

University of Szeged
Institute of Informatics

Parametric Estimation of Affine Deformations without Correspondences

Summary of the Ph.D. Dissertation

by

Csaba Domokos

Advisor:

Dr. Zoltán Kató

Szeged

2010

Introduction

Image registration is a crucial step in almost every image processing application where images of different views or sensors of an object need to be compared or combined. Typical application areas include object recognition, target tracking in video sequences, monitoring land usage on satellite images, super resolution, image mosaicking, and 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, the problem is inherently *ill-defined* unless this variability is taken into account.

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 (Fig. 1) are good examples as they usually exhibit highly nonlinear radiometric distortions, making registration hard to solve. If perfect greylevel images were available, the estimation of an aligning transformation could be reduced to solving of a linear system of equations [1; 2]. In real applications, however, such a strict requirement cannot be satisfied. We will show that registration can be solved without making use of any intensity information. In real situations, the images obtained are related by a projective transformation (also called the planar homography). Although the projective transformation is nonlinear, it can often be successfully modelled by an affine transformation, which is linear. Owing to its linear property, the affine transformation is of great importance in image registration.

This thesis summarizes the author's research results in binary image registration without correspondences, where affine transformations are considered.

Basic Solution

We will introduce our preliminary notations and also formulate the basic solution of the problem. Let us denote the homogenous coordinates of the n dimensional *template* and *observation* points by $\mathbf{x} = [x_1, x_2, \dots, x_n, 1]^T$, $\mathbf{y} = [y_1, y_2, \dots, y_n, 1]^T \in \mathbb{P}^n$, respectively. The *identity relation* between the shapes is then as follows [3; 4]

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

where \mathbf{A} is the unknown affine transformation that we want to recover:

$$\mathbf{A} = \begin{bmatrix} a_{11} & \dots & a_{1n} & a_{1(n+1)} \\ \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \dots & a_{nn} & a_{n(n+1)} \\ 0 & \dots & 0 & 1 \end{bmatrix}, \quad \text{and} \quad \mathbf{A}^{-1} = \begin{bmatrix} q_{11} & \dots & q_{1n} & q_{1(n+1)} \\ \vdots & \ddots & \vdots & \vdots \\ q_{n1} & \dots & q_{nn} & q_{n(n+1)} \\ 0 & \dots & 0 & 1 \end{bmatrix}.$$

Here we are interested in a *direct approach* that does not involve solving the correspondence problem.

When we observe some image features (e.g. grey-levels of the pixels [2]) that are invariant under this transformation, we can define an additional relation

$$f(\mathbf{x}) = g(\mathbf{Ax}) = g(\mathbf{y}), \quad (2)$$

where $f, g : \mathbb{P}^n \rightarrow \mathbb{R}$ are *covariant functions* under the transformation \mathbf{A} . The above relations are still valid when a function acts on both sides of Eq. (1) [2; 3; 4] and Eq. (2) [5; 6; 7]. Indeed, for a properly chosen $\omega_p : \mathbb{P}^n \rightarrow \mathbb{R}$ and $\omega_c : \mathbb{R} \rightarrow \mathbb{R}$, we get

$$\omega_p(\mathbf{y}) = \omega_p(\mathbf{Ax}), \quad \text{and} \quad (3)$$

$$\omega_c(g(\mathbf{y})) = \omega_c(g(\mathbf{Ax})) = \omega_c(f(\mathbf{x})). \quad (4)$$

Hence we can generate as many linearly independent equations as needed by making use of the nonlinear ω_p and ω_c functions. The nonlinear function ω_p acts directly on the point coordinates and hence on the unknown parameters of \mathbf{A} and results in a *nonlinear* system of equations [3; 4]; in contrast ω_c acts on the *covariant functions* f and g that give rise to a *linear* system of equations [2; 5; 6]. Now let concrete shapes be represented by their characteristic function $\mathbb{1} : \mathbb{P}^n \rightarrow \{0, 1\}$, where 0 and 1 correspond to the background and foreground, respectively. If we denote the *template* by $\mathbb{1}_t$ and the *observation* by $\mathbb{1}_o$, then Eq. (1) implies that [3; 4]

$$\mathbb{1}_t(\mathbf{x}) = \mathbb{1}_o(\mathbf{Ax}) = \mathbb{1}_o(\mathbf{y}). \quad (5)$$

Parametric Estimation of Affine Deformations of Binary Images: Polynomial Solution

Here we propose a novel approach where the exact transformation (aligning a known 2D shape and its distorted *observation*), is found from the solution of a polynomial system of equations. Classical *feature-based* approaches would identify point pairs and solve the system of linear equations in Eq. (1). However, we are interested in a *direct solution* without solving the correspondence problem.

We will start with the *identity relation* defined by Eq. (1) and we will take the Lebesgue integral of both sides of the *identity relation* [3; 4]

$$\int_{\mathbb{P}^n} \mathbf{x} \, d\mathbf{x} = \frac{1}{|\mathbf{A}|} \int_{\mathbb{P}^n} \mathbf{A}^{-1}\mathbf{y} \, d\mathbf{y} , \quad (6)$$

where the integral transformation $\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$, $d\mathbf{x} = |\mathbf{A}^{-1}|d\mathbf{y}$ has been applied. The determinant $|\mathbf{A}|$ is the Jacobian and we can assume that $|\mathbf{A}|$ is always positive. Furthermore, it can be evaluated by integrating [3; 4]

$$\int_{\mathbb{P}^n} \mathbb{1}_t(\mathbf{x}) \, d\mathbf{x} = \frac{1}{|\mathbf{A}|} \int_{\mathbb{P}^n} \mathbb{1}_o(\mathbf{y}) \, d\mathbf{y} \quad \Leftrightarrow \quad |\mathbf{A}| = \frac{\int_{\mathbb{P}^n} \mathbb{1}_o(\mathbf{y}) \, d\mathbf{y}}{\int_{\mathbb{P}^n} \mathbb{1}_t(\mathbf{x}) \, d\mathbf{x}} ,$$

which can be directly computed from the input images. Since the characteristic functions only take values from $\{0, 1\}$, we can further simplify the above integrals [3; 4]:

$$\int_{\mathbb{P}^n} \mathbb{1}_t(\mathbf{x}) \, d\mathbf{x} \equiv \int_{\mathcal{F}_t} d\mathbf{x} ,$$

where the finite domain \mathcal{F}_t consists of the *template* foreground regions: $\mathcal{F}_t = \{\mathbf{x} \in \mathbb{P}^n | \mathbb{1}_t(\mathbf{x}) = 1\}$. Similarly, we can restrict the integral of $\mathbb{1}_o(\mathbf{y})$ to the *observation* foreground regions \mathcal{F}_o . Now multiplying Eq. (6) by Eq. (5), we get a finite integral equation [3; 4]:

$$\int_{\mathcal{F}_t} \mathbf{x} \, d\mathbf{x} = \frac{1}{|\mathbf{A}|} \int_{\mathcal{F}_o} \mathbf{A}^{-1}\mathbf{y} \, d\mathbf{y} . \quad (7)$$

This equation implies that the finite domains \mathcal{F}_t and \mathcal{F}_o are also related via $\mathcal{F}_o = \mathbf{A}\mathcal{F}_t$; *i.e.* we match the shapes as a whole instead of point to point correspondences. We will use Eq. (3) and get the following integral equation from Eq. (7) [3; 4]

$$\int_{\mathcal{F}_t} \omega(\mathbf{x}) \, d\mathbf{x} = \frac{1}{|\mathbf{A}|} \int_{\mathcal{F}_o} \omega(\mathbf{A}^{-1}\mathbf{y}) \, d\mathbf{y} . \quad (8)$$

The basic idea behind this approach is to generate enough linearly independent equations by

making use of ω functions. Note, however, that the generated equations contain no new information; they simply impose new linearly independent constraints.

Proposition 1.1 *Let $\omega : \mathbb{P}^n \rightarrow \mathbb{P}^n$ and $\mathbf{x} \in \mathbb{P}^n$ ($n \in \mathbb{N}$). If the k^{th} coordinate of $\omega(\mathbf{x})$, denoted by $\omega^{(k)}(\mathbf{x}) = p_k$ is a real n -variate polynome, $1 \leq k \leq n$, then applying ω in Eq. (8) results in a polynomial system of equations of degree $\deg(p_k)$ at most.*

From now on we will just consider the 2-dimensional case. The class of x^l ($l \in \mathbb{N}_0$) functions is an ideal choice for ω . We have the following polynomial equations for $k = 1, 2$ [3; 4]:

$$|\mathbf{A}| \int_{\mathcal{F}_t} x_k^l \, d\mathbf{x} = \sum_{i=1}^l \binom{l}{i} \sum_{j=0}^i \binom{i}{j} q_{k1}^{l-i} q_{k2}^{i-j} q_{k3}^j \int_{\mathcal{F}_o} y_1^{l-i} y_2^{i-j} \, d\mathbf{y} \quad l = 1, 2, 3. \quad (9)$$

Although we have constructed our equations in the continuum, in practice the integrals can only be *approximated* by a discrete sum. The images need to be scanned only once, and the integrals as well as the Jacobian can be evaluated during this scan. It is clear that the solution is obtained in a single pass without any loop or optimization.

Now we assume that the observed point coordinates are around the true ones, hence we will consider i.i.d. additive Gaussian noise model on the *observation* coordinates. The *identity relation* Eq. (1) thus becomes [4]

$$\mathbf{y}^* = \mathbf{y} + \varepsilon(\mathbf{y}) = \mathbf{A}\mathbf{x} + \varepsilon^*(\mathbf{y}^*) \quad \Leftrightarrow \quad \mathbf{x} = \mathbf{A}^{-1}(\mathbf{y}^* - \varepsilon^*(\mathbf{y}^*)),$$

where $\varepsilon(\mathbf{y}) \equiv \varepsilon^*(\mathbf{y}^*) = [\varepsilon_1^*(\mathbf{y}^*), \varepsilon_2^*(\mathbf{y}^*), 0]^T$ is the noise function which gives a random translation at each point $\mathbf{y}^* = [y_1^*, y_2^*, 1]^T$. Overall, we can say that the error caused by i.i.d. additive Gaussian noise with standard deviance σ_1, σ_2 on the point coordinates of the *observation* is as follows [4]:

Equation	Error term
$ \mathbf{A} $:	0
$\omega(\mathbf{x}) = \mathbf{x}$:	0
$\omega(\mathbf{x}) = [x_1^2, x_2^2, 1]^T$:	$q_{k1}^2 \sigma_1^2 + q_{k2}^2 \sigma_2^2$
$\omega(\mathbf{x}) = [x_1^3, x_2^3, 1]^T$:	$3q_{k3}(q_{k1}^2 \sigma_1^2 + q_{k2}^2 \sigma_2^2)$

The proposed algorithm was tested on a large database of binary images of size 1000×1000 . The dataset consisted of 56 different shapes and their transformed versions, giving a total of over 50 000 images. In order to quantitatively evaluate the registration results, we defined two kinds of error measures:

$$\epsilon = \frac{1}{|F_t|} \sum_{\mathbf{p} \in F_t} \|(\mathbf{A} - \hat{\mathbf{A}})\mathbf{p}\|, \quad \text{and} \quad \delta = \frac{|F_r \Delta F_o|}{|F_r| + |F_o|} \cdot 100\%, \quad (10)$$

where Δ means the symmetric difference, while F_t , F_o and F_r denote the set of pixels of the *template*, *observation* and *registered* shape, respectively. We also examined the robustness of the proposed approach in the case of incomplete objects. The results show that our method is quite robust, while the error rate of other state-of-the-art methods increases considerably.

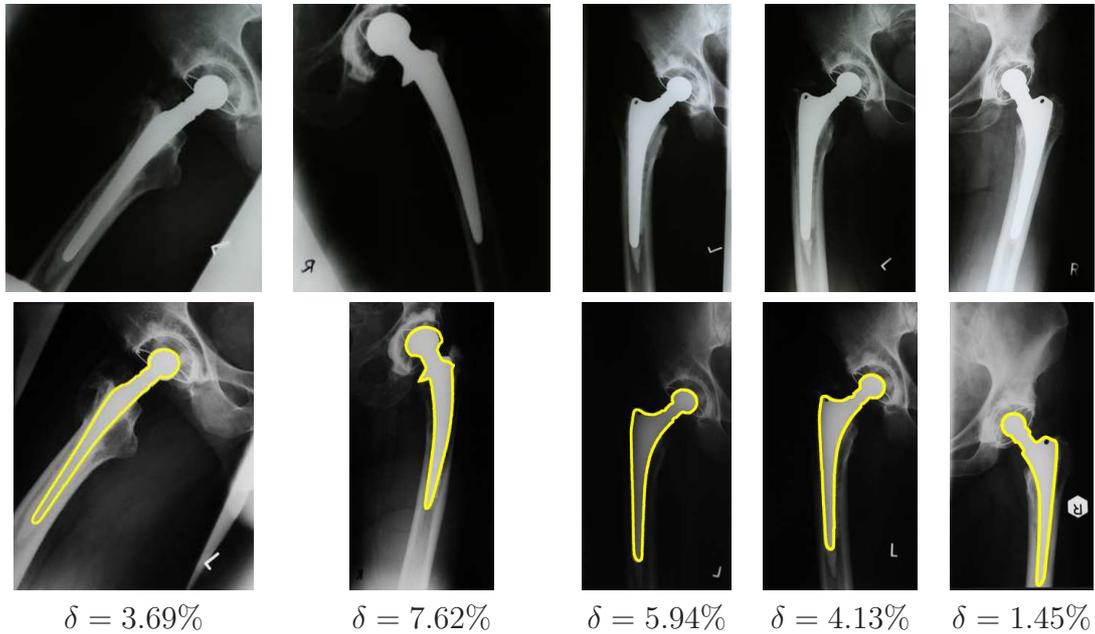


Figure 1: Registration of hip prosthesis X-ray images. The overlaid contour in the second row shows the aligned contour of the corresponding image in the first row.

The fundamental difference between classical image registration algorithms and ours is that our model works without any landmark, feature detection or optimization, incorporating a novel idea where the transformation is obtained as a solution of a set of polynomial equations. It uses all the available information in the input images and provides an exact solution regardless of the magnitude of the transformations. The experimental results show that the proposed method provides good alignment on both real and synthetic images (see Table 1). Moreover, it is robust in the case of noisy *observations*. The method was successfully applied in the registration of hip prosthesis X-ray images (Fig. 1).

Table 1: Registration results of the proposed method on 49 282 synthetic observations of 56 shapes. There was no solution in 5.47% of the test cases.

	ϵ (pixel)	δ (%)	CPU time (sec.)
Median	0.51	0.15	0.98
Mean	36.98	3.36	0.94
Variance	154.18	12.55	0.2

Affine Shape Alignment Using Covariant Gaussian Densities

Here, we propose a novel approach for the estimation of 2D affine transformations, where the exact transformation is got from a least-squares solution of a linear system of equations. It is constructed by fitting Gaussian densities to the shapes which preserve the effect of the unknown transformation. The crucial step of the proposed approach is to construct a pair of *covariant functions* satisfying Eq. (2). Unfortunately, the construction of such functions for binary images is quite a challenging task due to the lack of radiometric information.

Here we need to use inhomogeneous representations of the coordinates, *i.e.* $\mathbf{x} = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n$ and $\mathbf{y} = [y_1, y_2, \dots, y_n]^T \in \mathbb{R}^n$. Hence the *identity relation* Eq. (1) becomes [5; 6; 7]

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{t} \quad \Leftrightarrow \quad \mathbf{x} = \mathbf{A}^{-1}(\mathbf{y} - \mathbf{t}) = \mathbf{A}^{-1}\mathbf{y} - \mathbf{A}^{-1}\mathbf{t}, \quad (11)$$

where $(\mathbf{A}, \mathbf{t}) \in (\mathbb{R}^{n \times n} \times \mathbb{R}^{n \times 1})$ is the unknown affine transformation that we want to recover. We can safely consider the points of the *template* as a sample from a bivariate normally distributed random variable denoted by $X \sim N(\mu, \Sigma)$ with probability density function (PDF) [5; 6; 7]:

$$p(\mathbf{x}) = \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^{-1}(\mathbf{x} - \mu)\right).$$

We make use of the Gaussian PDF, which is over the shape, to construct a *covariant function pair*. We should add that the shape is not represented or modelled by a Gaussian; just the mean and covariance value of the points are estimated. Applying any linear transformation to X also results in a normal random variable $Y = \mathbf{A}X + \mathbf{t}$ with parameters $X \xrightarrow{(\mathbf{A}, \mathbf{t})} Y \sim N(\mu', \Sigma') = N(\mathbf{A}\mu + \mathbf{t}, \mathbf{A}\Sigma\mathbf{A}^T)$. The parameters of the PDFs $N(\mu, \Sigma)$ and $N(\mu', \Sigma')$ can be easily estimated from the input images. The relationship between p and s is [5; 6; 7]

$$s(\mathbf{y}) = \frac{1}{2\pi\sqrt{|\Sigma'|}} \exp\left(-\frac{1}{2}(\mathbf{y} - \mu')^T \Sigma'^{-1}(\mathbf{y} - \mu')\right) = \frac{1}{|\mathbf{A}|} p(\mathbf{x}),$$

where $|\mathbf{A}|$ can be readily derived from $|\mathbf{A}| = \sqrt{|\Sigma'|/|\Sigma|}$, since $\mathbf{A}\Sigma\mathbf{A}^T = \Sigma'$. By making the necessary equivalence conversions, we get the *Mahalanobis distance* [5; 6; 7]. We then define our *covariant functions* $P, S : \mathbb{R}^n \rightarrow \mathbb{R}$ by [5; 6; 7]

$$P(\mathbf{x}) = (\mathbf{x} - \mu)^T \Sigma^{-1}(\mathbf{x} - \mu) \quad \text{and} \quad S(\mathbf{y}) = (\mathbf{y} - \mu')^T \Sigma'^{-1}(\mathbf{y} - \mu').$$

Hence

$$P(\mathbf{x}) = S(\mathbf{A}\mathbf{x} + \mathbf{t}) = S(\mathbf{y}). \quad (12)$$

Multiplying Eq. (11) by Eq. (12), we can integrate out the individual point correspondences [5; 6; 7]

$$\int_{\mathcal{F}_t} \mathbf{x}P(\mathbf{x}) \, d\mathbf{x} = |\mathbf{A}|^{-1} \int_{\mathcal{F}_o} \mathbf{A}^{-1}(\mathbf{y} - \mathbf{t})S(\mathbf{y}) \, d\mathbf{y} ,$$

where we have used the integral transformation $\mathbf{x} = \mathbf{A}^{-1}(\mathbf{y} - \mathbf{t})$, $d\mathbf{x} = |\mathbf{A}|^{-1}d\mathbf{y}$. In order to generate more linearly independent equations, we will adopt suitable nonlinear functions $\omega : \mathbb{R} \rightarrow \mathbb{R}$ and generate new equations according to Eq. (4) [5; 6; 7]:

$$\int_{\mathcal{F}_t} \mathbf{x}\omega(P(\mathbf{x})) \, d\mathbf{x} = |\mathbf{A}|^{-1} \int_{\mathcal{F}_o} \mathbf{A}^{-1}(\mathbf{y} - \mathbf{t})\omega(S(\mathbf{y})) \, d\mathbf{y}. \quad (13)$$

Adopting a set of linearly independent functions $\{\omega_i\}_{i=1}^m$, we can expand the above integrals and get the following linear systems for all $k = 1, \dots, n$ [5; 6; 7]

$$|\mathbf{A}| \int_{\mathcal{F}_t} x_k \omega_j(P(\mathbf{x})) \, d\mathbf{x} = \sum_{i=1}^n q_{ki} \int_{\mathcal{F}_o} y_i \omega_j(S(\mathbf{y})) \, d\mathbf{y} + q_{k(n+1)} \int_{\mathcal{F}_o} \omega_j(S(\mathbf{y})) \, d\mathbf{y} ,$$

where $j = 1, \dots, m$. The solution of this linear system provides the parameters of the transformation. If $m > 3$ then the system is overdetermined and the result is obtained from a least-squares solution.

We could define our *covariant functions* $\mathcal{P}, \mathcal{S} : \mathbb{R}^n \rightarrow \mathbb{R}$ as in [6; 7]

$$\mathcal{P}(\mathbf{x}) = 2\pi\sqrt{|\Sigma|}p(\mathbf{x}) \quad \text{and} \quad \mathcal{S}(\mathbf{y}) = 2\pi\sqrt{|\Sigma'|}s(\mathbf{y}) . \quad (14)$$

Assuming that the *template* object consists of $\ell \geq 2$ disjoint shapes, each component has exactly one corresponding shape on the *observation*. As a consequence, we can construct *covariant functions* $\mathcal{P}_i(\mathbf{x}), \mathcal{S}_i(\mathbf{y})$ for each pair of shapes satisfying Eq. (12). Furthermore, the overall shape gives rise to a pair of covariant functions $\mathcal{P}_0(\mathbf{x})$ and $\mathcal{S}_0(\mathbf{y})$. As such a matching is usually not known, we will sum these relations, yielding *covariant function* [6; 7]

$$P(\mathbf{x}) \equiv \sum_{i=0}^m \mathcal{P}_i(\mathbf{x}) = \sum_{i=0}^m \mathcal{S}_i(\mathbf{y}) \equiv S(\mathbf{y}). \quad (15)$$

Note that these sums are mixtures of unnormalized Gaussian PDFs, which can also be interpreted as a consistent colouring of the *template* and *observation*, respectively (see Fig. 2), which preserves the effect of the transformation. Similar to Eq. (13), we may use the following relation [6; 7]:

$$\sum_{i=1}^{\ell} \int_{\mathcal{F}_t} \mathbf{x}\omega(\mathcal{P}_i(\mathbf{x})) \, d\mathbf{x} = |\mathbf{A}|^{-1} \int_{\mathcal{F}_o} \mathbf{A}^{-1}(\mathbf{y} - \mathbf{t}) \sum_{i=1}^{\ell} \omega(\mathcal{S}_i(\mathbf{y})) \, d\mathbf{y} .$$

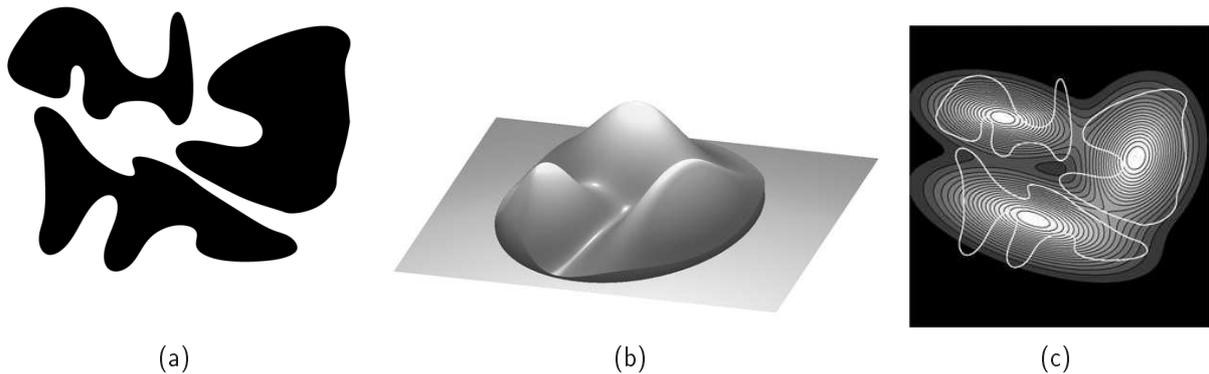


Figure 2: Gaussian PDFs fitted over a compound shape yield a consistent colouring. (a) Original shape. (b) A 3D plot of the Gaussian PDFs over the elliptic domain with $r = 2$. (c) Gaussian densities represented as a greyscale image. The white contour defines component boundaries.

A trivial choice for the domains in our integral equation Eq. (13) is the foreground regions \mathcal{F}_t and \mathcal{F}_o [5]. A clear disadvantage of this approach is that any segmentation error will inherently result in erroneous integrals and cause a misalignment. The key idea is to make use of the statistics of the whole *template* and *observation* objects. Since the equidensity contours of these PDFs are ellipsoids, it is natural to choose a corresponding pair of such ellipses as the integration domain.

We proposed two different approaches to construct our linear system of equations. First, when the object has only one part, the foreground regions of the objects could only be used for the integration domain. In this case, the *covariant functions* are constructed by making use of Mahalanobis distance, defined over the objects, and we have to integrate over the shape domains (Single density). In the second case, when we have compound objects we choose an elliptic domain for integration, and the *covariant functions* are defined by a mixture of Gaussians. Our equations were constructed in the continuum, but in practice we only have a limited precision digital image. Hence, the integrals will be approximated by finite sums over a grid with sufficient resolution (MPDF with finite sums over a grid). However, in the case of compound objects, we also provide an efficient numerical scheme in order to evaluate the integrals in closed forms (MPDF with an efficient numerical scheme).

While other choices of ω_i are also possible, the *power functions* allow a closed form computation of the integrals over the elliptic domains \mathcal{F}_t and \mathcal{F}_o . A clear benefit of this numerical scheme is near-real time performance. We found empirically that the l^{th} power and l^{th} root functions with odd l , i.e. the set $\{x, x^3, x^5, \sqrt[3]{x}, \sqrt[5]{x}\}$, produced satisfactory alignments in all of our test cases.

In order to analyze the performance of our algorithm, we created an image dataset con-

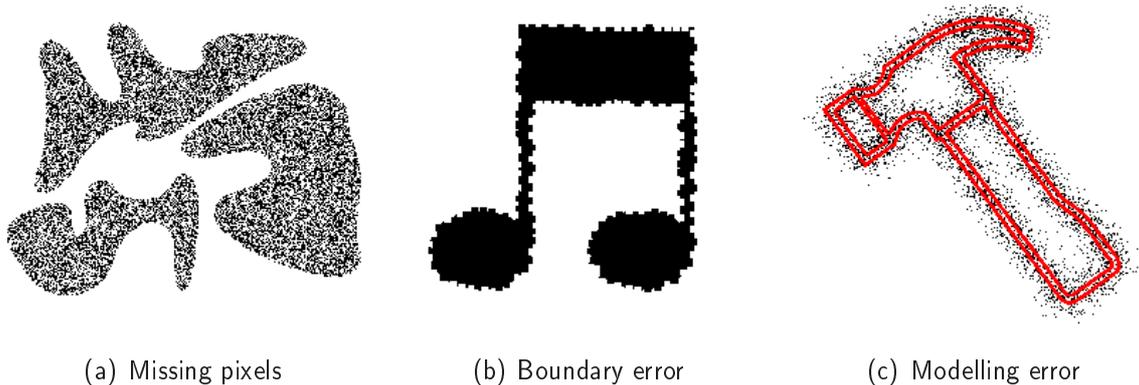


Figure 3: Sample *observations* for testing robustness. In c) the true affine contours are highlighted in red.

taining synthetically generated *observations*, where the applied transformations were randomly composed. The method was tested on synthetic as well as on real images. Then we evaluated the same kinds of error measure as those defined in Eq. (10). The robustness of the proposed algorithm was analyzed against missing pixels, boundary and modelling errors (see Fig. 3). Besides using real images with these kinds of errors, we also conducted a systematic test on simulated data. Table 2 shows the median of error measures obtained by our proposed algorithms.

Method	ϵ (pixel)	δ %	CPU time (sec.)
Single density [5]	0.64	0.31	0.48
MPDF with finite sums over a grid [6]	0.58	0.25	4.65
MPDF with an efficient numerical scheme [7]	0.54	0.19	0.33

Table 2: Median of error measures and runtimes obtained by the proposed methods on 1500 randomly generated *observations*.

With our method point correspondences are not required, nor the solution of a complex optimization problem. It has linear time complexity and yields an exact solution regardless of the size of the deformation. In the case of compound shapes, we also propose a robust and efficient numerical scheme for achieving near real-time performance. However, due to its linear time complexity, the proposed algorithm runs quite fast on large images, so we need not compromise quality when the CPU time is a critical issue. The experimental results demonstrate that our method provides good alignments on both real and synthetic images. Moreover, its robustness was demonstrated. In general, our method will perform well as long as the first and second order statistics of shapes do not change dramatically, hence its superiority can be fully exploited in applications where occlusion can be kept to a minimum.

Realignment of Deformed Object Fragments

Now we address the problem of simultaneously estimating different linear deformations, resulting in a global nonlinear transformation between an original object and its broken fragments without correspondences. Our purpose is to realign $2 \leq \ell \in \mathbb{N}$ shapes of the *observation* to their original position on the *template*. The transformation is nonlinear and is made from ℓ linear transformations, where the i^{th} transformation is denoted by \mathbf{A}_i . This is also known as the *puzzle* problem, which is not only interesting from a theoretical point of view, but also arises in many application domains like archaeology and medical imaging (e.g. bone fracture reduction).

We will assume that the overall segmentation of the *template* is unknown, i.e. its partitioning is hidden. The labelling of shapes on the input images is given by the functions $\mathbb{L}_t, \mathbb{L}_o : \mathbb{P}^n \rightarrow \{0, 1, \dots, \ell\}$, which assign a value 0 to the background. More precisely, \mathbb{L}_t is hidden, since the partitioning of the *template* is unknown. Furthermore $\mathcal{D}_i = \{\mathbf{x} \in \mathbb{P}^n | \mathbb{L}_t(\mathbf{x}) = i\}$ and $\mathcal{D}'_i = \{\mathbf{y} \in \mathbb{P}^n | \mathbb{L}_o(\mathbf{y}) = i\}$ denote the points of the i^{th} *template shape* and its distorted *observation*, respectively. If the shape correspondence were known, a pairwise alignment could be recovered by any standard binary registration method like that described in [4]. Unfortunately, to find this correspondence is far from trivial, hence we are interested in a *direct solution* without identifying corresponding object-pairs.

Let us now consider the i^{th} fragment, where the *identity relation* Eq. (1) [4; 8]:

$$\mathbf{x} = \mathbf{A}_i \mathbf{y} .$$

Notice that the *identity relation* remains valid when an arbitrary nonlinear $\omega : \mathbb{P}^n \rightarrow \mathbb{R}$ function acts on the both sides of (Eq. (3)) [4; 8]:

$$\omega(\mathbf{x}) = \omega(\mathbf{A}_i \mathbf{y}) . \tag{16}$$

Integrating over the domain \mathcal{D}_i , we get [4; 8]

$$\int_{\mathcal{D}_i} \omega(\mathbf{x}) \, d\mathbf{x} = |\mathbf{A}_i| \int_{\mathcal{D}'_i} \omega(\mathbf{A}_i \mathbf{y}) \, d\mathbf{y} , \tag{17}$$

where the integral transformation $\mathbf{x} = \mathbf{A}_i \mathbf{y}$, $d\mathbf{x} = |\mathbf{A}_i| d\mathbf{y}$ has been applied and $|\mathbf{A}_i|$ is the Jacobian determinant of the i^{th} transformation. The nonlinear function ω acts directly on the point coordinates and hence on the unknown parameters of \mathbf{A}_i resulting in a nonlinear system of equations [4; 8]. Based on Eq. (17), we can construct as many equations as needed by making use of a set of nonlinear functions.

Hence we know relations for Eq. (17) between the i^{th} shape-pair, but neither the partitioning (*i.e.* the hidden labeling \mathbb{L}_t) of the *template* nor correspondences between the shapes is known. A standard technique is to sum all the equations for all shape domains \mathcal{D}_i and solve the problem simultaneously, estimating all the parameters in one system of equations. By making use of a set of $\{\omega_j\}_{j=1}^m$ functions in Eq. (17), we get [8]:

$$\sum_{i=1}^{\ell} \int_{\mathcal{D}_i} \omega_j(\mathbf{x}) \, d\mathbf{x} = \sum_{i=1}^{\ell} |\mathbf{A}_i| \int_{\mathcal{D}'_i} \omega_j(\mathbf{A}_i \mathbf{y}) \, d\mathbf{y} .$$

Let $\mathcal{F}_t := \cup_{i=1}^{\ell} \mathcal{D}_i$, where $\mathcal{F}_t = \{\mathbf{x} \in \mathbb{P}^n | \mathbb{L}_t(\mathbf{x}) \neq 0\}$ is the shape domain corresponding to the whole *template*. Therefore the left hand side of the above equation can be written as [8]

$$\sum_{i=1}^{\ell} \int_{\mathcal{D}_i} \omega_j(\mathbf{x}) \, d\mathbf{x} = \int_{\cup_{i=1}^{\ell} \mathcal{D}_i} \omega_j(\mathbf{x}) \, d\mathbf{x} = \int_{\mathcal{F}_t} \omega_j(\mathbf{x}) \, d\mathbf{x} ,$$

which can be computed directly from the input image without knowing the partitioning \mathcal{D}_i . The resulting system of equations has $\ell n(n+1)$ unknowns [8]:

$$\int_{\mathcal{F}_t} \omega_j(\mathbf{x}) \, d\mathbf{x} = \sum_{i=1}^{\ell} |\mathbf{A}_i| \int_{\mathcal{D}'_i} \omega_j(\mathbf{A}_i \mathbf{y}) \, d\mathbf{y} \quad j = 1, \dots, m . \quad (18)$$

The solution of this system of equations provides all the unknown parameters of the overall deformation. Since each ω_j provides one equation, we need $m \geq \ell n(n+1)$ linearly independent functions to solve ℓ linear transformations. In practice, $m > \ell n(n+1)$, yielding an overdetermined system where a least-squares solution is obtained.

Theoretically, any function satisfying Eq. (16) could be used to construct the system of equations defined in Eq. (18). The solution is obtained via iterative least-squares minimization algorithms like the *Levenberg-Marquardt algorithm* and requires a carefully chosen numerical scheme. The solver needs to evaluate the equations for each iteration step, hence our aim is to apply the kind of ω s that produce a nonlinear system of equations instead of a system of integral equations (*i.e.* unknowns are emphasized in the integrals). It was shown in [4] that choosing a set of polynomial functions will result in a polynomial system of equations, where these integrals become precomputed constants. Based on these findings, the following set of polynomials will be adopted [8]

$$\{\omega_j\}_{j=1}^m = \{\mathbf{x} \mapsto x_1^{u_1} \dots x_n^{u_n} | u_1, \dots, u_n \in \mathbb{N}_0, \sum_{i=1}^n u_i \leq d\} , \quad (19)$$

where $\omega_j : \mathbb{P}^n \rightarrow \mathbb{R}$, d is the maximum degree, and the number of polynomials is given by

$$m = \frac{1}{n!} \prod_{i=1}^n (d + i).$$

The coordinates of both images are *normalized* to the unit hyper-cube $[-0.5, 0.5]^n$, in order to avoid numerical stability due to the high powers. Let \mathbf{N}_t and \mathbf{N}_o denote the normalizing transformations of the *template* and *observation*, respectively. Since a least-squares solution involves minimizing the algebraic error of Eq. (18), we expect an equal contribution from each equation in order to guarantee an unbiased error measure. In practice, only a limited precision digital image is available, thus the integrals can only be *approximated* by a discrete sum over the foreground pixels that introduce an inherent, although negligible error into our computations [8]

$$\frac{1}{c_j} \sum_{\mathbf{x} \in F_t} \omega_j(\mathbf{N}_t \mathbf{x}) = \frac{1}{c_j} \sum_{i=1}^{\ell} |\mathbf{A}_i| \sum_{\mathbf{y} \in D'_i} \omega_j(\mathbf{N}_o \mathbf{A}_i \mathbf{y}), \quad (20)$$

where c_j is an appropriate constant, corresponding to the integral $\int |\omega_j(\mathbf{x})| d\mathbf{x}$ over a hypersphere with centre at the origin and a radius $\sqrt{n}/2$.

We apply the proposed framework to well-known classes of linear deformations: 2D and 3D affine and rigid-body transformations. 2D affine transformations are often used as a linear approximation of projective distortions. When an object is broken into several parts, the fragments are generally distorted by different rigid-body transformations. 3D rigid-body transformations are important in many medical applications.

We quantitatively evaluated the proposed algorithm on a large synthetic dataset containing 2D, 3D images, where linear (affine and rigid-body) transformations are considered. We then conducted an extended analysis of numerical stability of the proposed algorithm. The results show that the method is robust against segmentation errors. We present experimental results on 2D real images as well as on volumetric medical images applied to surgical planning. In practice, segmentation never produces perfect shapes. Therefore, besides using various kinds of real images inherently subject to such errors, we also evaluated the robustness of the proposed approach against different types of segmentation errors.

A novel framework to solve shape realignment problem was proposed and applied to 2D and 3D affine and rigid-body transformations (Fig. 4). In contrast to classical solutions based on landmark extraction and correspondences, the proposed solution finds the aligning transformations without any additional information. Essentially, the method consists of constructing a polynomial system of equations whose solution directly provides the unknown parameters. The quantitative evaluations on both 2D and 3D synthetic datasets demonstrate the performance and robustness of the method and the results obtained on real images suggest that it can be applied to various application domains. The main advantages are that the proposed method

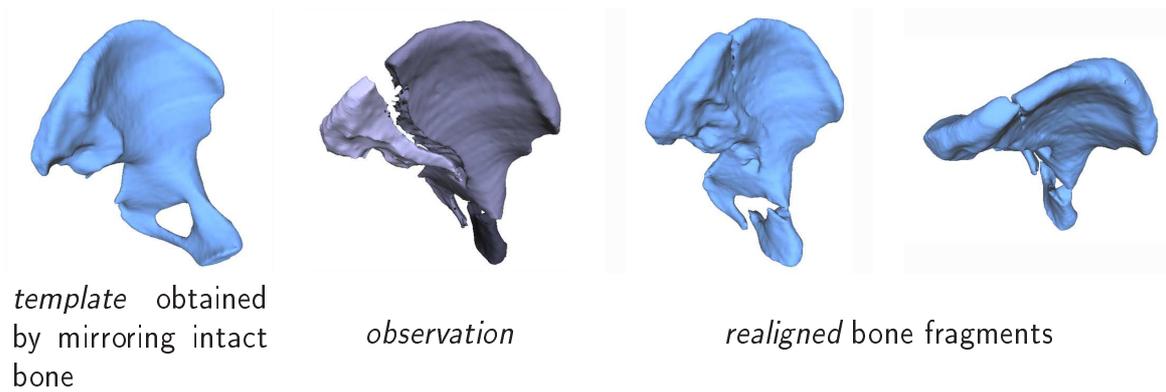


Figure 4: Bone fracture reduction.

does not require point correspondences, it is quite fast, and it is easy to implement.

Summary of the Author's Contributions

In the following a listing of the most important results of the dissertation is given. The work presented in this dissertation resulted in several publications. Table 3 summarizes which publication covers which item of the thesis points.

- I.) The author addresses the problem of the estimation of affine transformations for aligning a known 2D shape and its distorted observation. He proposes a novel approach where the exact transformation is obtained as the solution of a polynomial system of equations. He tested the proposed method on synthetic as well as on real images. The author demonstrated the robustness of the algorithm in the presence of segmentation errors and additive geometric noise too, then he successfully applied the method to the registration of hip prosthesis X-ray images.
- II.) The author proposes a novel approach for the estimation of 2D affine transformations for aligning a planar shape and its distorted observation. He shows how the exact transformation is obtained as a least-squares solution of a linear system of equations constructed by fitting Gaussian densities to the shapes which preserve the effect of the unknown transformation. In the case of compound shapes, the author also presents a robust and efficient numerical scheme that achieves a near real-time performance. He tested the method on synthetic as well as on real images and demonstrated its robustness in the case of missing pixels, boundary errors, and modelling errors.
- III.) The author considers the problem of realigning broken objects without correspondences, where the segmentation of the overall *template* is known, but the segmentation of the object parts is unknown. The author applies linear transformations between the object

fragments and presents the method by using 2D and 3D affine transformations. He shows that constructing and solving a polynomial system of equations provides the unknown parameters of the alignment. Here he quantitatively evaluated the proposed algorithm on a large synthetic dataset containing 2D and 3D images. He analyzes the numerical stability of the method as well as its robustness against segmentation errors. He also presents the results of experiments on 2D real images as well as on volumetric medical images applied to surgical planning.

	[3]	[4]	[5]	[6]	[7]	[8]
I.	•	•				
II.			•	•	•	
III.						•

Table 3: The connection between the thesis points and the corresponding publications

References

- [1] Rami Hagege and Joseph M. Francos. Parametric estimation of two dimensional affine transformations. In *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, volume 3, pages 305–308, Montreal, May 2004. IEEE.
- [2] Rami Hagege and Joseph M. Francos. Linear estimation of sequences of multi-dimensional affine transformations. In *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, volume 2, pages 785–788, Toulouse, France, May 2006. IEEE.
- [3] Csaba Domokos, Zoltan Kato, and Joseph M. Francos. Parametric estimation of affine deformations of binary images. In *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, pages 889–892, Las Vegas, NV, USA, April 2008. IEEE.
- [4] Csaba Domokos and Zoltan Kato. Parametric estimation of affine deformations of planar shapes. *Pattern Recognition*, 43(3):569–578, March 2010.
- [5] Csaba Domokos and Zoltan Kato. Binary image registration using covariant Gaussian densities. In Aurélio Campilho and Mohamed Kamel, editors, *Proceedings of International Conference on Image Analysis and Recognition*, volume 5112 of *Lecture Notes in Computer Science*, pages 455–464, Póvoa de Varzim, Portugal, June 2008. Springer.
- [6] Csaba Domokos and Zoltan Kato. Affine alignment of compound objects: A direct approach. In *Proceedings of International Conference on Image Processing*, pages 169–172, Cairo, Egypt, November 2009. IEEE.

- [7] Csaba Domokos and Zoltan Kato. Affine shape alignment using covariant Gaussian densities: A direct solution. *IEEE Transactions on Image Processing*, 2010. Submitted.
- [8] Csaba Domokos and Zoltan Kato. Affine puzzle: Realigning deformed object fragments without correspondences. In Kostas Daniilidis, Petros Maragos, and Nikos Paragios, editors, *Proceedings of European Conference on Computer Vision*, volume 6312 of *Lecture Notes in Computer Science*, pages 777–790, Heraklion, Crete, Greece, September 2010. Springer.