measurement noise, two shapes are related exactly by an affine transformation,  $n_x = n_y = n$ , and the  $i$ -th point in  $X$  corresponds to the  $i$ -th point in  $Y$ , then the following equation is satisfied:

$$Y = AX \quad (1)$$

where  $A$  is a nonsingular matrix representing the affine transformation. However, the equality in (1) does not hold exactly in real applications due to contour detection error, transformations which are not exactly affine, and non-uniqueness of the sequence of points. Therefore, for a given pair of shapes represented by  $X$  and  $Y$ , the aim is to find a permutation matrix  $P$  and estimate the affine transformation  $A$  such that the registration error is minimized as follows:

$$\min_{A, P} \|Y - AXP\|_F \quad (2)$$

such that template  $X$  lies on the target  $Y$ , where  $\|\bullet\|_F$  denotes the Frobenius norm.

For the minimization problem in (2), the objective function should be minimized over  $A$  and  $P$  under the constraint that template  $X$  lies on the target  $Y$ . However, in a real applica-

tion,  $X$  and  $Y$  are not precisely specified but only consists of unlabelled points as features extracted from shape images, and usually have different sizes. In the following sections, we will show how to make the minimization problem (2) equivalent to two optimization problems with single optimized parameters, and find out the correspondence between  $X$  and  $Y$  by  $P$ , and estimate the affine transformation  $A$ .

### 3 Matrix projection approach to affine invariant shape registration

The proposed method consists of two main steps. First, contours are normalized by singular value decomposition (SVD). Second, a matrix projection method is applied to solve the minimization problem (2). In the following phases, the proposed method will be described in detail.

#### 3.1 Normalization of shape contour

Generally, an affine transformation consists of shearing, scaling, rotation and translation, in which shearing and scaling are the most influential negative components for an application. This paper will weaken or remove the shearing and scaling effects by normalizing the curve into its canonical form, which is the whitened version of the curve.<sup>[8]</sup> A curve is in a canonical form if the covariance matrix of the set of points that belong to the curve is an identity matrix. If the covariance matrix has a factorization  $C = FF^T$ , the whitening transform is expressed by  $T = F^{-1}S$ , where  $\bar{S} = S - E\{S\}$ ,  $T$  is in the canonical form with the identity matrix as its covariance matrix.

Mathematically, for a given  $X$ , normalization can be achieved by SVD as follows:<sup>[16]</sup>

$$X = U_x \sum_x V_x \quad (3)$$

Where  $X = (x_1, x_2, \dots, x_n)$ ,  $U_x$  is a  $2 \times 2$  unitary matrix,  $\sum_x$  is a  $2 \times 2$  diagonal matrix, and  $V_x$  is a  $2 \times n$  matrix. From (3) we have  $V_x = \sum_x^{-1} U_x^T X$ , which is the canonical form of curve  $X$  satisfying  $V_x V_x^T = I$ , where  $I$  is the identity matrix, and  $V_x$  is the whitened version of the curve  $X$ .

The relationship between canonical forms of shapes related by affine transformation has a succinct formulation.<sup>[15,17]</sup> However, the drawback is that the order of points is not involved. In this paper, we improved the theories in those methods through a permutation matrix  $P$ . Based on the formulation in (3), points on curves  $X$  and  $Y$  can be projected into new coordinate spaces which are eigen-spaces of  $X$  and  $Y$ , and for shapes related by affine transformation we only need a rotation for matching in the new coordinate space with their canonical forms. This means that the shearing and scaling effects between template shape and target shape can be removed during the normalizing process, based on

the following theorem.

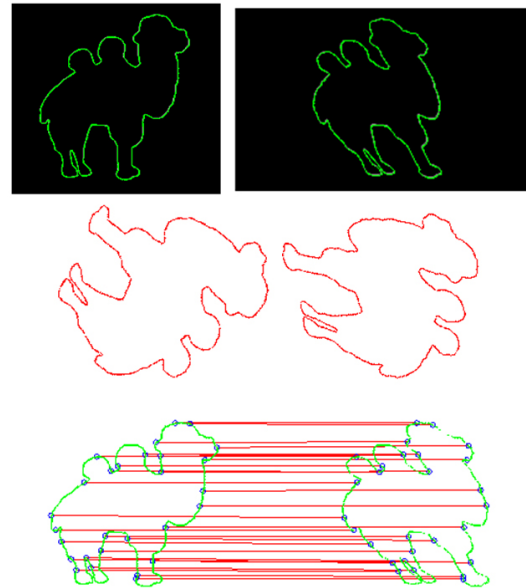
**Theorem 1:** Given point sets  $X = (x_1, x_2, \dots, x_n)$  and  $Y = (y_1, y_2, \dots, y_n)$ , there exists a rotation matrix  $R$  such that  $V_y = RV_x P$  if there exists an invertible matrix  $A$  and a permutation matrix  $P$  such that  $Y = AXP$ , where  $V_x$  and  $V_y$  are the canonical forms of  $X$  and  $Y$ .

**Proof:** According to Eq. (3),  $X$  and  $Y$  can be normalized into their canonical forms as  $X = U_x \sum_x V_x$  and  $Y = U_y \sum_y V_y$ . Using the equation  $Y = AXP$ , we get the following terms:

$$\begin{aligned} U_y \sum_y V_y &= AU_x \sum_x V_x P \Rightarrow V_y \\ &= \sum_y^{-1} U_y^T AU_x \sum_x V_x P \Rightarrow V_y = RV_x P \end{aligned}$$

where  $R = \sum_y^{-1} U_y^T AU_x \sum_x$  which means  $R$  is the angle between two basis of the same space spanned by  $U_x$  or  $U_y$ , and  $RR^T = V_y P^T V_x^T V_x P V_y^T = V_y V_y^T = I$ ,  $RR^T = V_y P^T V_x^T V_x P V_y^T = V_y V_y^T = I$ .

Theorem 1 shows that if two disordered unlabeled point sets can be matched under an affine transformation, their canonical forms are related only by rotation. Figure 1 shows an example illustrating Theorem 1, where (left) and (right) are a pair of camel images and (right) is obtained from an affine transform applied on (left). The detected unlabeled points on contours are represented by  $X(n_x = 1360)$  and  $Y(n_y = 1390)$ , which are marked on the original images. The canonical form of  $X$  and  $Y$  are shown in the middle, which are only related by a rotational transform.



**Figure 1:** Affine related camel images. Top: the original shape images marked by the unlabeled edge points; middle: the canonical forms of those unlabeled points; bottom: the aligned sample points of curves for the camel, 28 of 1360 pairs of matched sample points are shown with circles.

### Minimization by matrix projection

In this section, the minimization problem<sup>[2]</sup> will be converted to two optimization problems with a single optimization parameter, and permutation matrix  $P$  will be found by canonical forms of  $X$  and  $Y$ , which are only related by rotation, and an affine transformation  $A$  will be estimated by correspondence.

For a given curve  $X$ , SVD is applied to  $X$  to obtain its canonical form  $V_x$ :

$$V_x = \sum_x^{-1} U_x^T X$$

Let  $V_x = (v_1, v_2, \dots, v_n)$ , where  $v_i = (v_{1i}, v_{2i})^T$  is the  $i$  th point of  $X$  in the new coordinate system. We define a  $n \times n$  similarity matrix  $W$  with entry  $W_{i,j} = d(v_{\bullet i}, v_{\bullet j})$  being the distance between the  $i$ -th point and the  $j$ -th point of  $V_x$ ,  $i, j = 1, 2, \dots, n$ . We use the Gaussian distance metric  $d(v_{\bullet i}, v_{\bullet j}) = \exp(-\frac{\|v_{\bullet i} - v_{\bullet j}\|^2}{2\sigma^2})$ . Based on Theorem 1, the minimization problem (2) can be rewritten as:

$$\begin{cases} P^* = \min_{P, P^T P = I} \|W_y - P^T W_x P\|_F \\ A^* = \min_A \|Y - AX P^*\|_F \end{cases} \quad (4)$$

using the following theorem.

**Theorem 2:** Let  $W_x$  and  $W_y$  be similarity matrices of  $V_x$  and  $V_y$ , respectively, then:  $W_y = P^T W_x P$  if there exists a rotation matrix and a permutation matrix  $P$ , such that  $V_y = R V_x P$ , where  $V_x$  and  $V_y$  are canonical forms of  $X$  and  $Y$ .

**Proof:** Define  $\langle v_{\bullet i}, v_{\bullet j} \rangle = d(v_{\bullet i}, v_{\bullet j})$ , then  $W_y = V_y^T V_y = P^T V_x^T R^T R V_x P = P^T V_x^T V_x P = P^T W_x P$ .

Furthermore, according to the following equations:

$$\|W_y - P^T W_x P\|_F = \langle W_y - P^T W_x P, W_y - P^T W_x P \rangle = \langle W_y, W_y \rangle - 2 \langle W_y, P^T W_x P \rangle + \langle W_x, W_x \rangle$$

the minimization problem given in (4) can be rewritten as:

$$\begin{cases} P^* = \arg_P \max_{P^T P = I} \langle W_y, P^T W_x P \rangle \\ A^* = \min_A \|Y - AX P^*\|_F \end{cases} \quad (5)$$

This problem can be solved exactly, and the optimal value is obtained by performing spectral decompositions of  $W_x$  and  $W_y$ . Let us suppose that the orthogonal diagonalizations of  $W_x$  and  $W_y$  are  $W_x = U_x D_x U_x^T$  and  $W_y = U_y D_y U_y^T$  respectively, where the eigenvalues in  $D_x$  and  $D_y$  are ordered in a non-decreasing fashion to guarantee uniqueness. The optimal value of  $P^* = \arg_P \max_{P^T P = I} \langle W_y, P^T W_x P \rangle$  is  $\text{tr}(D_x D_y)$ , and the optimal solution is obtained by using the orthogonal matrices that yield the diagonalizations, i.e.  $P^* = U_x U_y^T$ . The affine transformation minimization  $A^* = \min_A \|Y - AX P^*\|_F$  can be obtained

by the standard least square technique as:

$$\begin{aligned} A^* &= Y(XP^*)^T((XP^*)(XP^*)^T)^{-1} \\ &= Y(XP^*)^T(XX^T)^{-1} \end{aligned} \quad (6)$$

The previous analysis in this subsection is under the condition  $n_x = n_y = n$  for curves  $X$  and  $Y$ . The following subsection will describe the problem (2) under  $n_x \neq n_y$  and also when the  $i$ -th point in  $X$  does not correspond to the  $i$ -th point in  $Y$ , or when for points in  $X$  there do not exist corresponding points in  $Y$ .

### 3.2 Minimization by matrix projection for the general situation

Even if two shapes are related by an affine transformation, it is common that there are different number of pixels along their contours, i.e.,  $n_x \neq n_y$  for  $X, Y$  and  $V_x, V_y$  in the previous subsection. Without loss of generality, let  $n_x > n_y$ , in the previous subsection. Without loss of generality, let  $A^*$  can be obtained by the standard least squares technique according to the first  $n_y$  correspondences between  $V_y$  and  $V_x P$  if there exists a  $n_x \times n_x$  matrix  $P$  such that the following formula:

$$\begin{bmatrix} W_y & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} E_{n_y} & 0 \\ 0 & 0 \end{bmatrix} P^T W_x P \begin{bmatrix} E_{n_y} & 0 \\ 0 & 0 \end{bmatrix}$$

is satisfied, where  $E_{n_y}$  is a  $n_y \times n_y$  identity matrix,  $P$  is a  $n_x \times n_x$  permutation matrix,  $W_x$  and  $W_y$  are  $n_x \times n_x$  and  $n_y \times n_y$  matrices respectively, which are similarity matrices for  $V_x$  and  $V_y$ .

If we denote  $\begin{bmatrix} W_y & 0 \\ 0 & 0 \end{bmatrix}$  as  $W_y$ , the affine transformation  $A^*$  to make template  $X$  lie on the target  $Y$ , can be obtained by solving the following maximization and minimization problems:

$$\begin{cases} P^* = \arg_P \max_{P^T P = I} \langle W_y, P^T W_x P \rangle \\ A^* = \min_A \|Y - AX [E_{n_y}, O_{n_y \times (n_x - n_y)}] P^*\|_F \end{cases} \quad (7)$$

where  $O_{n_y \times (n_x - n_y)}$  is a  $n_y \times (n_x - n_y)$  zero matrix. Figure 1, bottom, shows an example of the alignment result for the camel shapes given in Figure 1. The contours  $X$  and  $Y$  have 1360 and 1390 points respectively, and 1360 correspondences are obtained by solving the optimization problem (7). In this figure, 28 pairs of matched points are plotted with an interval of 50 points for each set according to the recorded sequence. The corresponding sample points are shown with circles.

## 4 Experimental results

To demonstrate the performance of the proposed approach, we compare our algorithm with the ‘‘shape context’’ method<sup>[14]</sup> and the ‘‘area based’’ method<sup>[10]</sup> which are contour based and region based respectively, for image

shape registration. The three methods are all implemented in MATLAB, and are tested on both synthetic and real images considering the following items: registration accuracy for synthetic images, robustness against noise, and registration accuracy on real images. The parameters for “shape context” method are consistent with those in Ref.<sup>[14]</sup> to make the algorithm perform the best, where the regularization parameter is 500, the iteration number is 40 and other parameters remain as their default values. The results show that the proposed method compares favorably with the other methods, both by visual inspection and quantitative measurements.

#### 4.1 Measurements for registration error

We use the error measurement defined in Ref.<sup>[10,15]</sup> For two given shapes,  $X^*$  for template and  $Y^*$  for target; Equation (1) shows the affine relationship between them. The registered shape  $Y_R^*$  is the shape contour transformed from the template  $X^*$  by the estimated affine transformation  $A^*$ , which is the solution of problem (4) or (7), as follows:

$$Y_R^* = A^* X^* \quad (8)$$

Therefore, the error measurement can be defined as the difference between the target shape and the registered shape in terms of the percentage of non-overlapping area between  $Y^*$  and  $Y_R^*$ , which can be expressed as:

$$\delta = \frac{|T \oplus S|}{|T| + |S|} \quad (9)$$

where  $T$  denotes the set of pixels belonging to the target shape, including the pixels on and inside the shape contour  $Y^*$ ,  $S$  denotes the set of pixels belonging to the registered shape, including the pixels on and inside the shape contour  $Y_R^*$ , and  $\oplus$  is the exclusive disjunction operation.

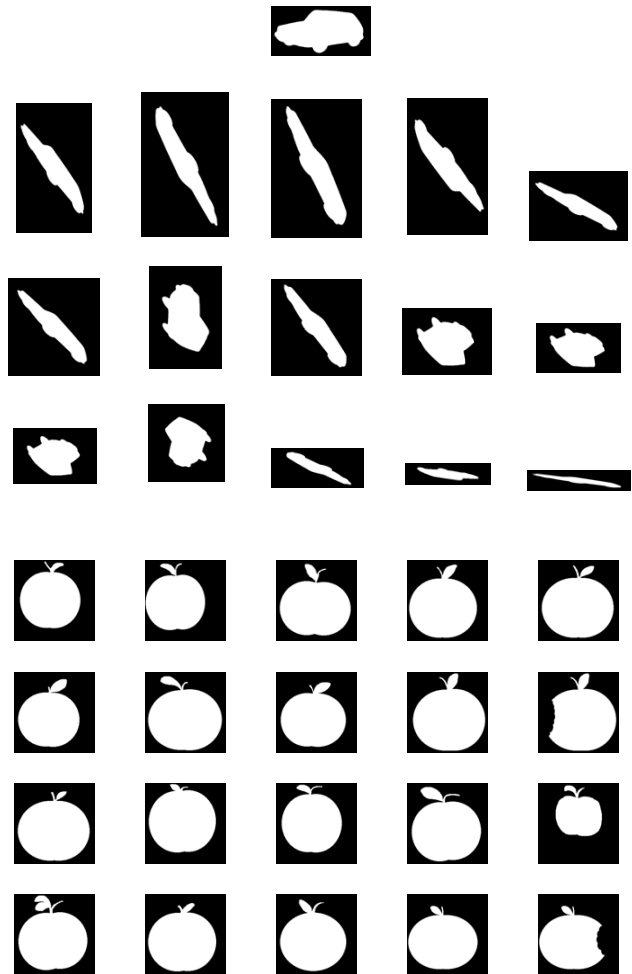
#### 4.2 Registering shapes from binary image databases

We consider two shapes to be “equivalent” if they are images of the same object and are related by an affine transformation. In order to test the proposed approach on binary image databases where the shapes in the same class are related by an affine transformation, two binary image databases are created from the MPEG-7 database: one is an affine shape database, denoted as  $D_1$ , and the other is a sub-database of MPEG-7, denoted as  $D_2$ .

The database  $D_1$  is constructed from the 30 shape classes chosen in MPEG-7. From each shape class, one shape is randomly selected as a template and distorted by 15 affine transformations. Each affine transformation is randomly created according three items: (1) a rotation randomly chosen between  $0^\circ$  and  $360^\circ$ ; (2) shearing parallel to both axes, with shear factors randomly chosen between 0.1 and 1.2;

(3) scaling along both axes, with scale factors randomly chosen between 0.3 and 1.2. Therefore,  $D_1$  has 480 images consisting of 30 different template shapes and their 450 different transformed versions. The contours are extracted by the contour calculation function in MATLAB. Table 1 (top) shows a class example in  $D_1$ , which consists of template shape “personal car” (in the first row) and its 15 random transformed versions (in the remaining three rows). It is worth noting that all the 30 classes have different transformed versions because of the randomly created affine transformations.

**Table 1:** Top: Image of Car in  $D_1$ ; Bottom: Images of Apple in  $D_2$ .



The  $D_2$  database is created by 15 classes in MPEG-7, each having 20 shapes, for a total of 300 images. For each class, the first shape is chosen as the template and the other 19 as targets. Table 1 (bottom) shows a class example in  $D_2$ , which consists of the first shape of “apple” as template, and the other 19 shapes as targets. Note that all 15 classes have different shapes.

For the databases  $D_1$  and  $D_2$ , the proposed method and the reference approach are tested. Table 2 shows three image pairs randomly chosen in each database, and the registra-

tion results with registration error  $\delta$  given below each image for numerical evaluation. The performance comparisons are given in Table 3 (top) and (bottom), respectively.

**Table 2:** Registration Results on Sample Image Pairs in  $D_1$  and  $D_2$ .

Template	Target	Shape context	Area based	Proposed
		 5.88	Unsolved	 1.03
		 9.13	 1.24	 0.49
		 2.49	 5.59	 1.87
		 15.68	 3.60	 2.57
		 0.75	Unsolved	 0.52
		 4.50	 4.00	 0.81

**Table 3:** Top: Performance Comparison on  $D_1$ ; Bottom: Performance Comparison on  $D_2$ .

$D_1$

Method	$\delta$ (%)	Accuracy (%)	Unsolved (%)
Shape context	13.75	31.69	0
Area based	13.19	75.00	18.75
Proposed	5.68	90.28	0

$D_2$

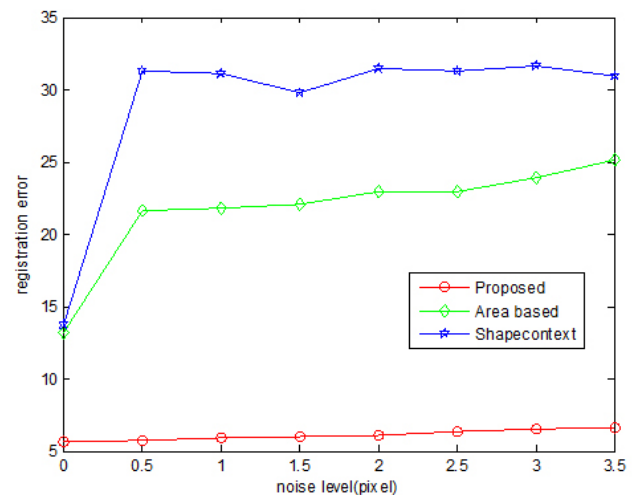
Method	$\delta$ (%)	Accuracy (%)	Unsolved (%)
Shape context	11.94	45.00	0
Area based	15.62	43.16	10.53
Proposed	2.58	100	0

For  $D_1$ , the affine transformations between shape templates of each class and its 15 randomly transformed versions (as targets) are computed. For  $D_2$ , the corresponding affine transformations between the first shape of each class (template) and the other 19 images (targets) are calculated. The items in the tables are registration error  $\delta$  accuracy ratio, and unsolved ratio. The registrations  $\delta$  in Table 3 are computed by averaging over all the 450 calculation results in  $D_1$  and 285 calculation results in  $D_2$ . The accuracy ratio is the number of accurate cases over all cases, where a transformation estimation is deemed to be accurate if the corresponding error satisfies  $\delta \leq 10\%$ . The unsolved ratio is the percentage of cases in which the algorithm finds no solution over all transformation estimation cases.

The comparisons in Table 3 show that the proposed method works better on binary image databases  $D_1$  and  $D_2$ , than the “shape context” and area based methods. The proposed method gives lower average error and higher accuracy ratio than the other methods. In Table 3 (bottom), the accuracy of the proposed method is 100%, which does not mean that the results are always perfect, but accurate for all the cases in  $D_2$ , because  $\delta \leq 10\%$  for every case. It also shows that only the area based method has unsolved cases for  $D_1$  and  $D_2$ , this is because it needs to solve a system of six equations with six unknown coefficients, which probably has no solution in some cases.

### 4.3 Robustness to geometric noise

In this section, Gaussian noise is added to the point coordinates to test the robustness of the proposed method. The 480 images in  $D_1$ , which consisted of 30 different template shapes and their 450 different transformed versions, are used as test data. Gaussian noise of different levels are introduced independently to two coordinates of all the points belonging to the curve of the target shape, and the noise level changes as the standard deviation of Gaussian ranges from 0 to 3.5 pixels in 0.5 pixel increments. On each general level of noise, the proposed method and the two reference methods are tested on 450 pairs of matching cases, and the average registration error  $\delta$  for each method is shown in Figure 3, which is considered as the quality of the three methods. It is clear that the proposed method has the best performance in terms of robustness against geometric noise.


















**Figure 2:** Average registration error versus noise level.

### 4.4 Registration on real images

In this section, the proposed approach is evaluated on real images. Table 4 shows three image pairs, "Balloon", "Green block" and "Jar" that are selected from the Amsterdam Library of Object Images. The regions and the contours are

extracted using parametric kernel graph cuts method<sup>[18]</sup> and contour calculation function in MATLAB respectively. For the extracted shapes, the first and third rows are non-planar, and the second row is approximately planar. The matching results and registration errors for the proposed and the two reference methods are also shown in Table 4. It shows that the proposed approach works well on real shape images too, and the comparison results for real images are consistent with those for binary image databases  $D_1$  and  $D_2$ . This shows that the proposed method has lower registration error than the reference methods. Furthermore, the proposed methods can not only deal with planar shapes, but also non-planar shapes.

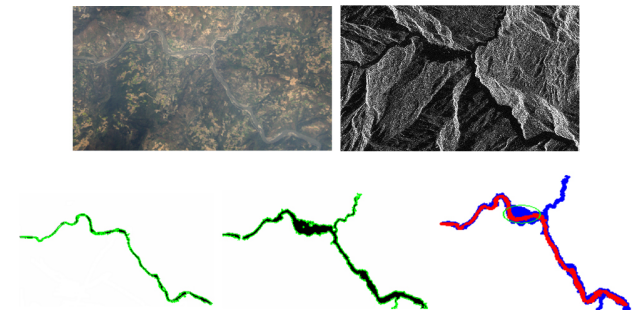
**Table 4:** Registration Results on Real Images Pairs

Template	Target	Shape context	Area based	Proposed
				
		0.26	1.48	0.20
				
		10.83	9.30	7.55
				
		5.76	16.60	3.70

#### 4.5 Experiments on remote sensing image pairs for change detection

In this section, in order to test the matching ability of our algorithm for practical application, two remote sensing image pairs are used to detect change of Tangjiashan barrier lake, which is one of the largest lakes that were created by the Wenchuan earthquake on May 12, 2008.<sup>[19]</sup> Figure 3 (top) shows the geographic location of Tangjiashan and the pre-earthquake image of ALOS AVNIR-2 and the post-earthquake images of Radarsat-1 SAR. The pre-earthquake image was acquired on March 31, 2007, and the post-earthquake image was acquired on May 16, 2008. Because of acquisition by different remote sensors, there are no obvious common features, such as points or lines, which can be used as features for registration. Our aim is to extract the area of water from two images as template shape and target shape, and match the water area by the proposed algorithm to detect the change between the two water areas. The Pyramid segmentation algorithm<sup>[20]</sup> is used to extract the area of water from the images. Figure 3 (bottom) shows the extracted areas of water for pre and post-earthquake and

the registration results between them. The red area is the original river and the blue area is the submerged zone after earthquake, the main flooded zone is marked with a green ellipse, which is the barrier lake induced by the earthquake. Figure 3 shows the changes between the original river and the river after earthquake more clearly. By counting the pixels belonging to the main flooded zone, we can find that the maximum width of the barrier lake increased from 93 m pre-earthquake to 552 m post-earthquake, and the surface area of the barrier lake increased from 0.955 km<sup>2</sup> to 1.485 km<sup>2</sup>.



**Figure 3:** Top: The geographical location of Tangjiashan area. Middle: pre-earthquake image of ALOS AVNIR-2 and the post-earthquake images of Radarsat-1 SAR. Bottom: The extracted areas of water of pre (left) and post-earthquake (middle) with marked curve; right: the registration results between two areas of water.

## 5 Conclusion

In this paper, a matrix projection method to matching a pair of shapes and estimating the affine transformation aligning the shapes was proposed. The main innovation made by the proposed method is the idea of matching a pair of shapes by minimizing an inner product between one matrix and a projection of the other. The proposed method is robust to geometric noise, easy to implement and has a clear mathematical and physical interpretation. It is capable of matching two curves of a pair of shapes by finding correspondences of points on those curves without re-sampling. The proposed approach is compared with two state-of-the-art reference methods, and the performance comparisons show that the proposed method compares favorably to the two reference ones in terms of registration errors, accuracy ratio, and unsolved ratio.

## Acknowledgements

This work is supported by the National Natural Science Foundation of China (Nos. 60972150, 61201323), the Basic Research Foundation of Northwestern Polytechnical University (JC201053), a Chinese Govt. Scholarship.



## References

- [1] Quang, M.T., & Wageeh, W.B. An application of wavelet-based affine representation. *Pattern Recognition Letter*. 1995; 16: 1287-1296. [http://dx.doi.org/10.1016/0167-8655\(95\)00079-1](http://dx.doi.org/10.1016/0167-8655(95)00079-1)
- [2] Tsang, P.W.M. A genetic algorithm for affine invariant recognition of object shapes from broken boundaries. *Pattern Recognition Letter*. 1997; 18: 631-639. [http://dx.doi.org/10.1016/S0167-8655\(97\)00034-2](http://dx.doi.org/10.1016/S0167-8655(97)00034-2)
- [3] Oirrak, A.E., Daoudi, M., & Aboutajdine, D. Affine invariant descriptors using Fourier series. *Pattern Recognition Letters*. 2002; 23(10): 1109-1118. [http://dx.doi.org/10.1016/S0167-8655\(02\)00027-2](http://dx.doi.org/10.1016/S0167-8655(02)00027-2)
- [4] [4]Chaker, F., Bannour, M.T., & Ghorbel, F. Contour retrieval and matching by affine invariant Fourier descriptors. in: *MVA*. 2007: 291-294. Tokyo, Japan
- [5] Khalil, M.I., & Bayoumi, M.M. A dyadic wavelet affine invariant function for 2d shape recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2001; 23(10): 1152-1164. <http://dx.doi.org/10.1109/34.954605>
- [6] Khalil, M.I., & Bayoumi, M.M. Affine invariants for object recognition using the wavelet transform. *Pattern Recognition Letters*. 2002; 23: 57-72. [http://dx.doi.org/10.1016/S0167-8655\(01\)00102-7](http://dx.doi.org/10.1016/S0167-8655(01)00102-7)
- [7] Rube, I.E., Ahmed, M., & Kamel, M.S. Wavelet approximation-based affine invariant shape representation functions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2006; 28(2): 323-327. <http://doi.ieeecomputersociety.org/10.1109/TPAMI.2006.43>
- [8] Heikkil, J. Pattern matching with affine moment descriptors. *Pattern Recognition*. 2004; 37(9): 1825-1834. <http://dx.doi.org/10.1016/j.patcog.2004.03.005>
- [9] Jin, L., Deren, L., Wenbing, T., & Yan, L. An automatic method for generating affine moment invariants. *Pattern Recognition Letters*. 2007; 28: 2295-2304. <http://dx.doi.org/10.1016/j.patrec.2007.07.016>
- [10] Domokos, C., & Kato, Z. Parametric estimation of affine deformations of planar shapes. *Pattern Recognition*. 2010; 43(3): 569-578. <http://dx.doi.org/10.1016/j.patcog.2009.08.013>
- [11] Xuming, H., Bin, W., & Liming, Z. A new scheme for extraction of affine invariant descriptor and affine motion estimation based on independent component analysis. *Pattern Recognition Letters*. 2005; 26(9): 1244-1255. <http://dx.doi.org/10.1016/j.patrec.2004.11.006>
- [12] Guney, N., & Ertuzun, A. Undoing the affine transformation using blind source separation. in: *ICA*. 2006: 360-367. [http://dx.doi.org/10.1007/11679363\\_45](http://dx.doi.org/10.1007/11679363_45)
- [13] Mei, Y., & Androustos, D. Affine invariant shape descriptors: the ica-fourier descriptor and the pca-fourier descriptor. in: *ICPR*. 2008: 1-4. <http://dx.doi.org/10.1109/ICPR.2008.4761381>
- [14] Belongie, S., Malik, J., & Puzicha, J. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2002; 24(4): 509-522. <http://dx.doi.org/10.1109/34.993558>
- [15] Mai, F., Chang, C.Q., & Hung, Y.S. A subspace approach for matching 2D shapes under affine distortions. *Pattern Recognition*. 2011; 44: 210-221. <http://dx.doi.org/10.1016/j.patcog.2010.08.032>
- [16] Alt, H., Melhorn, K., Wagener, H., & Welzl, E. Congruence, similarity, and symmetries of geometric objects. *Discrete Comput. Geom*. 1988; 3(1): 237-256. <http://dx.doi.org/10.1007/BF02187910>
- [17] Josef, S., & Michael, W. Affine point matching. *Pattern Recognition Letter*. 1994; 15: 337-339. [http://dx.doi.org/10.1016/0167-8655\(94\)90081-7](http://dx.doi.org/10.1016/0167-8655(94)90081-7)
- [18] Salah, M.B., Mitiche, A., & Ayed, I.B. Multiregion Image Segmentation by Parametric Kernel Graph Cuts. *IEEE Transaction on Image Processing*. 2011; 20(2): 545-556. <http://dx.doi.org/10.1109/TIP.2010.2066982>
- [19] Min, X., Chunxiang, C., Hao, Z., & Yong, X. Change detection of the Tangjiashan barrier lake based on multi-source remote sensing data. In: *Proceedings of IEEE international geoscience & remote sensing symposium*. 2009; 4: 30-36. <http://dx.doi.org/10.1109/IGARSS.2009.5417373>
- [20] Marfil, R., Molina-Tanco, L., Bandera, A., Rodriguez, J.A., & Sandoval, F. Pyramid segmentation algorithms revisited. *Pattern Recognition*. 2006; 39(8): 1430-1451. <http://dx.doi.org/10.1016/j.patcog.2006.02.017>