A Unifying Framework for Correspondence-Less Linear Shape Alignment

Zoltan Kato*

Department of Image Processing and Computer Graphics,
University of Szeged, P.O. Box 652., 6701 Szeged, Hungary
kato@inf.u-szeged.hu,
http://www.inf.u-szeged.hu/~kato

Abstract. We consider the estimation of linear transformations aligning a known binary shape and its distorted observation. The classical way to solve this registration problem is to find correspondences between the two images and then compute the transformation parameters from these landmarks. Here we propose a unified framework where the exact transformation is obtained as the solution of either a polynomial or a linear system of equations without establishing correspondences. The advantages of the proposed solutions are that they are fast, easy to implement, have linear time complexity, work without landmark correspondences and are independent of the magnitude of transformation.

Keywords: Affine deformation, shape registration, covariant function, Mahalanobis-distance.

1 Introduction

Registration is a crucial step in almost all image processing tasks where images of different views or sensors of an object need to be compared or combined. Typical application areas include visual inspection, target tracking in video sequences, super resolution, or medical image analysis. In a general setting, one is looking for a transformation which aligns two images such that one image (called the *observation*) becomes similar to the second one (called the *template*). Due to the large number of possible transformations, there is a huge variability of the object signature. In fact, each *observation* is an element of the orbit of the transformations applied to the *template*. Hence the problem is inherently *ill-defined* unless this variability is taken into account.

Several techniques have been proposed to address the affine registration problem, a good survey of these methods can be found in [1]. Basically registration algorithms fall into two main categories: *Feature-based* and *Area-based* methods.

* This research was partially supported by the grant CNK80370 of the National Innovation Office (NIH) & the Hungarian Scientific Research Fund (OTKA); the European Union and co-financed by the European Regional Development Fund within the project TÁMOP-4.2.1/B-09/1/KONV-2010-0005.

A. Campilho and M. Kamel (Eds.): ICIAR 2012, Part I, LNCS 7324, pp. 277–284, 2012.

Feature-based methods [2] aim at establishing point correspondences between two images. For that purpose, they extract some easily detectable features (*e.g.* intersection of lines, corners, etc.) from the images and then use these points to compute the closest transformation based on a similarity metric. Searching for the best transformation usually requires an iterative algorithm like the iterative closest point (ICP) algorithm, which requires that the deformation be close enough to identity. The main drawback of these methods is thus the assumption of a limited deformation and high computational cost. Their main advantage is that as long as a sufficient number of point matches are available, one can usually find an optimal aligning transformation implying that these algorithms are less sensitive to occlusions.

Area-based methods [3,4] treat the problem without attempting to detect salient objects. These methods are sometimes called correlation-like methods because they use a rectangular window to gain some preliminary information about the distortion. They search the position in the observation where the matching of the two windows is the best and then look for sufficient alignment between the windows in the template and in the observation. The drawback of this family of methods is also the high computational cost and the restricted range of distortions.

In many situations, the variability of image features is so complex that the only feasible way to register such images is to reduce them to a binary representation and solve the registration problem in that context. X-ray images are good examples as they usually exhibit highly nonlinear radiometric distortions [5] making registration hard to solve. Therefore binary registration (*i.e.* shape alignment) is an important problem for many complex image analysis tasks.

2 Correspondence-Less Alignment Framework

Let us denote the points of the *template* and the *observation* by $\mathbf{x}, \mathbf{y} \in \mathbb{P}^2$ respectively (*i.e.* we use homogeneous coordinates). $\mathbf{A}$ is the unknown non-singular linear transformation that we want to recover. We can define the identity relation as follows

$$\mathbf{A}\mathbf{x} = \mathbf{y} \quad \Leftrightarrow \quad \mathbf{x} = \mathbf{A}^{-1}\mathbf{y}. \tag{1}$$

If we can observe some image features (*e.g.* gray-level of the pixels [4]) that are invariant under the transformation $\mathbf{A}$ then the following equality also holds

$$g(\mathbf{x}) = h(\mathbf{A}\mathbf{x}) \quad \Leftrightarrow \quad g(\mathbf{A}^{-1}\mathbf{y}) = h(\mathbf{y}). \tag{2}$$

Furthermore, the above equations still hold when an *invariant function* $\omega : \mathbb{R}^n \to \mathbb{R}^n$ is acting on both sides of the equations. Indeed, for a properly chosen ω

$$\omega(\mathbf{x}) = \omega(\mathbf{A}^{-1}\mathbf{y}), \quad \text{and} \tag{3}$$

$$\omega(g(\mathbf{x})) = \omega(h(\mathbf{A}\mathbf{x})) = \omega(h(\mathbf{y})). \tag{4}$$

The basic idea of the proposed approach is to generate enough linearly independent equations by making use of the relations in Eq. (1)–(4). Furthermore, we can get rid of the need for point correspondences by integrating both sides of the equations over the corresponding segmented domains.

Recently, similar ideas have been successfully applied to graylevel image registration [4], where one can make use of rich radiometric features to build a system of linear equations. However, these approaches cannot be used in the binary case due to the missing radiometric information. Therefore the main challenge of the proposed approach is to find a way to construct a direct method to estimate linear deformations without making use of feature correspondences or complex optimization algorithms. In this project, we propose two ways to tackle this fundamental problem.

2.1 Solution via a Polynomial System

The first one [6,7] makes use of Eq. (3) and constructs a system of polynomial equations

$$\int \omega(\mathbf{x})d\mathbf{x} = \frac{1}{|\mathbf{A}|}\int \omega(\mathbf{A}^{-1}\mathbf{y})d\mathbf{y}. \tag{5}$$

Obviously, the choice of ωs is crucial as our goal is to construct a system which can be solved. It is easy to see that a polynomial system, which is certainly straightforward to solve, is obtained when $\omega(x) = x^i$. From a geometric point of view, for $\omega(x) \equiv x$ Eq. (5) simply matches the center of mass of the *template* and *observation* while for $\omega(\mathbf{x}) = [x_1^i, x_2^i, 1]^T$ Eq. (5) matches the center of mass of the shapes obtained by the nonlinear transformations ω. In the affine case, we have to solve a system of polynomial equations of the following form, where q_{ki} denotes the unknown elements of the inverse transformation $\mathbf{A}^{-1}$

$$|\mathbf{A}|\int x_k = q_{k1}\int y_1 + q_{k2}\int y_2 + q_{k3}\int 1, \tag{6}$$

$$\begin{aligned}|\mathbf{A}|\int x_k^2 = {} & q_{k1}^2\int y_1^2 + q_{k2}^2\int y_2^2 + q_{k3}^2\int 1 \\ & + 2q_{k1}q_{k2}\int y_1y_2 + 2q_{k1}q_{k3}\int y_1 \\ & + 2q_{k2}q_{k3}\int y_2,\end{aligned} \tag{7}$$

$$\begin{aligned}|\mathbf{A}|\int x_k^3 = {} & q_{k1}^3\int y_1^3 + q_{k2}^3\int y_2^3 + q_{k3}^3\int 1 \\ & + 3q_{k1}^2q_{k2}\int y_1^2y_2 + 3q_{k1}^2q_{k3}\int y_1^2 \\ & + 3q_{k2}^2q_{k3}\int y_2^2 + 3q_{k1}q_{k2}^2\int y_1y_2^2 \\ & + 3q_{k2}q_{k3}^2\int y_2 + 3q_{k1}q_{k3}^2\int y_1 \\ & + 6q_{k1}q_{k2}q_{k3}\int y_1y_2.\end{aligned} \tag{8}$$

The above system of equations can be readily solved either by a direct solver found *e.g.* in Matlab [6] or by a classical LSE solver like the *Levenberg-Marquardt* algorithm [7].

2.2 Solution via a Linear System Using Covariant Functions

The second solution [8,9] is based on Eq. (4). The advantage of this approach is that it yields a linear system of equations which is numerically much more stable. The key idea is to construct two *covariant functions* satisfying Eq. (2). Once this is achieved, we can construct a linear system using Eq. (4) to solve for the unknown transformation $\mathbf{A}$. Since we do not have any radiometric information, this is a quite challenging task as we have to define these functions based on the only available geometric information. For example, we can consider the points of the *template* as a sample from a normally distributed random variable $X \sim N(\mu, \Sigma)$. It is well known, that for any linear transformation, when $Y = \mathbf{A}X$ then Y has also a normal distribution

$$X \mapsto Y \sim N(\mu', \Sigma') = N(\mathbf{A}\mu, \mathbf{A}\Sigma\mathbf{A}^T), \tag{9}$$

furthermore

$$p'(\mathbf{y}) = \frac{1}{|\mathbf{A}|} p(\mathbf{x}), \tag{10}$$

where p' and p are the Gaussian density functions. It is clear that p and p' are *covariant* and the Jacobian can also be computed as $|\mathbf{A}| = \sqrt{|\Sigma'|/|\Sigma|}$. Obviously, the above relation is only valid when $\mathbf{A}$ is positive definite. The parameters of the probability densities $N(\mu, \Sigma)$ and $N(\mu', \Sigma')$ can be easily estimated as the sample means and covariances (*i.e.* the mean and covariance of the point coordinates). From a geometric point of view, the mean values μ and μ' represent the center of mass of the *template* and *observation* respectively, while Σ and Σ' capture the orientation and eccentricity of the shapes. Note that the densities p' and p can be further reduced to the corresponding Mahalanobis distances g and h

$$\begin{aligned} g(\mathbf{x}) &= (\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) \quad \text{and} \\ h(\mathbf{y}) &= (\mathbf{y} - \mu')^T \Sigma'^{-1} (\mathbf{y} - \mu'). \end{aligned} \tag{11}$$

New equations can then be generated by making use of appropriate *invariant functions* $\omega : \mathbb{R} \to \mathbb{R}$. Thus we get

$$\begin{aligned} \int \mathbf{x}\omega(g(\mathbf{x}))d\mathbf{x} &= \int \mathbf{x}\omega(h(\mathbf{A}\mathbf{x}))d\mathbf{x} = \\ &= \frac{1}{|\mathbf{A}|} \int \mathbf{A}^{-1}\mathbf{y}\omega(h(\mathbf{y}))d\mathbf{y}. \end{aligned} \tag{12}$$

Theoretically any *invariant function* could be applied. For example the following set of functions gave us good results [8]: x, $\cos(x)$, $\cos(2x)$, $\sin(x)$ and $\sin(2x)$

(see Fig. 2 for another function set). In the affine case, we can write the linear system in matrix form

$$\begin{pmatrix} \int y_1\omega_1(h(\mathbf{y})) & \int y_2\omega_1(h(\mathbf{y})) & \int \omega_1(h(\mathbf{y})) \\ \vdots & \vdots & \vdots \\ \int y_1\omega_n(h(\mathbf{y})) & \int y_2\omega_n(h(\mathbf{y})) & \int \omega_n(h(\mathbf{y})) \end{pmatrix} \begin{pmatrix} q_{k1} \\ q_{k2} \\ q_{k3} \end{pmatrix}$$

$$= |\mathbf{A}| \begin{pmatrix} \int x_k\omega_1(g(\mathbf{x})) \\ \vdots \\ \int x_k\omega_n(g(\mathbf{x})) \end{pmatrix}. \tag{13}$$

The solution of the above linear system directly provides the parameters of the aligning transformation [8,9].

Compound Objects. When we have objects composed of several parts, yielding a group of disjoint shapes when segmented, the topology of such compound shapes will not change under the action of the affine group. Thus we can construct covariant functions $\mathcal{P}_i(\mathbf{x})$, $\mathcal{S}_i(\mathbf{y})$ for each pair of these shape parts and then sum these relations yielding [9]

$$P(\mathbf{x}) \equiv \sum_{i=0}^{m} \mathcal{P}_i(\mathbf{x}) = \sum_{i=0}^{m} \mathcal{S}_i(\mathbf{y}) \equiv S(\mathbf{y}). \tag{14}$$

where $\mathcal{P}_i(\mathbf{x}) = \exp\Big(-\frac{1}{2}(\mathbf{x}-\mu_i)^T \Sigma_i^{-1}(\mathbf{x}-\mu_i)\Big)$ are unnormalized Gaussians.

The big advantage of this representation is that we can get rid of the segmented shape used as the integration domain in Eq. (12). A clear disadvantage of using the segmented shape itself as the domain [4,8] is that any segmentation error will inherently result in erroneous integrals causing misalignment. In the case of compound shapes, however, it is natural to chose one of the ellipses of the density fitted to the overall shape as the integration domain (see Fig. 3). An additional benefit is that these domains are analytical hence the computation of the integrals used as coefficients in Eq. (13) can be computed efficiently by a closed form formula.

3 Experiments and Discussion

The methods have been quantitatively evaluated on a set of more than 1000 synthetically generated observations for 60 different shapes. The applied transformations were randomly composed of $0°, 10°, \ldots, 350°$ rotations; $0, 0.4, \ldots, 1.2$ shearings; $0.5, 0.7, \ldots, 1.9$ scalings, and $-20, 0, 20$ translations along both axes. The resulting images are of size $\approx 1400 \times 1400$. For evaluation, we have computed two error measures: the Dice coefficient as $\delta = \frac{|R \triangle O|}{|R|+|O|} \cdot 100\%$, where $\triangle$ denotes symmetric difference, while T, R and O are the pixels of the *template*, *registered* object and *observation* respectively; and $\epsilon = \frac{1}{|T|}\sum_{\mathbf{p}\in T} \|(\mathbf{A}-\tilde{\mathbf{A}})\mathbf{p}\|$, which measures the average distance between the true $\mathbf{A}$ and the estimated $\tilde{\mathbf{A}}$ transformation. All algorithms were implemented in Matlab. These results are shown in Table 1. Based on these numbers, it is clear that the polynomial solution provides rather good alignments at the price of $\approx$ 1sec. CPU time. The linear system based on a single pair of covariant functions given in Eq. (11) works well when there are no segmentation errors, but deteriorates quickly when pixels are missing. On the other hand, the linear system with multiple pairs of covariant functions given in Eq. (14) clearly outperforms the polynomial solution in terms of CPU time as well as in robustness: even for 90% missing pixels [9], this method still provides acceptable alignments while the polynomial system fails over 50% [6]. We also remark, that -like any other area based method- both approaches are quite sensitive to occlusions as it yields large errors in the system of equations.

In Fig. 1 and Fig. 2, we show registration results on real X-ray images. These results also confirm the higher precision of the polynomial system. While the multiple covariant function approach cannot be applied on these images since we only have a single shape, Fig. 3 shows the alignment of traffic signes where -due to the compound shape of these signs- the multiple covariant function approach works pretty well.

Table 1. Registration results on a benchmark dataset of synthetic shapes

	Runtime (sec.)	ϵ (pixel)	δ (%)
Polynomial	0.98	0.51	0.15
Linear	1.5	5.42	2.6
Mult. covar. functions	0.33	0.54	0.19

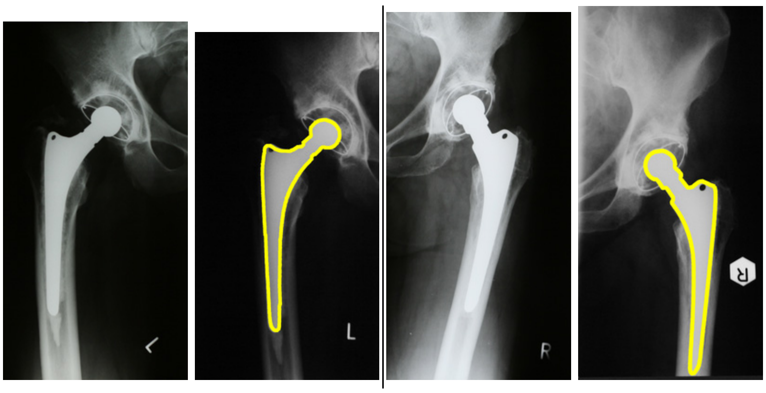

Fig. 1. Alignment of hip prosthesis X-ray images using a polynomial system of equations with ω functions $\{x, x^2, x^3\}$. Registration results are shown as an overlayed contour on the second image.

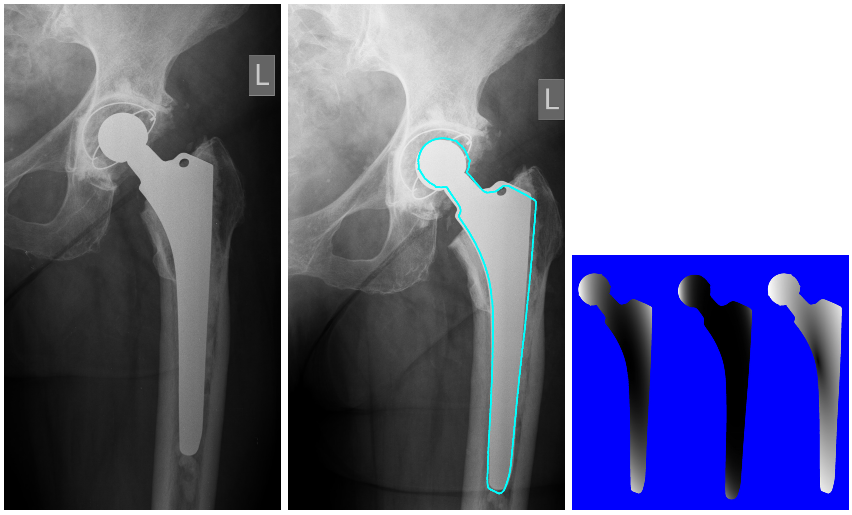

Fig. 2. Alignment of a hip prosthesis X-ray image using a linear system of equations with ω functions $\{x, x^3, x^{1/3}\}$. Registration result is shown as an overlayed contour on the second image.

Fig. 3. Alignment of a traffic sign images using a linear system of equations with multiple shape parts. The first image shows the elliptic integration domain with the compound covariant function fitted over the *template*. Registration results are shown as an overlayed contour on the second image.

4 Conclusion

We have proposed a unified framework to estimate affine deformations between binary images. The main contribution here is a generic solution without established correspondences and the evaluation and comparison of the related methods in terms of computational efficiency and registration quality. We have also shown results on real images of important application domains. Although we only considered affine transformations, other commonly used linear transformations, like rigid-body or similarity, are special cases of the affine family. The main advantages of the proposed framework are that it does not require any correspondences or time consuming optimization step; it is fast and easy to implement while being insensitive to the strength of deformations.

References

1. Zitová, B., Flusser, J.: Image registration methods: A survey. Image and Vision Computing 21, 977–1000 (2003)
2. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape context. IEEE Transactions on Pattern Analysis and Machine Intelligence 24, 509–522 (2002)
3. Heikkilä, J.: Pattern matching with affine moment descriptors. Pattern Recognition 37, 1825–1834 (2004)
4. Hagege, R., Francos, J.M.: Parametric estimation of multi-dimensional affine transformations:an exact linear solution. In: Proceedings of International Conference on Acoustics, Speech, and Signal Processing, Philadelphia, USA, vol. 2, pp. 861–864. IEEE (2005)
5. Pluim, J.P.W., Likar, B., Gerritsen, F.A. (eds.): WBIR 2006. LNCS, vol. 4057. Springer, Heidelberg (2006)
6. Domokos, C., Kato, Z.: Parametric estimation of affine deformations of planar shapes. Pattern Recognition 43, 569–578 (2010)
7. Tanács, A., Sladoje, N., Lindblad, J., Kato, Z.: Estimation of linear deformations of 3D objects. In: Proceedings of International Conference on Image Processing, Hong Kong, China, pp. 153–156. IEEE (2010)
8. Domokos, C., Kato, Z.: Binary Image Registration Using Covariant Gaussian Densities. In: Campilho, A., Kamel, M. (eds.) ICIAR 2008. LNCS, vol. 5112, pp. 455–464. Springer, Heidelberg (2008)
9. Domokos, C., Kato, Z.: Affine alignment of compound objects: A direct approach. In: Proceedings of International Conference on Image Processing, Cairo, Egypt, pp. 169–172. IEEE (2009)