

NONLINEAR REGISTRATION OF BINARY SHAPES

Jozsef Nemeth, Csaba Domokos, Zoltan Kato

Image Processing and Computer Graphics Department, University of Szeged,
P.O. Box 652, 6701 Szeged, Hungary, Fax: +36 62 546-397

ABSTRACT

A novel approach is proposed to estimate the parameters of a diffeomorphism that aligns two binary images. Classical approaches usually define a cost function based on a similarity metric and then find the solution via optimization. Herein, we trace back the problem to the solution of a system of nonlinear equations which directly provides the parameters of the aligning transformation. The proposed method works without any time consuming optimization step or established correspondences. The advantage of our algorithm is that it is easy to implement, less sensitive to the strength of the deformation, and robust against segmentation errors. The efficiency of the proposed approach has been demonstrated on a large synthetic dataset as well as in the context of an industrial application.

Index Terms— Registration, nonlinear deformation, binary image, industrial inspection.

1. INTRODUCTION

Image registration is a fundamental problem for various image processing tasks (*e.g.* change detection, visual inspection, or medical image analysis). In many cases, nonlinear transformation models [1] (*e.g.* polynomial, elastic, thin plate) are assumed between the images. Most nonrigid registration methods first extract a set of corresponding landmarks and then use information about the correspondences to find the aligning transformation between the images [2]. However, the correspondence problem itself is challenged by strong deformations. Furthermore, many approaches define a cost function based on a similarity metric which is then optimized [3]. Another class of methods adopt a variational approach [4]. Both approaches have a rather high computational cost. Francos *et al.* [5] proposed an elegant solution for estimating a homeomorphism between graylevel images by transforming the matching problem into the solution of a linear system of equations. Although this solution doesn't use point correspondences nor optimization, it relies on the availability of rich radiometric information.

In many cases, however, radiometric information is limited (*e.g.* document images, prints...) or intensity values are heavily distorted (*e.g.* X-ray images). In such cases, assuming segmentation of the objects is available, binary image registration is a valid alternative. Belongie *et al.* proposed a novel similarity metric [6], called *shape context*, which can be used to establish correspondences between contour points and then use them to align shapes by a general thin plate spline model. Another approach, similar to [5], has been proposed to recover affine deformations [7] and planar homographies [8] of binary shapes. The basic idea of [8] is to set up a system of nonlinear equations by integrating a set of nonlinear functions over the image domains and then solve it by classical *Levenberg-Marquardt* algorithm. In this paper, by extending the ideas in [5, 8], a novel method is proposed to estimate the parameters of a diffeomorphism that aligns two binary shapes avoiding both establishing correspondences and the optimization step. We tested the performance and robustness of the algorithm on a large set of synthetic images. Furthermore, the proposed method has been successfully used for visual inspection of signs printed on hoses in a real industrial application.

2. REGISTRATION FRAMEWORK

In the general case, we are looking for the parameters of a $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ diffeomorphism aligning a pair of binary shapes. Let us denote the coordinates of the *template* and *observation* by $\mathbf{x} = [x_1, x_2]^T \in \mathbb{R}^2$ and $\mathbf{y} = [y_1, y_2]^T \in \mathbb{R}^2$, respectively. We then have the following identity relation

$$\mathbf{y} = \varphi(\mathbf{x}) \Leftrightarrow \mathbf{x} = \varphi^{-1}(\mathbf{y}), \quad (1)$$

where $\varphi^{-1} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is the corresponding inverse transformation. The deformation field φ can be decomposed as $\varphi(\mathbf{x}) = [\varphi_1(\mathbf{x}), \varphi_2(\mathbf{x})]^T$, where $\varphi_1, \varphi_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$. Since we are interested in a direct solution without solving the correspondence problem, we will integrate both sides of Eq. (1) over the foreground regions \mathcal{F}_t and \mathcal{F}_o [8]:

$$\int_{\mathcal{F}_o} \mathbf{y} d\mathbf{y} = \int_{\mathcal{F}_t} \varphi(\mathbf{x}) |J_\varphi(\mathbf{x})| d\mathbf{x}, \quad (2)$$

where the integral transformation $\mathbf{y} = \varphi(\mathbf{x})$, $d\mathbf{y} = |J_\varphi(\mathbf{x})| d\mathbf{x}$ has been applied with $|J_\varphi| : \mathbb{R}^2 \rightarrow \mathbb{R}$ denoting the Jacobian

This research was partially supported by the Hungarian Scientific Research Fund (OTKA) – K75637 and *ContiTech Fluid Automotive Hungária Ltd.*

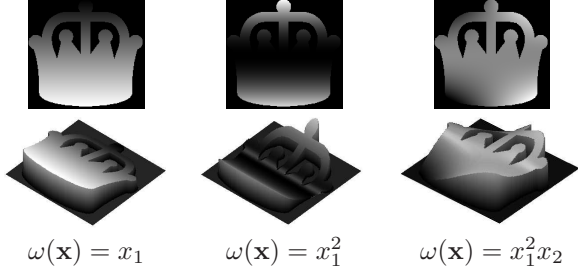


Fig. 1. The effect of the ω functions. **Top:** the generated coloring of a binary shape for various functions. **Bottom:** the corresponding volumes over the shape.

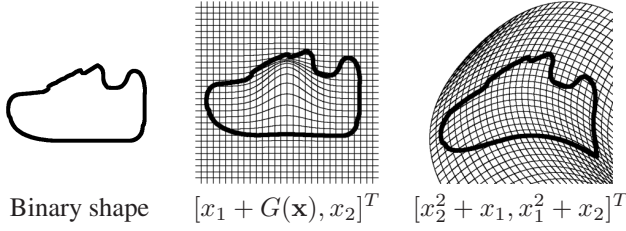


Fig. 2. Various deformation fields (G is a Gaussian function).

of the transformation. Unfortunately, the above two equations alone are not enough to solve for $k > 2$ unknowns.

In order to generate more equations, let us remark that the identity relation in Eq. (1) remains valid when a function $\omega : \mathbb{R}^2 \rightarrow \mathbb{R}$ is acting on both sides of the equation [7, 8]. We thus obtain the following integral equation from Eq. (2)

$$\int_{\mathcal{F}_o} \omega(\mathbf{y}) d\mathbf{y} = \int_{\mathcal{F}_t} \omega(\varphi(\mathbf{x})) |J_\varphi(\mathbf{x})| d\mathbf{x}. \quad (3)$$

The basic idea of the proposed approach is to generate enough linearly independent equations by making use of nonlinear ω functions. Clearly, each applied ω generates one nonlinear equation. Let $\omega_i : \mathbb{R}^2 \rightarrow \mathbb{R}$ ($i = 1, \dots, \ell$) denote the set of adopted nonlinear functions. In order to solve for all unknowns, we need at least k equations, hence $\ell \geq k$. Intuitively, ω generates a consistent coloring of the binary shapes as shown in Fig. 1. From a geometric point of view, Eq. (2) simply matches the center of mass of the *template* and *observation* while Eq. (3) matches the volumes over the shapes constructed by the nonlinear functions ω (see Fig. 1).

While gray-level images allow for a wide range of deformation fields, not every transformation model causes visible distortion on a binary image. For example, the effect of a *localized* deformation is not always observable (see Fig. 2). Therefore in the remaining part of this paper, we will focus on two important class of deformations (polynomial and cylindrical) for which distortions are noticeable on binary shapes.

3. POLYNOMIAL DEFORMATIONS

A broadly used class of deformations is the polynomial family. Let $\varphi(\mathbf{x}) \equiv p(\mathbf{x}) = [p_1(\mathbf{x}), p_2(\mathbf{x})]^T$ be a polynomial deformation field. Without loss of generality, we can assume that $d = \deg(p_1) = \deg(p_2)$, furthermore

$$p_1(\mathbf{x}) = \sum_{i=0}^d \sum_{j=0}^{d-i} a_{ij} x_1^i x_2^j, \quad \text{and} \quad p_2(\mathbf{x}) = \sum_{i=0}^d \sum_{j=0}^{d-i} b_{ij} x_1^i x_2^j.$$

Note that a_{ij} and b_{ij} are the unknown parameters of the transformation and the number of these parameters is $k = (d+2)(d+1)$. The Jacobian of the transformation $|J_p| : \mathbb{R}^2 \rightarrow \mathbb{R}$ used in Eq. (3) is given by

$$\begin{aligned} |J_p(\mathbf{x})| &= \left(\sum_{i=1}^d \sum_{j=0}^{d-i} i a_{ij} x_1^{i-1} x_2^j \right) \left(\sum_{j=1}^d \sum_{i=0}^{d-j} j b_{ij} x_1^i x_2^{j-1} \right) \\ &\quad - \left(\sum_{j=1}^d \sum_{i=0}^{d-j} j a_{ij} x_1^i x_2^{j-1} \right) \left(\sum_{i=1}^d \sum_{j=0}^{d-i} i b_{ij} x_1^{i-1} x_2^j \right), \end{aligned}$$

which is clearly a polynomial. The parameters of the aligning transformation are then simply obtained as the solution of the nonlinear system of equations from Eq. (3). Theoretically any nonlinear function could be applied for constructing our system, but the solution of a system of polynomial equations is numerically more stable. If we use the polynomial transformation

$$\omega_i(\mathbf{x}) = x_1^{n_i} x_2^{m_i} \quad (n_i, m_i \in \mathbb{N}_0), \quad (4)$$

then we will get a system of polynomial equations. Hence Eq. (3) can be written as for all $i = 1, \dots, \ell$

$$\int_{\mathcal{F}_o} y_1^{n_i} y_2^{m_i} d\mathbf{y} = \int_{\mathcal{F}_t} p_1(\mathbf{x})^{n_i} p_2(\mathbf{x})^{m_i} |J_p(\mathbf{x})| d\mathbf{x}. \quad (5)$$

Note, that the left hand side of the equations are constant and can be computed directly using the coordinates of the *observation*. In practice, these integrals are approximated via finite sums, *i.e.* $\int_{\mathcal{F}_o} y_1^q y_2^r d\mathbf{y} \approx \sum_{\mathbf{y} \in \mathcal{F}_o} y_1^q y_2^r$, where \mathcal{F}_o denotes the set of the foreground pixels of the *observation*. Since both $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ as well as $|J_p(\mathbf{x})|$ are polynomes, obviously their multiplication is also a polynome. Thus, the integrals on the right hand side can be written as

$$\int_{\mathcal{F}_t} p_1(\mathbf{x})^{n_i} p_2(\mathbf{x})^{m_i} |J_p(\mathbf{x})| d\mathbf{x} = \sum_{q=0}^{c_i} \sum_{r=0}^{c_i-q} h_{iqr} \int_{\mathcal{F}_t} x_1^q x_2^r d\mathbf{x},$$

where $c_i = d(n_i + m_i) + d(d-1)$ is the degree of the polynome in the i th equation, and h_{iqr} are nonlinear terms consisting of the unknown parameters of the transformation. As we noted before, the integrals $\int_{\mathcal{F}_t} x_1^q x_2^r d\mathbf{x}$ are constants and can be computed directly using the coordinates of the *template*. In fact the i th equation in Eq. (5) is a polynomial equation with degree c_i and the unknowns are the parameters of the aligning transformation.

	Runtime (sec.)			δ (%)		
	m	μ	σ	m	μ	σ
SC [6]	79.75	85.08	30.06	2.48	2.92	1.78
Proposed	15.32	15.62	6.99	0.35	2.14	6.29

Table 1. Registration results (m denotes the median while μ and σ denotes the mean and standard deviation of the values).

Removed pixels (%)	5	10	15	20
Shape Context [6]	25.07	27.03	27.96	28.62
Proposed method	5.62	8.95	11.88	14.45

Size of occlusion (%)	1	2.5	5	10
Shape Context [6]	10.54	10.89	11.32	13.08
Proposed method	3.56	6.32	9.27	14.12

Table 2. Median of δ measures versus amount of randomly removed pixels (top) and size of random occlusions (bottom).

3.1. Experimental results

The proposed method has been tested on a synthetic database of 37 different shapes and their deformed versions, a total of ≈ 1700 images of size 800×800 . A second order ($d = 2$) polynomial deformation field with 12 parameters has been applied to generate random observations with $a_{20}, a_{02}, a_{11}, b_{20}, b_{02}, b_{11}$ parameters chosen randomly from $[-0.5; 0.5]$; a_{10}, b_{01} from $[0.5; 1.5]$; and a_{01}, b_{10} from $[-0.25; 0.25]$. Since image coordinates were normalized into $[-0.5; 0.5]$, we set a_{00} and b_{00} zero. We adopted the ω_i functions in Eq. (4) with exponents $\{(m_i, n_i)\}_{i=1}^{12} = \{(a, b) | 0 \leq a, b \leq 3, 1 < a+b \leq 4\}$, resulting in 12 equations of the form Eq. (5) which was solved by classical *Levenberg-Marquardt* algorithm. The *registered* image was then obtained by applying the resulting transformation to the *template*. The registration error has been quantitatively evaluated based on the absolute difference of the aligned shapes

$$\delta = \frac{|F_r \Delta F_o|}{|F_r| + |F_o|} \cdot 100\%,$$

where F_r is the set of foreground pixels of the *registered* shape and Δ denotes symmetric difference.

We have also compared the performance of our method to that of *Shape Context* (SC) [6], whose parameters were set empirically to their optimal value (`beta_init` = 30, `n_iter` = 30, and annealing rate `r` = 1). A summary of these results is presented in Table 1 and in Fig. 3.

The robustness of the proposed approach against two types of segmentation error has been also tested: In the first case we removed 5%, ..., 20% of the foreground pixels from the *observation* before registration, while in the second case we occluded continuous square shaped regions of size equal to 1%, ..., 10% of the whole shape (see Fig. 3). Table 2

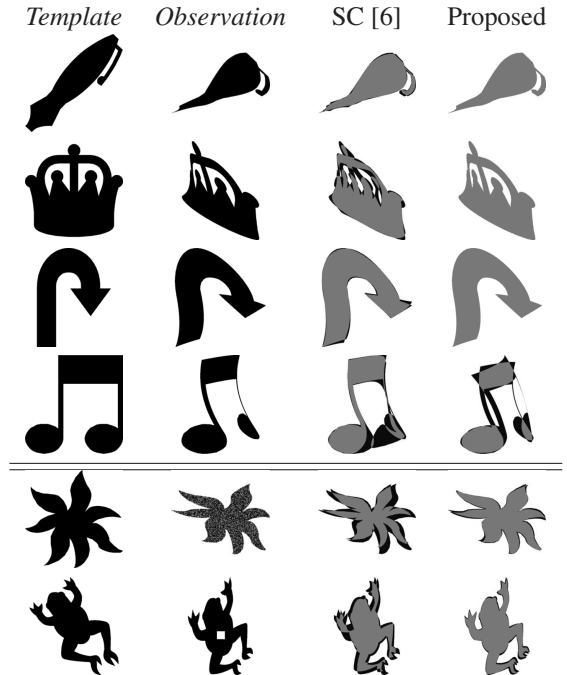


Fig. 3. Synthetic images and registration results. The last two rows show examples from the robustness tests. The *observation* and *registered* shapes were overlaid, overlapping pixels are shown in gray while non-overlapping ones in black.

shows the median of δ measures versus the percentage of segmentation errors and a comparison to the results of SC. As almost all registration methods based on point correspondences, SC is less sensitive to segmentation errors, whereas our method's error rate increases with the segmentation error. Nevertheless, it provides good results up to as high as 5% removed pixels and 2.5% occlusions; and it outperforms SC up to 20% removed pixels and 5% occlusions.

4. APPLICATION: VISUAL INSPECTION

An important step in hose manufacturing for automotive industry is to print various signs on the hose surface in order to facilitate installation (see Fig. 4). The quality control of this process involves visual inspection of the printed signs. In an automated inspection system, this can be implemented by comparing images of the printed sign to its *template*, which requires the alignment of the *template* and *observation* shapes. The main challenges are segmentation errors and complex distortions. The physical model of the contact printing procedure is as follows: (1) the stamp (basically a planar template of the sign) is positioned on the hose surface; (2) then it is pressed onto the surface. This can be described by a 2D rotation and scaling $\mathbf{S} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ of the *template* followed by a transformation $\gamma : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ that maps the

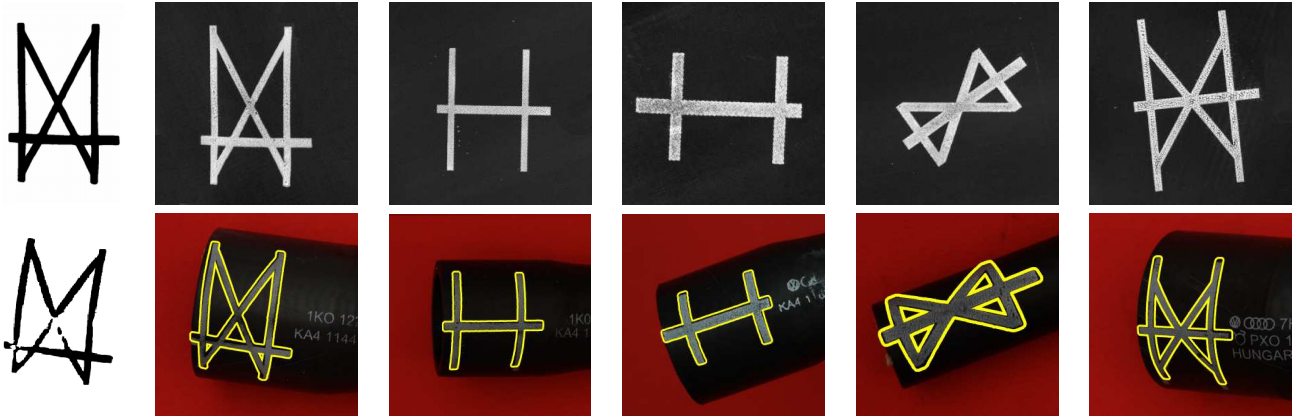


Fig. 4. Registration results of printed signs. **Top:** The images used as *templates*. **Bottom:** the corresponding *observations* with the overlaid contour of the registration results. The first image pair shows the segmented regions used for registration. Note the typical segmentation errors of the *observation*. (Images provided by *ContiTech Fluid Automotive Hungaria Ltd.*)

template's plane to a cylinder with radius r :

$$\gamma(\mathbf{x}) = \left[r \sin \frac{x_1}{r}, x_2, -r \cos \frac{x_1}{r} \right]^T.$$

Then a picture is taken with a camera, which is described by a classical projective transformation $\mathbf{P} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ (basically the camera matrix). Thus the transformation $\mathbf{P} \circ \gamma \circ \mathbf{S}$ acting between a planar *template* and its distorted *observation* has 11 parameters: \mathbf{S} has 3 parameters, γ has one (r), and \mathbf{P} has 7 (six extrinsic parameters and the focal length). We have empirically tested a number of potential $\{\omega_i\}$ sets and found that the following gives the best result:

$$\omega_i(\mathbf{x}) = (x_1 \cos \alpha_i - x_2 \sin \alpha_i)^{n_i} (x_1 \sin \alpha_i + x_2 \cos \alpha_i)^{m_i}$$

for $i = 1, \dots, 12$, using all combinations for $\alpha_i \in \{0, \frac{\pi}{6}, \frac{\pi}{3}\}$ and $(n_i, m_i) \in \{(1, 2), (2, 1), (1, 3), (3, 1)\}$. The method has been tested on more than 150 real images and it proved to be efficient in spite of segmentation errors and severe distortions.

5. CONCLUSION

In this paper we have introduced a novel featureless method to estimate the parameters of a nonlinear transformation that aligns two binary images. We have demonstrated the performance of the method on a large synthetic dataset and analyzed its robustness against various segmentation errors. The method has been validated in the context of visual inspection of printed signs.

6. REFERENCES

- [1] L. Zagorchev and A. Goshtasby, "A comparative study of transformation functions for nonrigid image registration," *IEEE Trans. on Image Processing*, vol. 15, pp. 529–538, March 2006.
- [2] H. Guo, A. Rangarajan, S. Joshi, and L. Younes, "Non-rigid registration of shapes via diffeomorphic point matching," *IEEE Int. Symposium on Biomedical Imaging: Nano to Macro*, vol. 1, pp. 924–927, April 2004.
- [3] M. S. Hansen, M. F. Hansen, and R. Larsen, "Diffeomorphic statistical deformation models," in *Proc. of Int. Conf. on Computer Vision*, Rio de Janeiro, Brazil, October 2007, IEEE, pp. 1–8.
- [4] S. Marsland and C. J. Twining, "Constructing diffeomorphic representations for the groupwise analysis of non-rigid registrations of medical images," *IEEE Trans. on Medical Imaging*, vol. 23, pp. 1006–1020, August 2004.
- [5] J. M. Francos, R. Hagege, and B. Friedlander, "Estimation of multidimensional homeomorphisms for object recognition in noisy environments," in *Proc. of Conference on Signals, Systems and Computers*, Pacific Grove, California, USA, November 2003, IEEE, vol. 2, pp. 1615–1619.
- [6] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape context," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 509–522, April 2002.
- [7] C. Domokos, Z. Kato, and J. M. Francos, "Parametric estimation of affine deformations of binary images," in *Proc. of Int. Conf. on Acoustics, Speech, and Signal Processing*, Las Vegas, Nevada, USA, April 2008, pp. 889–892, IEEE.
- [8] J. Nemeth, C. Domokos, and Z. Kato, "Recovering planar homographies between 2D shapes," in *Proc. of Int. Conference on Computer Vision*, Kyoto, Japan, September 2009, accepted.