J Math Imaging Vis
DOI 10.1007/s10851-009-0188-4

Parametric Estimation of Affine Transformations: An Exact Linear Solution

Rami Hagege · Joseph M. Francos

Abstract We consider the problem of estimating the geometric deformation of an object, with respect to some reference observation on it. Existing solutions, set in the standard coordinate system imposed by the measurement system, lead to high-dimensional, non-convex optimization problems. We propose a novel framework that employs a set of non-linear functionals to replace this originally high dimensional problem by an *equivalent problem* that is *linear* in the unknown transformation parameters. The proposed solution includes the case where the deformation relating the observed signature of the object and the reference template is composed both of the geometric deformation due to the affine transformation of the coordinate system and a constant amplitude gain. The proposed solution is *unique and exact* and is applicable to any affine transformation *regardless of its magnitude*.

Keywords Affine transformations · Registration · Parameter estimation · Deformable templates · Linear least squares

1 Introduction

Registration is the procedure of bringing two or more observations on the same object (usually, in the form of images) to a common coordinate system. These images are usually referred to as the template (or reference) image, and the observed image. The difficulty of the registration problem results from its most basic characteristic: although the template (or a set of templates) is known, the variability associated with the object, such as its location and pose in the observed scene, the illumination conditions, or its deformation are unknown *a-priori*, and only the family of transformations causing this variability in the observation can be defined, based for example on the physical characteristics of the problem. This huge variability in the object signature (for any single object) due to the tremendous set of possible deformations that may relate the template and the observed signature, makes any detection and recognition problem ill-defined unless this variability is taken into account.

There is a vast literature on registration in general and on image registration in particular (see, *e.g.*, [2–4] for comprehensive surveys). Thus in the following we shall not attempt to review the entire range of available techniques, but rather to classify them such that the method proposed in this paper is placed in the proper framework, and its uniqueness is made clear.

In the following we shall attempt to classify the existing methods based on two different categorization criteria, that together enable characterization of the available methods. These classifications are based on the type of features used (salient local features vs. global reference to the object as a whole); and the approach of estimating the deformation as an explicit or implicit problem. We begin with a brief definition of the problem to be addressed: Assume we are observing functions defined on compact supports. Given a pair of functions, $(h(\mathbf{x}), g(\mathbf{x}))$, $\mathbf{x} \in R^n$ ($n = 2$ for planar objects), representing two observations on the same object, where the two are related by a geometric transformation such that $h(\mathbf{x}) = g(\varphi(\mathbf{x}))$, $\varphi \in G$, and G is the affine group of transformations, the goal is to find φ.

R. Hagege · J.M. Francos (✉)
Electrical and Computer Engineering Department,
Ben-Gurion University, Beer-Sheva 84105, Israel
e-mail: francos@ee.bgu.ac.il

R. Hagege
e-mail: hagege@ee.bgu.ac.il

Published online: 13 January 2010

Local Features vs. Global Reference Intuitively, it seems that the simplest way to relate any two objects, where one is assumed to be a deformed version of the other, is by associating ordered sequences of some salient features, that can be located on both objects (see, *e.g.*, [3, 15]). These are usually selected to be line features or point features (thus much of the information contained in the intensity structure of the images is being ignored). In this methodology, the shape of an object is described by the configuration of the landmarks, projected onto the image plane. The key concept, common to all such methods is the requirement that the features are labeled such that we know the correspondence between feature points across different images. Therefore, to make this approach feasible, the correspondence problem must be solved first. This approach may be feasible when the features are easily detectable, the deformation is close enough to the identity, the number of features is relatively small (for combinatoric reasons, yet big enough to allow for a meaningful estimation), and there is a strong contextual evidence to guide the solution to the correspondence problem. Yet, in many problems the feature points are not easily identifiable, and their number may be large—in which case the correspondence problem rapidly becomes very difficult to solve. Recently, methods based on identifying local invariants such as SIFT [18], MSER [19] (see also [20] and the references therein) have made the solution to this crucial problem more feasible and tractable.

A family of methods that attempt to overcome the need for a perfect solution to the correspondence problem, is the moment-based invariants approach. In these methods low order moments of the set of labeled points, or of closed boundary regions, in both images are computed by integrating over all combinatorially possible polygons created by the set of labeled points in order to overcome the need for an exact solution to the correspondence problem, *e.g.*, [11]. In a second stage the affine transformation parameters can be computed by solving a linear system of equations based on the estimated moments. A related approach is based on computing the dominant eigenvalues and the corresponding eigenvectors of the labeled points scattering matrix for each of the objects to be registered, from which the deformation parameters are evaluated in a second step.

The alternative, global approach, calls for a direct utilization of the observed intensity functions in both the reference and the observed image, without identifying any features in any of the images. Among the methods in this class of registration algorithms are those based on the maximization of the mutual information, *e.g.*, [14]. These methods obtain the transformation parameters by employing iterative optimization procedures, and hence are applicable only when the deformation is small, which lowers the risk of obtaining a local minimum as the solution.

Explicit (or Global) vs. Implicit (or Local, Optimization-based) Solutions An implicit method is one that finds some map ψ such that, ideally, $\psi(h, g, \varphi) = 0$. All registration methods based on minimizing some metric, are therefore implicit. On the other hand, in an explicit solution one obtains a map H (or an operator) such that the unknown deformation can be expressed by $\varphi = H(h, g)$.

The common principle in the implementation of all the implicit methods is the definition of a cost function penalizing both the 'distance' between a deformed version of the template and the observation, and a measure of the 'size' of the deformation. The aim is then to find the deformation that minimizes the cost. More specifically, let $d(h, g)$ be some metric on the function space that contains h and g. A solution to the deformation estimation problem is given by $\hat{\varphi} = \arg\min_{\varphi \in G}(d(h(\mathbf{x}), g(\varphi(\mathbf{x}))) + D(I, \varphi))$, where $D(I, \varphi)$ is a regularization term specifying some *a-priori* knowledge about the distance of φ from the identity [1, 16]. In principle, in order to find the global minimum of $d(h(\mathbf{x}), g(\varphi(\mathbf{x}))) + D(I, \varphi)$ one has to check each and every element of G, which is usually impossible. Nevertheless, application of some optimization procedure allows for finding a local minimum of this type of cost function, in the affine case [17], as well as for more complex deformations [4]. Unfortunately, there is no systematic way to obtain the global minimum. Thus, in general, implicit methods fail to find the global solution and hence can be considered to be local methods, as the solution holds only in a small neighborhood of the correct φ.

Obviously, an explicit solution is preferable due to many reasons. These include, computational complexity as optimization is avoided, and more importantly uniqueness of the solution. Moreover the explicit solution is always global in nature, since no local minimization operations are involved. Many such global methods exist (see, *e.g.*, [17] and the references therein) however their scope is restricted to a relatively small family of transformations. Thus, there are explicit methods in the cases of translation only, rotation only, or scale (moderate factor) only, but they turn into combined explicit/implicit methods for the combined transformation of rotation, scaling and translation, [6, 7]. Translation estimation is conveniently carried out in the Fourier domain based on the phase shift of the Fourier transforms of the two images to be registered, by employing the normalized phase-correlation algorithm, *e.g.*, [5, 8]. The combined method for estimating translation, scale and rotation, first transforms the pair of object images into the Fourier plane. In this domain the relation between the absolute values of the Fourier transforms is a function of the scale and rotation angle, but is independent of translations—as translations are related only to the phase. Estimation of rotation and scale is implemented iteratively in the polar Fourier domain, [6, 7]. Nevertheless, to the best of our knowledge *no explicit* method for estimating an arbitrary affine deformation was known prior to the presentation of preliminary versions of the method proposed in this paper [10, 24].

For the case where the deformation is affine, while assuming the deformation is small and the observations differentiable, a widely used approach is to linearize the problem using a first order Taylor series expansion. See, *e.g.*. [21]. The major advantage of this approach is that in case the deformation is indeed small, a solution to the problem of estimating the affine transformation parameters is formulated as a solution to a system of linear equations. We shall further elaborate on this method in Sect. 5 where we compare it to the method proposed in this paper. The method proposed here also provides a solution to the problem of finding the affine transformation parameters by solving a system of linear equations. However, this new method, that originates from entirely different considerations, is *exact* and not approximate, while being applicable to deformations of *any* size.

In the method described in this paper our aim is to find an explicit global operator $H(h, g)$ such that for every pair (h, g) for which $h(\mathbf{x}) = g(\varphi(\mathbf{x}))$, $\varphi \in G$, where G is the affine group, we have $\varphi = H(h, g)$. The center of the solution proposed in this paper is a method to replace the high dimensional and computationally intensive problem of evaluating the orbit created by applying to a given template the whole set of transformations in the affine group, by an *equivalent problem* which is *linear* in the unknown parameters of the affine transformation. In this setting, the problem of finding the parametric model of the affine deformation is mapped, by a set on non-linear functionals, into a set of *linear* equations which is then solved for *the* affine transformation parameters.

The basic solution is further extended to include the case where the deformation relating the observed signature of the object and the template, is composed both of a geometric deformation due to the affine transformation of the coordinate system and a constant amplitude gain. The proposed solution is *unique and exact*, as it provides a closed form expression for evaluating each of the affine transformation parameters using *only* measurements of the intensity information of the observed and reference signals (or images). The solution is applicable to any affine transformation *regardless of its magnitude*. Moreover, in forthcoming papers we show that the methodology presented here can be extended to handle groups of deformations that are much larger and richer than the affine group. These include elastic deformations modeled by the group of homeomorphisms (see, [9, 24, 26]), and time varying deformations, [25].

The structure of the paper is as follows: We begin by rigorously defining the scope of the problem of finding the affine transformation parameters, given an observation h and a template g of a planar object, where the two are known to be related through an affine transformation. To simplify the notations we assume in Sect. 3 that the translation is null and consider the problem of finding $\mathbf{A} \in GL_n(R)$, given the observations on h and g. Then, in Sect. 4 we solve using a least squares approach the problem of estimating the parameters of the affine transformation model, for the case where the model is only an approximation of the true physical distortion. In Sect. 5 the algorithm for finding the affine transformation parameters is extended so that $\mathbf{A}$, and the translation vector are jointly estimated. In Sect. 6 we further extend the framework of the model and consider the case where the observation and template are related by a geometric affine transformation as well as an amplitude deformation in the form of an unknown gain that have to be estimated jointly with the parameters of the affine transformation. In Sect. 8 we present some numerical examples to illustrate the operation and robust performance of the proposed parameter estimation algorithm to image based object registration, and for a broad range of deformations. Finally, we provide our conclusions in Sect. 9.

2 Estimation of Affine Transformations: Problem Definition

In this section we shall briefly set the mathematical framework we adopt in order to formalize the analysis of the deformation estimation problem. This framework enables accurate representation and analysis of our problem, leading to rigorous criteria on the existence and uniqueness of the solution, and under some mild restrictions to be explained below to the derivation of an explicit solution.

We note that due to the inherent physical properties of the problem, it is natural to model and solve it in the continuous domain. Inherently, the mapping φ of R^n into itself is of a continuous nature, as is the physical phenomenon of geometric deformation of real-life objects it represents. Thus, if we impose a discrete model (*e.g.*, $\mathbf{x} \in Z^n$), we find that, in general, the natural φ to consider is incompatible (as for "almost all" $\mathbf{x} \in Z^n$, $\varphi(\mathbf{x}) \notin Z^n$). Thus, the problem and its solution are formulated in the continuous domain, while the sampling and quantization effects that accompany the digital implementation of the method, are handled as noise contributions.

2.1 Group Theory Setting

Let M denote the space of compact support, bounded, and Lebesgue measurable (or more simply, integrable) functions from R^n to R, which in our case is the set of objects (real valued images, where $n = 2$) on which we observe. Let $\mathbf{x}$ be some vector in R^n.

Let G be a group representing the set of deformations the objects may undergo—the affine group in the case studied in this paper. G is said to act as a transformation group on M if there is a mapping $G \times M \to M$, denoted by $(\phi, m) \mapsto m \circ \phi = m(\phi(\mathbf{x}))$ such that $(m \circ \phi_1) \circ \phi_2 = m \circ (\phi_1 \circ \phi_2)$ for every $\phi_1, \phi_2 \in G$ and $m \in M$; and if $m \circ e = m$ for all $m \in M$, where e is the identity element of G.

For a given $m \in M$, the set $\{m \circ \phi : \phi \in G\}$ is called the orbit of m. It is the entire set of possible observations on the object—the result of applying to it any of the deformations in the group.

The stabilizer of the function $m \in M$ with respect to the group G is the set of group elements $\phi \in G$ such that $m \circ \phi = m$, *i.e.*, the set of group elements that map m to itself.

Thus the group G naturally defines an equivalence relation on M in terms of the orbits of M induced by the action of G: Any two functions h and g are equivalent if they are on the same orbit, *i.e.*, if there exists some $\phi \in G$ such that $g \circ \phi = h$.

Let $M_{Aff} \subseteq M$ be the subset of functions in M with no affine symmetry, *i.e.*, the set of functions in M whose stabilizer is trivial and includes only e, the identity element of G. Thus, M_{Aff} is the subset of functions in M where *uniqueness of the solution* to the defined problem is guaranteed in the sense that if $h, g \in M_{Aff}$ such that they are on the same orbit, then there exists a *single* ϕ such that $g \circ \phi = h$.

In contrast, examples of functions with affine symmetry include any constant function defined on all of R^n; any periodic function defined on all of R^n; and in the two dimensional case, functions with radial symmetry, such as a circle (as $SO_2(R) \subset GL_2(R)$). Note however that functions with compact support are not translation nor scale invariant.

2.2 Problem Statement

To simplify notations, we first assume that translation is null and consider the most elementary problem: Let $GL_n(R)$ denote the group of real valued invertible $n \times n$ linear transformations, and let $\mathbf{A}$ be some matrix in $GL_n(R)$. Then, given two bounded, Lebesgue measurable functions h, g with compact supports, and with no affine symmetry, such that

$$h : R^n \to R$$
$$g : R^n \to R$$

where

$$h(\mathbf{x}) = g(\mathbf{A}\mathbf{x}), \quad \mathbf{A} \in GL_n(R),\ \mathbf{x} \in R^n \tag{1}$$

the problem is to find the matrix $\mathbf{A}$. In the special case where h and g are images, we have $n = 2$, and $\mathbf{A}$ is a real valued invertible 2×2 matrix.

The direct approach for solving the problem of finding the parameters of the unknown transformation $\mathbf{A} \in GL_n(R)$ is to apply the set of all possible transformations (*i.e.*, every element of $GL_n(R)$), to the given template g, thus evaluating the entire orbit of g. Since h and g are affine related, one of the points on the orbit represents the action of the desired group element $\mathbf{A}$. Nevertheless, since $\mathbf{A}$ is an $n \times n$ matrix it is clear that implementation of such a search on the orbit requires a search over an n^2-dimensional manifold embedded in an infinite dimensional function space, which is infeasible.

In this paper we show that the problem of finding the parameters of the unknown affine transformation, whose direct solution requires a highly complex search in a function space, can be formulated as an *explicit parameter estimation problem*. Moreover, it is shown that the original problem can be formulated in terms of an *equivalent* problem which is expressed in the form of a *linear* system of equations in the unknown parameters of the affine transformation. A solution of this linear system of equations provides *the* unknown transformation parameters. To increase noise immunity, it employs integral operators, rather than differential ones.

3 An Algorithmic Solution

In this section we specify the conditions, and provide a constructive proof showing that given an observation on $h(\mathbf{x}) \in M_{Aff}$ and an observation on $g(\mathbf{x}) \in M_{Aff}$ where $h(\mathbf{x}) = g(\mathbf{A}\mathbf{x})$, $\mathbf{A}$ can be *uniquely* determined.

Let $\mathbf{x}, \mathbf{y} \in R^n$, *i.e.*,

$$\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$$
$$\mathbf{y} = [y_1, y_2, \ldots, y_n]^T$$

such that

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \quad \mathbf{x} = \mathbf{A}^{-1}\mathbf{y} \tag{2}$$

Since $\mathbf{A} \in GL_n(R)$, also $\mathbf{A}^{-1} \in GL_n(R)$. It is therefore possible to solve for $\mathbf{A}^{-1}$ and the solution for $\mathbf{A}$ is guaranteed to be in $GL_n(R)$. Moreover, as shown below, in the proposed procedure the transformation determinant is evaluated first, and by a different procedure than the one employed to estimate the elements of $\mathbf{A}^{-1}$. Hence, a non-zero Jacobian guarantees the existence of an inverse to the transformation matrix.

Let $f \in M_{Aff}$ and let $\int_{R^n} f$ denote the Lebesgue integral of f with respect to the Lebesgue measure on R^n. Note that in the following derivation it is assumed that the functions are bounded and have compact support, as they are measurable (and hence integrable) but not necessarily continuous. It is further assumed that $\mathbf{A} \in GL_n(R)$ has a positive determinant.

Let us define an auxiliary function space, W, such that every function $w : R \to R$ in W is Lebesgue measurable, and vanishes at zero. Next, we define the mapping from the space of compact support, bounded, measurable functions to itself induced by the functions in W. More specifically, we define a mapping from M_{Aff} to itself, such that a function $g \in M_{Aff}$ is mapped by $w \in W$ to some function $w(g(x)) \in M_{Aff}$. This operator is, in general, non linear.

The first step in the solution is to find the determinant of the matrix $\mathbf{A}$. Let $w : R \to R$ be some Lebesgue measurable function, such that $w(0) = 0$, *i.e.*, $w \in W$. Then for every such function we have by interchanging the variables $\mathbf{y} = \mathbf{Ax}$ that

$$\int_{R^n} w(h(\mathbf{x})) = \int_{R^n} w(g(\mathbf{Ax})) = \left|\mathbf{A}^{-1}\right| \int_{R^n} w(g(\mathbf{y})) \tag{3}$$

Hence,

$$\left|\mathbf{A}^{-1}\right| = \frac{\int_{R^n} w(h(\mathbf{x}))}{\int_{R^n} w(g(\mathbf{y}))} \tag{4}$$

and $|\mathbf{A}^{-1}| = |\mathbf{A}|^{-1}$. Thus, by an arbitrary choice of the function w the Jacobian $|\mathbf{A}^{-1}|$ is evaluated, as long as $\int_{R^n} w(g(\mathbf{y})) \neq 0$. For example, for any non-zero, measurable g, whose support is not of measure zero, the choice $w(x) = x^2$ satisfies the requirement.

In the second stage we obtain $\mathbf{A}^{-1}$ itself: For every Lebesgue measurable function $w : R \to R$, such that $w(0) = 0$, we have by interchanging the variables that

$$\begin{aligned} \int_{R^n} \mathbf{x} w(h(\mathbf{x})) &= \int_{R^n} \mathbf{x} w(g(\mathbf{Ax})) \\ &= \left|\mathbf{A}^{-1}\right| \int_{R^n} (\mathbf{A}^{-1}\mathbf{y}) w(g(\mathbf{y})) \\ &= \left|\mathbf{A}^{-1}\right| \mathbf{A}^{-1} \int_{R^n} \mathbf{y} w(g(\mathbf{y})) \end{aligned} \tag{5}$$

Repeating this procedure by applying a family of Lebesgue measurable, left-hand compositions $\{w_\ell\}_{\ell=1}^{p} \in W$, to the known relation $h(\mathbf{x}) = g(\mathbf{Ax})$, and rewriting it in a matrix form by ordering the columns on the right (left) of (5) next to each other, yields

$$\begin{aligned} &|\mathbf{A}| \left[\int_{R^n} \mathbf{x} w_1(h(\mathbf{x})) \cdots \int_{R^n} \mathbf{x} w_p(h(\mathbf{x})) \right] \\ &\quad = \mathbf{A}^{-1} \left[\int_{R^n} \mathbf{y} w_1(g(\mathbf{y})) \cdots \int_{R^n} \mathbf{y} w_p(g(\mathbf{y})) \right] \end{aligned} \tag{6}$$

Thus let

$$\mathbf{G}_p = \left[\int_{R^n} \mathbf{y} w_1(g(\mathbf{y})) \cdots \int_{R^n} \mathbf{y} w_p(g(\mathbf{y})) \right] \tag{7}$$

and let

$$\mathbf{H}_p = \left[\int_{R^n} \mathbf{x} w_1(h(\mathbf{x})) \cdots \int_{R^n} \mathbf{x} w_p(h(\mathbf{x})) \right] \tag{8}$$

Rewriting (6) we have

$$|\mathbf{A}|\, \mathbf{H}_p = \mathbf{A}^{-1} \mathbf{G}_p \tag{9}$$

We have just proved the following theorem:

Theorem 1 *Let $\mathbf{A} \in GL_n(R)$. Assume $h, g \in M_{Aff}$ such that $h(\mathbf{x}) = g(\mathbf{Ax})$. Then, given measurements of h and g, $\mathbf{A}$ can be uniquely determined if there exists a function $w \in W$ such that $\int_{R^n} w(g(\mathbf{y})) \neq 0$ and a set of functions $\{w_\ell\}_{\ell=1}^{n} \in W$, such that the matrix $\mathbf{G}_n$ is full rank. Then, $\mathbf{A}^{-1} = |\mathbf{A}|\mathbf{H}_n\mathbf{G}_n^{-1}$.*

Remark 1 Note that the denominator of (4), as well as the elements of the matrix $\mathbf{G}_n$ depend only on the template and its coordinate system and thus have to be evaluated only *once*. In fact the denominator of (4) together with the matrix $\mathbf{G}_n$ represent all the information in the template, required for finding the affine transformation parameters. Thus the denominator of (4) together with $\mathbf{G}_n$ form a "*sufficient representation*" of the template (similarly to the notion of sufficient statistics), so that the template itself is not needed for solving the estimation problem once $\mathbf{G}_n$ and the denominator of (4) have been evaluated.

Remark 2 It should be noted that although we use the term "estimation" throughout this paper, the solution in Theorem 1 for the affine transformation parameters is *exact* and is not an estimate in the usual sense of the word, but rather a procedure for solving for the transformation.

Remark 3 Inspecting equations (4) and (5) (or more generally (9)), we see that application of each of the non linear functionals to the known relation $h(\mathbf{x}) = g(\mathbf{Ax})$ amounts to obtaining a linear constraint on the linear transformation $\mathbf{A}$ between the centers of mass of the functions $w_\ell(h(\mathbf{x}))$ and $w_\ell(g(\mathbf{y}))$ obtained by an identical nonlinear operation on the amplitudes of h and g. Hence, it can be concluded that the proposed solution is explicit and global solution that employs stable "features" (centers of mass) rigorously extracted such that the correspondence between them is explicitly known, rather than (in general) unstable features extracted based on local properties where the difficult correspondence problem has to be solved in a later stage.

Remark 4 The above basic procedure is not directly applicable to binary images. This is because in this case all the w_ℓ's are equivalent and produce linearly dependent columns in (7). Thus, the result is a rank-1 system, which cannot be solved for the deformation parameters.

Remark 5 Note that the solution for $\mathbf{A}$ employs only zero (the Jacobian) and first order constraints (obtained by multiplying $w_\ell(h(\mathbf{x}))$ by $\mathbf{x}$) and avoids the use of higher order moments. However, imposing such a restriction (which is clearly convenient due to the simplicity of the resulting equations) may result in cases where a system of the type (9) does not exist, as shown for example, in the previous remark. It is then obvious that higher order moments are needed to obtain a system similar to (9) (yet nonlinear) with enough equations to solve for all the unknowns (for example a system of polynomial equations).

Remark 6 The application of a set $\{w_\ell\}_{\ell=1}^n$ to $g(\mathbf{y})$ yielding $\mathbf{G}_n$ is in fact a mapping from the space of compact support, bounded, and measurable functions to the space of $n \times n$ matrices. In Theorem 2, we will show that the subset of functions $g \in M_{Aff}$, for which there exists a set $\{w_\ell\}_{\ell=1}^n$ such that $\mathbf{G}_n$ is full-rank, is dense in M_{Aff} in the supremum norm. Hence, for every g, or for an infinitesimal modification of it, the matrix $\mathbf{G}_n$ is invertible.

In order to simplify the proof and to avoid a lengthy technical derivation, we restrict our attention to the (practical) case where $g(\mathbf{y})$ is piecewise continuous. Thus, the range of $g(\mathbf{y})$ in each of its continuity subsets, is an interval in R. Referring to Remarks 4 and 5 above, we note that in the special case where the range of $g(\mathbf{y})$ contains less than n distinct values, $g(\mathbf{y})$ can be replaced by a modified version $g^1(\mathbf{y}) = g(\mathbf{y}) + q(\mathbf{y})$, such that $q(\mathbf{y})$ is continuous and $\|q\|_\infty$ is arbitrarily small, as detailed below.

Lemma 1 *Let S be some subset of R^n that contains an open set, and let $q(\mathbf{y}) : S \to R$. Then $\int_{R^n} \mathbf{y} q(\mathbf{y})$ is a vector in R^n. Moreover, let ϵ be some arbitrarily small real number. Then, for any arbitrarily small ball of functions q such that $\|q\|_\infty < \epsilon$, the range of the mapping $\int_{R^n} \mathbf{y} q(\mathbf{y})$ contains an open ball in R^n.*

Proof The orthogonality of the Cartesian coordinate system, implies the linear independence of its axes, as functions. Hence the mapping from q to $\int_{R^n} \mathbf{y} q(\mathbf{y}) \in R^n$ is a nonsingular linear projection to R^n. Hence the image of an open ball of functions q is an open ball of vectors in R^n. □

Next, we consider the range of g: Let $\{R_k\}$ be a partition into disjoint sets of the range of g, such that no discontinuity point of g has its image in any interval R_k (*i.e.*, on the pre-image of R_k, g is continuous). Let $R_k = (a_k, b_k)$, such that $b_k - a_k > 2\epsilon$. Thus each interval R_k contains an open subset.

Lemma 2 *Let $g \in M_{Aff}$. Then there exists a set of measurable functions defined on the range of g, $\{w_k\}_{k=1}^n \in W$ with disjoint supports such that for every k the pre-image of $w_k(g(\mathbf{y}))$ contains an open set.*

Proof Let $w_k(x) = x\mathbf{1}_{R_k}$, where $\mathbf{1}_{R_k}$ denotes the indicator function of the interval R_k. Since for every k, R_k contains an open set, and since on the pre-image of R_k under g, the function g is continuous, we conclude that $g^{-1}(R_k)$ contains an open set. □

Theorem 2 *Let $g \in M_{Aff}$. Then for every $\epsilon > 0$, there exist some function $g^1 \in M_{Aff}$ such that $\|g - g^1\|_\infty < \epsilon$, and a set $\{w_\ell\}_{\ell=1}^n$ such that*

$$\mathbf{G}_n^1 = \left[\int_{R^n} \mathbf{y} w_1(g^1(\mathbf{y})) \cdots \int_{R^n} \mathbf{y} w_n(g^1(\mathbf{y})) \right] \tag{10}$$

is full rank.

Proof If $\mathbf{G}_n$ in (7) is full-rank the result is obvious. Otherwise, choose some n arbitrary intervals from the partition $\{R_k\}$, and re-enumerate them such that $\{R_k\}_{k=1}^n$. Let ϵ be some arbitrarily small real number, and define $R_k^\epsilon = (a_k + \epsilon, b_k - \epsilon)$. Due to the original definition of the partition, we have that $\{R_k\}_{k=1}^n$ are disjoint and hence their pre-images with respect to g, *i.e.*, $\{g^{-1}(R_k)\}_{k=1}^n$ are disjoint as well. Next define $q(\mathbf{y}) = \sum_{k=1}^n q_k(\mathbf{y}) \mathbf{1}_{g^{-1}(R_k^\epsilon)}$, where for every k, $q_k(\mathbf{y})$ is continuous, $\|q_k\|_\infty < \epsilon$ and $q_k(\mathbf{y})$ is non-zero only on the support defined by $g^{-1}(R_k^\epsilon)$. Let $g^1 = g + q$, and let $w_k(x) = x\mathbf{1}_{R_k}$. Hence for every column of $\mathbf{G}_n^1$ we have

$$\int_{R^n} \mathbf{y} w_k(g^1(\mathbf{y})) = \int_{R^n} \mathbf{y} w_k(g(\mathbf{y})) + \int_{R^n} \mathbf{y} q_k(\mathbf{y}) \tag{11}$$

as by the definitions of w_k and q, we have that $w_k(g^1(\mathbf{y})) = w_k(g(\mathbf{y})) + q_k(\mathbf{y})$. Next, define the set of functions

$$B_{R_k^\epsilon} = \{q_k \mid \|q_k\|_\infty < \epsilon \;\&\; \mathrm{supp}(q_k) = g^{-1}(R_k^\epsilon)\}$$

Since $B_{R_k^\epsilon}$ is an open ball, by Lemmas 2 and 1, its corresponding map of $\int_{R^n} \mathbf{y} w_k(g^1(\mathbf{y}))$ contains an open ball in R^n. Hence, each of the columns of $\mathbf{G}$ can be modified within an open ball, by an additive term, independently of the modifications of the other columns. This implies that the matrix $\mathbf{G}$ can be modified within an arbitrarily small open ball in R^{n^2}. We therefore conclude that arbitrary small modifications of g (by adding q defined above) result in an arbitrarily small ball around $\mathbf{G}$. As $GL_n(R)$ is dense in R^{n^2} this arbitrarily small ball around $\mathbf{G}$ contains an element in $GL_n(R)$. This element is $\mathbf{G}^1$ and the corresponding function is the desired g^1. □

4 Estimation in the Presence of Model Mismatch

In general, it may happen that there exists a mismatch between the assumed model and the physical one, for example in the presence of noise (of different types: sampling, quantization, measurement), partial occlusions, or when the transformation is not affine but close to it in some sense so that it is desired to best approximate the deformation by an affine transformation. Thus, following the above solution, additional constraints can be added by considering additional compositions $\{w_\ell\}_{\ell=1}^{p} \in W$ with $p \geq n$. Hence, (3) is replaced by

$$\int_{R^n} w_\ell(h(\mathbf{x})) \approx \int_{R^n} w_\ell(g(\mathbf{A}\mathbf{x})) = \left|\mathbf{A}^{-1}\right| \int_{R^n} w_\ell(g(\mathbf{y})) \quad (12)$$

Thus, (12) produces an overdetermined system

$$\begin{bmatrix} \int_{R^n} w_1(h(\mathbf{x})) \\ \vdots \\ \int_{R^n} w_p(h(\mathbf{x})) \end{bmatrix} \approx \left|\mathbf{A}^{-1}\right| \begin{bmatrix} \int_{R^n} w_1(g(\mathbf{y})) \\ \vdots \\ \int_{R^n} w_p(g(\mathbf{y})) \end{bmatrix} \quad (13)$$

which by a least squares solution provides an estimate for $|\mathbf{A}|$. Similarly, (9) becomes now an overdetermined system as p, the number of columns in $\mathbf{G}_p$, is greater than n, the number of rows. Rewriting (9) in the conventional form where p is the number of rows yields

$$\mathbf{G}_p^T(\mathbf{A}^{-1})^T \approx |\mathbf{A}|\,\mathbf{H}_p^T \quad (14)$$

and the solution for $\mathbf{A}^{-1}$ becomes a linear least squares solution

$$\hat{\mathbf{A}}^{-1} = |\mathbf{A}|\,\mathbf{H}_p\mathbf{G}_p^T[\mathbf{G}_p\mathbf{G}_p^T]^{-1} \quad (15)$$

In [26] where we consider the case of a noisy observation h, we elaborate on the model mismatch problem in great detail.

Remark 7 As indicated above, the subset of functions $g \in M_{Aff}$, for which there exists a set $\{w_\ell\}_{\ell=1}^{n}$ such that $\mathbf{G}_n$ is full-rank, is dense. Hence, augmenting the matrix $\mathbf{G}_n$ with additional columns by employing $p > n$ constraints yields on the L.H.S. of (14) a matrix which is rank-n. Thus, for every $g \in M_{Aff}$, or for an infinitesimal modification of it, there exists a least squares solution for $\mathbf{A}^{-1}$.

Remark 8 When model mismatch is considered the question of how to optimally choose the set $\{w_\ell\}$ becomes critical. While in the ideal case discussed in Sect. 3 any arbitrary choice of the set $\{w_\ell\}$ is equally optimal—as long as the resulting matrix $\mathbf{G}_n$ is full rank—the solution to the problem of how to optimally select the set $\{w_\ell\}$ in the presence of model mismatch is entirely different. In the latter case, effects and models of error sources such as sampling, quantization, noise, and illumination variations, must be incorporated into the process of selecting the set $\{w_\ell\}$. The rigorous analyzes of these issues, and their implications on the optimal choice of the set $\{w_\ell\}$ are beyond the scope of the present paper. In [26] (see also [27–29]), these analyzes have been carried out and expressions for the first- and second-order moments of the errors in (13), (14) as functions of the different error models have been derived. This analysis enables the replacement of the above linear least squares solution by a weighted linear least squares, which eliminates the statistical dependencies that result from an arbitrary choice of the set $\{w_\ell\}$. This estimator can be equivalently expressed in terms of an optimal set $\{w_\ell\}$. In the following we briefly elaborate on the principles that lead to optimal choice of the set of non-linear operators:

The derivation is based on the understanding that since our goal is to estimate the geometric transformation, the appropriate noise model for the problem is a model that explicitly relates the presence of noise and the measures of the geometric entities in the observed image. In our case these are the zero- and first-order moments of the observation, after the nonlinear operator w_ℓ was applied to the observation, namely $\int_{R^n} w_\ell(h(\mathbf{x}))$ and $\int_{R^n} \mathbf{x}w_\ell(h(\mathbf{x}))$. In the above mentioned analysis we analytically evaluate the first- and second-order moments of these quantities for each of the different error sources as well as for the combined error model. Having obtained closed form expressions for the first- and second-order moments of $\int_{R^n} w_\ell(h(\mathbf{x}))$ and $\int_{R^n} \mathbf{x}w_\ell(h(\mathbf{x}))$, they can be substituted into the classic equations of the weighted least-squares estimator to yield the optimal linear estimator of the deformation parameters. This estimator is therefore equivalent to optimally calculating the $\{w_\ell\}$'s, using some arbitrary initial choice of the $\{w_\ell\}$'s that is only required to provide a full-rank $\mathbf{G}_p$, when applied to the template.

5 Finding the Affine Transformation Parameters in the Presence of Translation

In this section we extend the solution presented in Sect. 3 by considering the more general problem where the observed object is subject to an affine transformation which includes an unknown translation, *i.e.*, the observation model is given by

$$h(\mathbf{x}) = g(\mathbf{A}\mathbf{x} + \mathbf{c}) \quad (16)$$

In this case the goal is to find the parameters of the affine transformation including the translation. The transformation model is thus given by

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{c}, \quad \mathbf{x} = \mathbf{A}^{-1}\mathbf{y} + \mathbf{b} \quad (17)$$

where $\mathbf{x}$, $\mathbf{y}$, $\mathbf{A}$, $\mathbf{A}^{-1}$ are defined as in Sect. 3, while $\mathbf{c}$ and $\mathbf{b} = -\mathbf{A}^{-1}\mathbf{c}$ are n-dimensional vectors of unknown constants, each representing the translation along a different axis, in the coordinate transformation model and its inverse, respectively. More specifically, let $\tilde{\mathbf{y}} = [1, y_1, \ldots, y_n]^T$. Hence, using (17)

$$\mathbf{x} = \mathbf{T}\tilde{\mathbf{y}} \quad (18)$$

where $\mathbf{T} = [\mathbf{b}|\mathbf{A}^{-1}]$.

The Jacobian of the transformation (18) is the same as in the case where there is no translation (analyzed in Sect. 3) and hence it is found using (4). The evaluation of $\mathbf{T}$ is performed in a similar way to the procedure in Sect. 3: Applying some composition $w \in W$ to the known relation $h(\mathbf{x}) = g(\mathbf{A}\mathbf{x} + \mathbf{c})$ we obtain

$$\begin{aligned}\int_{R^n} \mathbf{x} w(h(\mathbf{x})) &= \int_{R^n} \mathbf{x} w(g(\mathbf{A}\mathbf{x}+\mathbf{c})) = \left|\mathbf{A}^{-1}\right| \int_{R^n} (\mathbf{A}^{-1}\mathbf{y}+\mathbf{b}) w(g(\mathbf{y})) \\ &= \left|\mathbf{A}^{-1}\right| \mathbf{A}^{-1} \int_{R^n} \mathbf{y} w(g(\mathbf{y})) + \left|\mathbf{A}^{-1}\right| \mathbf{b} \int_{R^n} w(g(\mathbf{y})) \quad (19)\end{aligned}$$

Repeating this procedure by applying a family of left-hand compositions $\{w_\ell\}_{\ell=1}^p \in W$ to the known relation $h(\mathbf{x}) = g(\mathbf{A}\mathbf{x} + \mathbf{b})$, and rewriting it in a matrix form yields

$$\begin{aligned}|\mathbf{A}| \left[\int_{R^n} \mathbf{x} w_1(h(\mathbf{x})) \cdots \int_{R^n} \mathbf{x} w_p(h(\mathbf{x}))\right] \\ = \mathbf{T} \left[\int_{R^n} \tilde{\mathbf{y}} w_1(g(\tilde{\mathbf{y}})) \cdots \int_{R^n} \tilde{\mathbf{y}} w_p(g(\tilde{\mathbf{y}}))\right] \quad (20)\end{aligned}$$

Let

$$\tilde{\mathbf{G}}_p = \left[\int_{R^n} \tilde{\mathbf{y}} w_1(g(\mathbf{y})) \cdots \int_{R^n} \tilde{\mathbf{y}} w_p(g(\mathbf{y}))\right] \quad (21)$$

and hence (20) can be written in a more compact form as $|\mathbf{A}|\mathbf{H}_p = \mathbf{T}\tilde{\mathbf{G}}_p$. We therefore have the following conclusion:

Theorem 3 *Let $\mathbf{A} \in GL_n(R)$. Assume $h, g \in M_{Aff}$ such that $h(\mathbf{x}) = g(\mathbf{A}\mathbf{x} + \mathbf{c})$. Given measurements of h and g, then $\mathbf{A}$ and $\mathbf{c}$ can be uniquely determined if there exists a function $w \in W$ such that $\int_{R^n} w(g(\mathbf{y})) \neq 0$ and a set of functions $\{w_\ell\}_{\ell=1}^{n+1} \in W$ such that $\tilde{\mathbf{G}}_{n+1}$ is full rank. Then, $\mathbf{T} = |\mathbf{A}|\mathbf{H}_{n+1}(\tilde{\mathbf{G}}_{n+1})^{-1}$.*

Remark 9 As in the previous case, the elements of the matrix $\tilde{\mathbf{G}}_{n+1}$ depend only on the template and its coordinate system and thus have to be evaluated only *once*. Therefore, the denominator of (4) together with $\tilde{\mathbf{G}}_{n+1}$ represent all the information in the template, required for finding the affine transformation parameters including the translation. Hence, the denominator of (4) together with $\tilde{\mathbf{G}}_{n+1}$ form a sufficient representation of the template.

Remark 10 As in Sect. 4, if the model is only an approximate one due to model mismatch, the Jacobian is determined by (13), while $\mathbf{T}$ is evaluated by a least squares solution obtained by taking $p \geq n + 1$. In that case

$$\hat{\mathbf{T}} = |\mathbf{A}|\, \mathbf{H}_p \tilde{\mathbf{G}}_p^T [\tilde{\mathbf{G}}_p \tilde{\mathbf{G}}_p^T]^{-1} \quad (22)$$

6 Finding the Affine Transformation in the Presence of an Unknown Spatially Constant Amplitude Gain

In the analysis carried out so far it has been assumed that there is no amplitude variation (illumination variation, in the case of images) between the template and the observation, and hence the observed deformation is only due to the geometric distortion of the coordinate system caused by the affine transformation. In this section we generalize the proposed solution and address a more general deformation model where the model given by (1) is replaced by

$$\begin{aligned}&h(\mathbf{x}) = a g(\mathbf{A}\mathbf{x} + \mathbf{c}) \\ &\mathbf{A} \in GL_n(R),\ \mathbf{x}, \mathbf{c} \in R^n,\ a \in R,\ a > 0 \quad (23)\end{aligned}$$

where a, $\mathbf{A}$ and $\mathbf{c}$ are unknown and need to be determined. As we prove in this section, the problem of finding the gain factor amounts to replacing the step in which the Jacobian of the transformation is being determined in the case where there is no gain change, by a step in which both the gain and the Jacobian are jointly determined. More specifically, let $w_1, w_2 \in W$ be a pair of Lebesgue measurable and separable functions such that $w_i(xy) = w_i(x)w_i(y)$ (for example, let $w_1(x) = x^2$, $w_2(x) = x^4$), and where $\int_{R^n} w_i(g(\mathbf{y})) \neq 0$. Since $h(\mathbf{x}) = a g(\mathbf{A}\mathbf{x} + \mathbf{c})$, we have

$$\begin{aligned}\int_{R^n} w_1(h(\mathbf{x})) &= w_1(a) \int_{R^n} |\mathbf{A}^{-1}| w_1(g(\mathbf{y})) \\ &= w_1(a)|\mathbf{A}^{-1}| \int_{R^n} w_1(g(\mathbf{y})) \quad (24)\end{aligned}$$

and a similar expression is obtained by applying w_2.

Hence,

$$|\mathbf{A}^{-1}| w_1(a) = \frac{\int_{R^n} w_1(h(\mathbf{x}))}{\int_{R^n} w_1(g(\mathbf{y}))} \quad (25)$$

and

$$|\mathbf{A}^{-1}| w_2(a) = \frac{\int_{R^n} w_2(h(\mathbf{x}))}{\int_{R^n} w_2(g(\mathbf{y}))} \quad (26)$$

Thus, both the Jacobian, $|\mathbf{A}^{-1}|$, and the gain factor, a, can be evaluated using (25)–(26). Having estimated a, the original problem in (23) can be rewritten as $\frac{1}{a}h(\mathbf{x}) = g(\mathbf{A}\mathbf{x} + \mathbf{c})$. This however is exactly the problem solved in Sect. 5, with $h(\mathbf{x})$ replaced by $\frac{1}{a}h(\mathbf{x})$. Hence, defining H_p in this case to have the form

$$\mathbf{H}_p = \left[\int_{R^n} \mathbf{x} w_1(\frac{1}{a}h(\mathbf{x})) \cdots \int_{R^n} \mathbf{x} w_p\left(\frac{1}{a}h(\mathbf{x})\right)\right] \quad (27)$$

we obtain along the same lines of arguments that:

Theorem 4 *Let $\mathbf{A} \in GL_n(R)$. Assume $h, g \in M_{Aff}$ such that $h(\mathbf{x}) = a g(\mathbf{A}\mathbf{x} + \mathbf{c})$, and a is an unknown real and positive* *gain coefficient. Given measurements of h and g, then* $\mathbf{A}$, $\mathbf{c}$ *and* a *can be uniquely determined if there exists a pair of separable functions* $\{w_i\}_{i=1}^2 \in W$ *such that* $\int_{R^n} w_i(g(\mathbf{y})) \neq 0$ *and a set of functions* $\{w_\ell\}_{\ell=1}^{n+1} \in W$ *such that* $\tilde{\mathbf{G}}_{n+1}$ *is full rank. Then,* $\mathbf{T} = |\mathbf{A}|\mathbf{H}_{n+1}(\tilde{\mathbf{G}}_{n+1})^{-1}$.

Remark 11 Similarly to the previous cases, where the gain is assumed fixed, the denominators of (25) and (26), as well as the elements of $\tilde{\mathbf{G}}_{n+1}$ depend only on the template and its coordinate system and thus have to be evaluated only *once*. Therefore, the denominators of (25), (26) together with the matrix $\tilde{\mathbf{G}}_{n+1}$ represent all the information in the template, required for finding the affine transformation parameters including the translation, in the case where the gain is unknown. Hence, the denominators of (25) and (26) together with $\tilde{\mathbf{G}}_{n+1}$ form a sufficient representation of the template.

Remark 12 As in Sect. 4, if the model is only an approximate one due to model mismatch, the solution for $\mathbf{A}$ and $\mathbf{c}$ becomes a least squares solution to yield

$$\hat{\mathbf{T}} = |\mathbf{A}|\,\mathbf{H}_p\tilde{\mathbf{G}}_p^T[\tilde{\mathbf{G}}_p\tilde{\mathbf{G}}_p^T]^{-1} \tag{28}$$

where $\mathbf{H}_p$ is defined in (27), and $p \geq n+1$.

7 Discussion, Analysis and Comparison with Existing Solutions

As the proposed solution employs evaluation of moments, and solution to a linear system of equations, it may seem that it bares similarities with existing linear solutions to the problem and with moment based methods. In the sequel we explain in some detail why the methodology presented in this paper is completely different from these methods, and how it employs an entirely different approach to the problem.

7.1 Linear Methods

As indicated earlier, for the case where the affine deformation can be assumed small and the observed functions are differentiable, a widely used approach is to linearize the problem using an approximation based on a *first order* Taylor series expansion. See, *e.g.*, [21]. In order to simplify the exposition we address the two dimensional case. More specifically, using the small deformation assumption, the original problem is approximated by $h(x,y) \cong g(x + (ax+by+c), y + (dx+ey+f))$, where $(ax+by+c)$ and $(dx+ey+f)$ are small enough such that for g itself

$$\begin{aligned} &g(x + (ax+by+c), y + (dx+ey+f)) \\ &\quad\cong g(x,y) + (ax+by+c)\frac{\partial g}{\partial x}(x,y) \\ &\qquad + (dx+ey+f)\frac{\partial g}{\partial y}(x,y) \end{aligned} \tag{29}$$

One may next define the distance between the observation and the approximating deformed template:

$$\begin{aligned} &V(a,b,c,d,e,f) \\ &\quad= \sum_{(x,y)}[h(x,y) - g(x + (ax+by+c), \\ &\qquad y + (dx+ey+f))]^2 \end{aligned} \tag{30}$$

Substitution of (29) into (30) yields

$$\begin{aligned} &V(a,b,c,d,e,f) \\ &\quad= \sum_{(x,y)}\bigg[h(x,y) - g(x,y) - (ax+by+c)\frac{\partial g}{\partial x}(x,y) \\ &\qquad - (dx+ey+f)\frac{\partial g}{\partial y}(x,y)\bigg]^2 \end{aligned}$$

Let $I_x = \frac{\partial g}{\partial x}(x,y)$, $I_y = \frac{\partial g}{\partial y}(x,y)$, $I_d = g(x,y) - h(x,y)$. Minimizing $V(a,b,c,d,e,f)$ by taking its partial derivatives with respect to the six transformation parameters and equating each result to zero, one obtains

$$\begin{aligned} \frac{\partial V}{\partial a} &= \sum 2\bigg[h(x,y) - g(x,y) - (ax+by+c)\frac{\partial g}{\partial x}(x,y) \\ &\quad - (dx+ey+f)\frac{\partial g}{\partial y}(x,y)\bigg]x\frac{\partial g}{\partial x}(x,y) = 0 \end{aligned} \tag{31}$$

and similarly for the derivatives with respect to the remaining parameters. Reorganization of the six equations of the form (31) results in the following system of linear constraints:

$$\begin{bmatrix} \sum xI_xxI_x & \sum yI_xxI_x & \sum I_xxI_x & \sum xI_yxI_x & \sum yI_yxI_x & \sum I_yxI_x \\ \sum xI_xyI_x & \sum yI_xyI_x & \sum I_xyI_x & \sum xI_yyI_x & \sum yI_yyI_x & \sum I_yyI_x \\ \sum xI_xI_x & \sum yI_xI_x & \sum I_xI_x & \sum xI_yI_x & \sum yI_yI_x & \sum I_yI_x \\ \sum xI_xxI_y & \sum yI_xxI_y & \sum I_xxI_y & \sum xI_yxI_y & \sum yI_yxI_y & \sum I_yxI_y \\ \sum xI_xyI_y & \sum yI_xyI_y & \sum I_xyI_y & \sum xI_yyI_y & \sum yI_yyI_y & \sum I_yyI_y \\ \sum xI_xI_y & \sum yI_xI_y & \sum I_xI_y & \sum xI_yI_y & \sum yI_yI_y & \sum I_yI_y \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} = \begin{bmatrix} \sum I_dxI_x \\ \sum I_dyI_x \\ \sum I_dI_x \\ \sum I_dxI_y \\ \sum I_dyI_y \\ \sum I_dI_y \end{bmatrix} \tag{32}$$

Provided that the matrix is invertible, a solution for the affine transformation parameters is obtained for the case where the deformation is indeed small. On the other hand, the linear solution proposed in this paper is *exact* and is applicable to any affine transformation *regardless of its magnitude*. Moreover, while in the presence of noise the suggested method allows for an increase in the number of equations which leads to a LS solution for the affine transformation parameters, the linearized approximate solution is based on only six equations, and no additional linear constraints on the parameters can be added to the system. Also note that the linearized approximation requires the evaluation of the image spatial derivatives. This evaluation is highly problematic since the observed objects are usually not continuous nor differentiable everywhere. This restriction becomes even more problematic in the presence of noise. On the other hand, the method derived in this paper employs only integration of the observed data and hence is much more robust.

7.2 Moment Based Methods

In this section we briefly describe moment based methods, their advantages and limitations, in the context of deformation estimation and registration problems. We then elaborate on the major conceptual and practical differences between moment based methods and the solution we proposed in the previous sections.

Again, in order to simplify the exposition we address the two dimensional case. Thus the relation between the template and observation is given by $g(x,y) = h((ax + by + c, dx + dy + f)^{-1})$. Define the functional $\mu_{p,q}(h) = \int\int x^p y^q h(x,y)dxdy$. This functional is called the moment of order (p,q). Thus,

$$\begin{aligned}
&\mu_{p,q}(h((ax+by+c, dx+ey+f)^{-1})) \\
&\quad = \int\int x^p y^q h((ax+by+c, dx+ey+f)^{-1})dxdy \\
&\quad = \frac{1}{\Delta}\int\int (ax+by+c)^p(dx+ey+f)^q h(x,y)dxdy \\
&\quad = \frac{1}{\Delta} P_{p,q}(a,b,c,d,e,f,\{\mu_{k,j}(h(x,y))\}_{k\le p, j\le q}) \qquad (33)
\end{aligned}$$

where $P_{p,q}$ is a homogenous polynomial of degree $p+q$ in a,b,c,d,e,f, linear in $\{\mu_{k,j}(h)\}_{k\le p,j\le q}$ and Δ is the Jacobian of the transformation.

The first two moments $\mu_{1,0}$ and $\mu_{0,1}$ have a special geometric meaning, as they provide the center of mass of the "object", h, when normalized by $\mu_{0,0}$. Thus, the centers of mass of h and g are related linearly by the above analysis, and therefore we have two linear constraints on the 6 unknown parameters. Obviously, this is not enough in order to find all the parameters we are after. We further note that an increase in the dimension, n, of the space the affine transformation is defined on, increases the number of parameters as $n^2 + n$. Yet, the number of first order moments is n. Hence, we *must* use higher order moments in order to solve for all the transformation parameters.

We note that in the special case where it is assumed that the transformation is composed of only rotation and translation (so that we are looking for only three parameters: two for translation and one for rotation), one can derive an explicit and simultaneous solution for the translation and rotation parameters in terms of moments. However when the transformation is some unrestricted element of the affine group, no explicit solution exists.

As a result of the lack of general methods to solve systems of polynomial equations, the attempt to solve the system of polynomial equations obtained by employing high order moments was generally abandoned. Instead, research turned to invariant theories in order to find relations between the moments of the template and observation. See *e.g.*, [12, 13] and the references therein. These relations enable the description of some properties of the function, independently of the transformation itself. Hence, such descriptions enable the design of object recognition methods that are independent of the affine transformation. In conclusion, moment based methods force us either to skip the problem of estimating the transformation parameters, or alternatively, to solve a system of polynomial equations.

On the other hand, the methodology developed in this paper provides an entirely different view point on how the problem should be solved. This novel approach allows for an explicit estimation of the deformation model, while employing a solution to a linear system of equations obtained by using *only* first order moments of various nonlinear operators applied to the template and observation. In other words, the application of the nonlinear operators maps the problem into a new "coordinate system" where now a *linear* mapping relates the first order moments of these nonlinear functions of the observation and the template.

8 Experimental Results

8.1 Numerical Examples

In this section we present some numerical examples to illustrate the operation and robust performance of the proposed parameter estimation algorithm, for a broad range of deformations. Note that in these experiments the applied nonlinear operators $\{w_\ell\}$ were chosen such that the resulting system is full rank, yet they are not optimal in any sense.

The examples illustrates the operation of the proposed algorithm on a car image. The template image dimensions are 3100×1200. It is shown in the bottom image of Fig. 1. The

Fig. 1 From *bottom* to *top*: Template; Estimated deformed object obtained by applying the deformation estimated from the observation to the template; Observation on the deformed object

observed deformed image is shown in the upper image of the figure. The image coordinate system is $[-1,1]\times[-1,1]$. The translation vector is $[-0.1,-0.14]$ and the translation estimation error vector is $[2\cdot 10^{-5}, 2.4\cdot 10^{-4}]$. The deforming transformation is given by

$$\mathbf{A}=\begin{pmatrix}-0.988 & 0.454\\ 0.156 & -0.891\end{pmatrix}$$

where the estimate obtained by the proposed procedure is

$$\hat{\mathbf{A}}=\begin{pmatrix}-0.987 & 0.453\\ 0.156 & -0.891\end{pmatrix}$$

To provide some measure for the magnitude of the error in estimating the deformation matrix we list the value of $\mathbf{A}^{-1}\hat{\mathbf{A}}-\mathbf{I}$:

$$\mathbf{A}^{-1}\hat{\mathbf{A}}-\mathbf{I}=\begin{pmatrix}9.943 & -2.301\\ -1.298 & 0.228\end{pmatrix}\cdot 10^{-4}$$

Finally, the estimated deformation is applied to the original template in order to obtain an estimate of the deformed object (middle image in Fig. 1) which can be compared with the deformed observation shown in the upper image.

In [22] a method that employs the commutativity of the affine and scale operators is employed to estimate the parameters of an affine transformation by inducing a series of scaling operations to the observed image and template. In [22] an experimental evaluation of the method relative to other known methods such as the cross-weighted moments [11] and the moment descriptor method [23] is provided. Following this evaluation we adopted the same metric used in [22], *i.e.*,

$$d=\frac{1}{2}\sum_{i=1}^{2}\frac{\|(\mathbf{A}-\hat{\mathbf{A}})\mathbf{p}_i\|}{\|\mathbf{A}\mathbf{p}_i\|} \tag{34}$$

where $\mathbf{p}_1=(1,0)^T$ and $\mathbf{p}_2=(0,1)^T$, to compare the performance of the method proposed in this paper to the performance of existing ones. Application of the proposed method yields an average error of 0.001, compared with 0.02, and 0.03 obtained by the scaling-operator based method and by the moment descriptor method, respectively. Note that the results achieved by the proposed method, using a sub-optimal set of operators, yield performance improvement of a magnitude of order over existing solutions.

The second example employs the same template as the first one and it is shown in the bottom image in Fig. 2. The observed deformed image is shown in the upper image of the figure. This image is both smaller in scale than in the first example and is observed with lower illumination, such that the illumination gain is $a=0.58$. The error in estimating the gain is $\hat{a}-a=0.726\cdot 10^{-7}$. The image coordinate system is $[-1,1]\times[-1,1]$. The translation vector is $[-0.4,-0.5]$ and

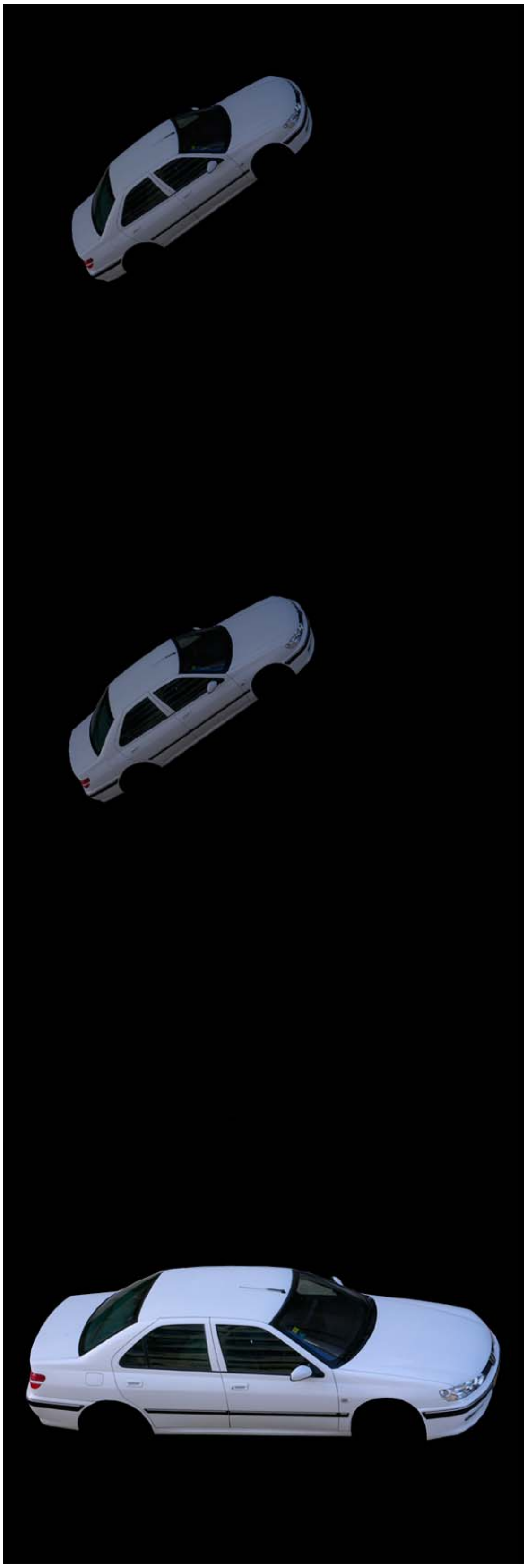

Fig. 2 From *bottom* to *top*: Template; Estimated deformed object obtained by applying the deformation estimated from the observation to the template; Observation on the deformed object

the translation estimation error vector is $[1.85 \cdot 10^{-3}, 3.92 \cdot 10^{-5}]$. The deforming transformation is given by

$$\mathbf{A} = \begin{pmatrix} 0.4854 & 0.3527 \\ -0.3527 & 0.4854 \end{pmatrix}$$

where the estimate obtained by the proposed procedure is

$$\hat{\mathbf{A}} = \begin{pmatrix} 0.4722 & 0.3626 \\ -0.353 & 0.4858 \end{pmatrix}$$

while

$$\mathbf{A}^{-1}\hat{\mathbf{A}} - \mathbf{I} = \begin{pmatrix} 0.81 & -2.65 \\ 0.001 & -0.11 \end{pmatrix} \cdot 10^{-2}$$

Finally, the estimated deformation is applied to the original template in order to obtain an estimate of the deformed object (middle image in Fig. 2) which can be compared with the deformed observation shown in the upper image.

Evaluation of the estimation results using the metric proposed in [22], yields an average error of $d = 0.019$, which is similar to error measurements obtained by the scaling-operator based method and the moment descriptor method, when *no* scaling nor illumination variations are involved.

It should be noted that the estimation errors documented in the above experiments result from numerical errors in synthesizing the observation, sampling, quantization effects, and from approximating integrals by sums. Detailed analysis of the effects of these error sources, on the evaluation of $\mathbf{H}_p$ can be found in [26], where the mean and covariance of the estimates are derived. This detailed analysis enables us to successfully employ the suggested method to estimate transformations in realistic scenarios as illustrated in the next example.

8.2 Implementation in the Presence of Model Mismatch: Estimation of Orientation in Space

The final example describes a scheme for estimating the orientation in space (pan and tilt angles) of a planar object (a picture) in a realistic scenario where from a single image, the orientation angles of the plane need to be determined. Note, that even at the level of the problem definition, the current problem deviates from the previously derived analytic model, as the affine transformations of the observed planar object are only an approximation to the perspective deformation induced by the camera. The setting of the experiment is as follows: In each repetition of the experiment the pan and tilt angles of the plane are drawn at random from uniform distributions on $[-45, 45]$ degrees for the pan and on $[-15, 15]$ degrees for the tilt (these are fed to a computer controlled stage); an image of the plane is then taken by the camera (standard consumer grade, 1280×1024 pixels) and

Fig. 3 A set of observations on a planar surface

the pan and tilt angles are evaluated from this single observation, using the linear least squares procedure derived in the previous sections. A set of example images taken in this experiment is shown in Fig. 3. To isolate the picture from the black background, a crud segmentation, based only on the color range information was employed. The experiment itself was performed for 4000 independent trials. The ground truth measurements vs. the estimated pan and tilt angles, in a short segment of the experiment are shown in Fig. 4. Since in this resolution, the graphs overlap we depict in Fig. 5 the error curves in estimating the surface tilt and pan angles along the entire sequence of 4000 experiments. For the entire sequence of 4000 experiments, the bias in estimating the pan angle is $-2.15 \cdot 10^{-13}$ degrees and the standard deviation is $1.13 \cdot 10^{-2}$, with a maximally encountered error of 0.45 degrees. The bias in estimating the tilt angle is $-5.46 \cdot 10^{-14}$ degrees and the standard deviation is $7 \cdot 10^{-3}$, with a maximally encountered error of 0.46 degrees. We therefore conclude that the estimates of both the pan and tilt angles are unbiased. For qualitative comparison, recent results, [30], on estimating the orientation of an object using specular flow report on average errors between 2.3 and 5 degrees, where using the proposed solution the maximal encountered error is less than 0.5 degrees. Moreover, in approximating the real deformation by an affine deformation and evaluating the error metric (34) using the true and estimated angles, we obtain an average error measure of 0.0002 in this real-world setting, which is by two orders of magnitude lower than the error obtained in synthetic experiments, by [22], and by an order of magnitude lower than the error measure obtained by the proposed method using an arbitrary choice of the non-linear operators. Furthermore, the *maximal* error in our real-world experiment when measured by the error metric (34), is 0.0014, which is by an order of magnitude lower than the *average* error of [22] in a synthetic setting. In terms of computational requirements, on a Pentium 3 1.13 GHz system, using a code written in Matlab, it takes 0.01 s to estimate the tilt and pan angles for each frame.

Note that the robustness of the proposed solution is demonstrated by the fact that despite all the mismatches between the assumed affine deformation model and the deformations encountered in the actual application, namely, the geometric transformation is only approximately affine (due to the perspective projection), the spatially varying changes in the illumination of the object surface due to its movements relative to the light source, and the poor quality of the

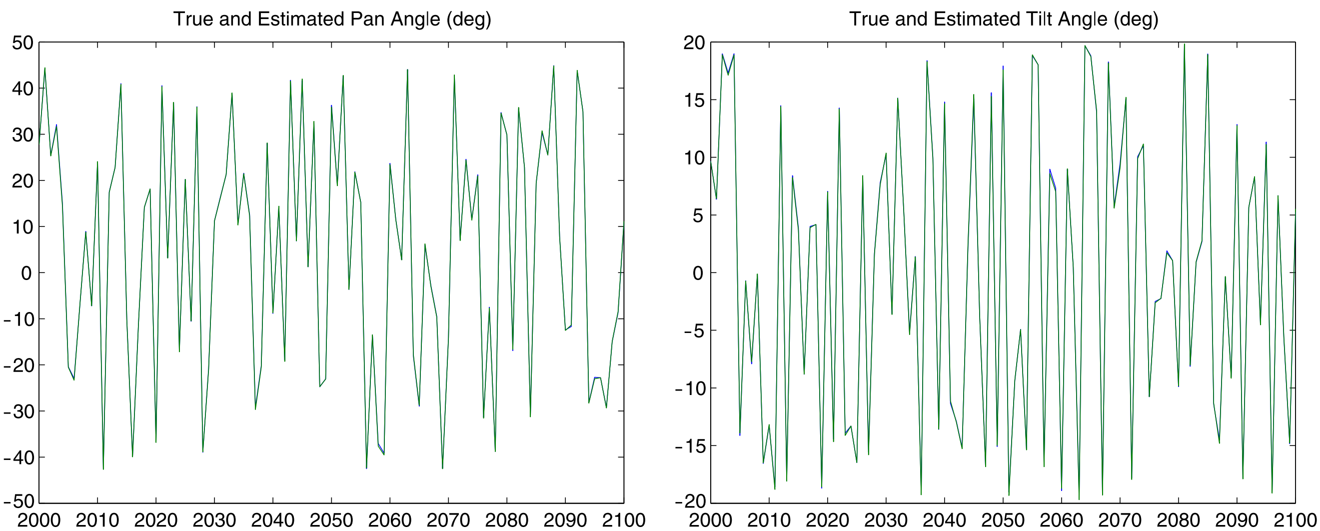

Fig. 4 (Color online) Surface pan and tilt angles: True (*blue*) vs. estimated (*green*)

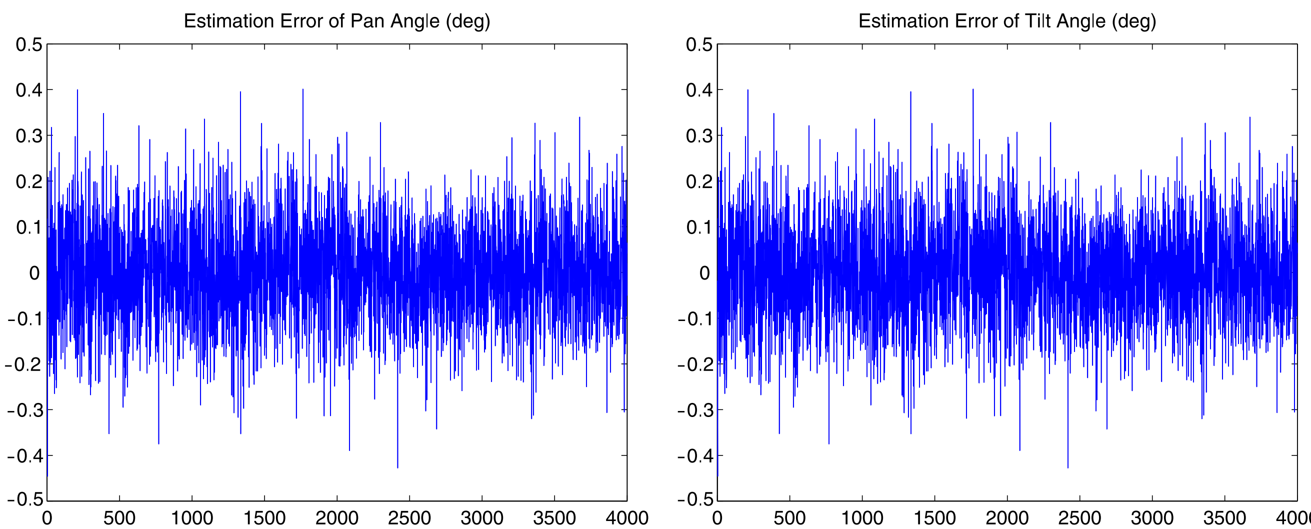

Fig. 5 Surface pan and tilt angles estimation errors along the entire experiment

segmentation procedure—the parametric estimates are unbiased and the standard deviation is in the order of a hundredth of a degree.

9 Conclusions

In this paper we have considered the problem of finding the affine transformation relating a given observation on a planar object with some pre-chosen template of this object. The direct approach for estimating the transformation is to apply each of the deformations in the affine group to the template in a search for the deformed template that matches the observation. We propose a method that employs a set of non-linear functionals to replace this high dimensional problem by an *equivalent linear problem*, expressed in terms of the unknown affine transformation parameters. Thus, the problem of finding the parametric model of the affine deformation is mapped, by this set of non-linear functionals, into a set of *linear* equations which is then solved for the affine transformation parameters. The proposed solution has been further extended to include the case where the deformation relating the observed signature of the object and the template, is composed of both a geometric deformation due to the affine transformation of the coordinate system, and an unknown amplitude gain factor. The proposed solution is *unique and exact* and is applicable to any affine transformation *regardless of its magnitude*.

In conclusion, the novel framework for estimating affine transformations presented in this paper has several advantages over the existing methods. Although each of the advantages individually is not unique to this method, the combination of properties is unique and to the best of our knowledge provides superior performance over all known methods. The method is explicit, global, deals simultaneously with all the affine transformations, and the map $\varphi = H(h, g)$ is continuous and involves only elementary linear analysis in the same dimension as that of the group model. To increase noise immunity, it employs integral operators, rather than differential ones. Although in this paper we concentrate on the solution of the most fundamental problems in this field, the same framework can be extended to include much more complicated problems, such as the estimation of elastic deformations, [9, 24]. In forthcoming papers we shall extend the scope of the method presented here, to such problems as the analysis of its performance in the presence of noise, the optimal selection of the nonlinear operators in the presence of noise, and its extensions to various scenarios where the deformations are both of geometry and amplitude. As illustrated in Sect. 8.2, these steps allow one to elegantly and accurately cope with realistic scenarios using the methodology derived in this paper.

References

1. Grenander, U.: General Pattern Theory. Oxford University Press, London (1993)
2. Brown, L.: A survey of image registration techniques. ACM Comput. Surv. **24**(4), 326–376 (1992)
3. Zitova, B., Flusser, J.: Image registration methods: a survey. Image Vis. Comput. **21**, 977–1000 (2003)
4. Amit, Y.: 2-D Object Detection and Recognition. MIT Press, Cambridge (2002)
5. Milanfar, P.: Two-dimensional matched filtering for motion estimation. IEEE Trans. Image Process. **8**, 438–443 (1999)
6. Keller, Y., Averbuch, A., Israeli, M.: Pseudopolar-based estimation of large translations, rotations, and scalings in images. IEEE Trans. Image Process. **14**, 12–22 (2005)
7. Keller, Y., Shkolnisky, Y., Averbuch, A.: The angular difference function and its application to image registration. IEEE Trans. Pattern Anal. Mach. Intell. **27**, 969–976 (2005)
8. Foroosh, H., Zerubia, J., Berthod, M.: Extension of phase correlation to subpixel registration. IEEE Trans. Image Process. **11**, 188–200 (2002)
9. Francos, J.M., Hagege, R., Friedlander, B.: Estimation of multidimensional homeomorphisms for object recognition in noisy environments. In: Proc. Thirty Seventh Asilomar Conference on Signals, Systems, and Computers (2003)
10. Hagege, R., Francos, J.M.: Parametric estimation of two-dimensional affine transformations. In: Proc. Int. Conf. Acoust., Speech, Signal Processing, Montreal (2004)
11. Yang, Z., Cohen, F.S.: Cross-weighted moments and affine invariants for image registration and matching. IEEE Trans. Pattern Anal. Mach. Intell. **21**(8), 804–814 (1999)
12. Flusser, J., Suk, T.: Pattern recognition by affine moment invariants. Pattern Recogn. **26**, 167–174 (1993)
13. Kotoulas, L., Andreadis, I.: Accurate calculation of image moments. IEEE Trans. Image Process. **16**(8), 2028–2037 (2007)
14. Viola, P., Wells, W.M.: Alignment by maximization of mutual information. Int. J. Comput. Vis. 134–154 (1997)
15. Kendall, D.G., Barden, D., Carne, T.K., Le, H.: Shape and Shape Theory. Wiley, New York (1999)
16. Younes, L.: Tutorial: deformation, warping and object comparison. In: Proc. ECCV'2000
17. Lucchese, L.: A frequency domain technique based on energy radial projections for robust estimation of global 2D affine transformations. Comput. Vis. Image Underst. **81**, 72–116 (2001)
18. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **60**, 91–110 (2004)
19. Matas, J., Chum, O., Urban, M., Pajdla, T.: Robust wide baseline stereo from maximally stable extremal regions. In: Proc. British Mach. Vis. Conf., pp. 384–393 (2002)
20. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. IEEE Trans. Pattern Anal. Mach. Intell. **27**, 1615–1630 (2005)
21. Keren, D., Peleg, S., Brada, R.: Image sequence enhancement using sub-pixel displacements. In: Proc. IEEE Conf. Comp. Vision and Patt. Recog., pp. 742–746, Ann Arbor (1988)
22. Kannala, J., Rahtu, E., Heikkilä, J., Salo, M.: A new method for affine registration of images and point sets. In: Proc. Scandinavian Conf. Image Anal. (Joensuu, Finland). Lecture Notes in Computer Science, vol. 3540, pp. 224–234. Springer, Berlin (2005)
23. Heikkilä, J.: Pattern matching with affine moment descriptors. Pattern Recogn. **37**, 1825–1834 (2004)
24. Francos, J.M., Hagege, R.: Parametric estimation of multidimensional homeomorphic transformations. United States Patent 7,406,197, November 2004
25. Hagege, R., Francos, J.M.: Linear estimation of sequences of multi-dimensional affine transformations. In: Int. Conf. Acoust., Speech, Signal Processing, Toulouse (2006)
26. Hagege, R.: Estimation of geometric transformations of the domain of functions. Ph.D. Dissertation, Ben-Gurion University (2008)
27. Kovalsky, S.Z., Cohen, G., Francos, J.M.: Registration of joint geometric and radiometric image deformations in the presence of noise. In: Proc. IEEE 14th Workshop on Statistical Signal Processing, pp. 561–565 (2007)
28. Cohen, G., Francos, J.M., Hagege, R.: Joint segmentation and registration of elastically deformable objects. In: Proc. Int. Conf. Patt. Rec., ICPR (2008)
29. Hagege, R., Francos, J.M.: A random sets framework for error analysis in estimating geometric transformations—a first order analysis. In: Proc. Int. Symp. Info. Theory and Appl., Auckland, New Zealand (2008)
30. Chang, J.Y., Raskar, R., Agrawal, A.: 3D pose estimation and segmentation using specular cues. In: Proc. CVPR (2009)

Rami Hagege received the B.Sc. (Summa Cum Laude), M.Sc. (Summa Cum Laude), and Ph.D. degrees in electrical and computer engineering in 2002, 2004 and 2009, respectively, all from Ben-Gurion University, Beer Sheva, Israel. Currently he is with the Massachusetts Institute of Technology Laboratory of Information and Decision Systems.

Joseph M. Francos received the B.Sc. degree in computer engineering in 1982, and the D.Sc. degree in electrical engineering in 1991, both from the Technion-Israel Institute of Technology, Haifa.

From 1982 to 1987, he was with the Signal Corps Research Laboratories, Israeli Defense Forces. From 1991 to 1992 he was with the Department of Electrical Computer and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY, as a Visiting Assistant Professor. During 1993, he was with Signal Processing Technology, Palo Alto, CA. In 1993 he joined the Department of Electrical and Computer Engineering, Ben-Gurion University, Beer-Sheva, Israel, where he is now a Professor. He heads the Mathematical Imaging Group, and the Signal Processing track. He also held visiting positions at the Massachusetts Institute of Technology Media Laboratory, Cambridge, at the Electrical and Computer Engineering Department, University of California, Davis, at the Electrical Engineering and Computer Science Department, University of Illinois, Chicago, and at the Electrical Engineering Department, University of California, Santa Cruz. His current research interests are in image registration, estimation of object deformations form images, parametric modeling and estimation of 2-D random fields, random fields theory, and texture analysis and synthesis.

Dr. Francos served as an Associate Editor for the IEEE Transactions on Signal Processing from 1999 to 2001.