

# Combined Affine Geometric Transformations and Spatially Dependent Radiometric Deformations: A Decoupled Linear Estimation Framework

Jacob Bentolila and Joseph M. Francos, *Senior Member, IEEE*

**Abstract**—This paper considers the problem of registering two observations of the same object, where the observations differ due to a combined effect of an affine geometric transformation and nonuniform illumination changes. The problem of deriving new representations of the observations that are both invariant to geometric transformations and linear in the illumination model is analyzed. In this framework, we present a novel method for linear estimation of illumination changes in an affine invariant manner, thus, decoupling the original problem into two simpler ones. The computational complexity of the method is low as it requires no more than solving a linear set of equations. The prior step of illumination estimation is shown to improve the accuracy of state-of-the-art registration techniques by a factor of two.

**Index Terms**—Affine invariance, affine invariant features, geometric distortion, illumination invariance, object tracking, pose estimation.

## I. INTRODUCTION

THE same object, viewed from different points, or under different illumination conditions, can have many different appearances. Relating the different appearances to the object and analyzing the changes in the appearance are very complex problems.

In many cases, the change of viewing point results in a geometric transformation of the observation like shift, rotation, or scale of the image coordinates. The problem of “aligning” the different coordinate systems to a single coordinate system is called image registration. Image registration is essential for integrating data from different measurements. The information derived from the geometric transformation is used in applications such as object tracking [1] and pose estimation [2]. Applications that detect changes in objects, such as in image-based quality control, require both the registration of the object and compensation for illumination changes. A survey of registration techniques is given in [3].

Manuscript received July 21, 2010; revised December 03, 2010; accepted March 10, 2011. Date of publication April 05, 2011; date of current version September 16, 2011. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Nikolaos V. Boulgouris.

The authors are with the Department of Electrical and Computer Engineering, Ben-Gurion University, Beer Sheva 84105, Israel (e-mail: bentolil@bgu.ac.il; francos@ee.bgu.ac.il).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2011.2136353

In this paper, we focus on affine transformations of the coordinate system. The affine transformation consists of linear transformations (rotation, scaling, and shear) and a translation (shift). It provides a good approximation for the changes in pose of objects that are relatively far from the camera. The affine transformation also serves as the basic block in the analysis and registration of general and nonrigid geometric transformations. When the area considered is small enough, the affine transformation serves as a first-order Taylor series approximation of any differentiable geometric transformation.

The most popular methods for estimating the geometric transformations today are based on local features, such as intensity-based regions (IBR) and edge-based region (EBR) [4], and scale-invariant feature transform (SIFT) [5] and maximally stable extremal regions (MSER) [6]. These methods identify features of small regions in the image and extract the transformation from the correspondence of the features. The correspondence problem is solved by using features that are invariant to the geometric transformation. Affine invariant features include multiscale autoconvolution (MSA) [7], affine invariant moments [8], cross-weighted (CW) moments [9], and trace transform [10]. Affine and illumination invariant features for color images are presented in [11]; however, the illumination changes are global and not location dependent.

Most methods employing local features handle illumination changes by normalization of the illumination in the considered area, or by using edge (corner) information, which is less sensitive to the variations in illumination. However, as shown in the following, the localization of the features and the success of the registration process are effected by the changes in lighting conditions.

Global registration methods estimate the parameters of the geometric transformation from the whole image without a prior stage of local feature extraction [12]–[14]. Since global algorithms treat the image as a whole, the background should be separated from the registered part of the image prior to the registration. Global algorithms tend to show robustness to noise. Sensitivity to radiometric changes, on the other hand, is a major disadvantage of most global approaches.

Some affine registration methods avoid dealing with the illumination changes by discarding the radiometric information and treating the image as a binary image. A comparison of several binary registration methods is presented in [15]. It is shown in [15] that although the method complexity grows linearly with the number of pixels, the calculation of moments is only a marginal factor in the total calculation time, mainly due to the need

to solve high-order polynomial equations. The reported calculation time is about 1 s even for small images.

The variation in illumination between images of the same object creates a major complication for recognition and registration algorithms. Most registration methods measure properties in the image that are either robust or invariant to the illumination changes [16]. However, such methods assume that the changes in the illumination are location independent; therefore, they are only applicable in small regions of the image.

The dominating approach for representing the effects of location-dependent illumination changes, when *no* geometric transformation is involved (i.e., camera and object are fixed), is by a linear combination of basis images. The idea of representing illumination changes by a linear combination of images was proposed by Shashua [17]. Hallinan [18] represented illumination changes by a linear combination of basis images and used principal component analysis (PCA) to find an approximation of the basis images. In cases, where the shape of an object is convex and the object is Lambertian (the apparent brightness of the surface is the same regardless of the viewing angle), it was shown by Belhumeur and Kriegman [16] that the set of all images under arbitrary illuminations forms a convex cone. The cone can be constructed from as few as three images. Further study was done by Basri and Jacobs [19], who used spherical harmonics to show that the set of images produced by a Lambertian convex surface lies close to a 9-D space.

As previously mentioned, the difficulties associated with the joint geometric–radiometric estimation problem have led to the current state, where only a few attempts have been made to solve it. The lack of pointwise correspondence (due to the geometric transformation) and the lack of intensitywise alignment (due to the radiometric mapping) do not allow for a simple direct usage of the intensity information of the images. Seemingly, the geometric and radiometric problems are strongly coupled and may not be answered separately. As such, straightforward approaches for solving this problem typically lead to a high-dimensional nonlinear nonconvex optimization problem. Only a few works have *explicitly* modeled joint geometric–radiometric deformations. Indeed, among these, most evade the inherent nonlinearity of this estimation problem through linear approximation and/or variational optimization-based approaches [20], [21]. An explicit solution to joint geometric–radiometric estimation problem, where the radiometric change is a nonlinear mapping of the image intensities is given [22]. The assumed radiometric changes, however, are not location dependent. The estimation of location-dependent radiometric changes in the presence of an affine transformation is described in [23]. It requires several images of different spectral bands. Therefore, it is not suitable for grayscale images.

The goal of this paper is to fill the gap between the existing methods for estimating geometric changes only and those for estimating radiometric changes only. We propose a global and *linear* solution to the problem of estimating the illumination in an invariant manner to the affine geometric changes. Linear, affine invariant constraints on the illumination of an isolated object are derived using image moments. We show that using the derived constraints, the illumination is estimated in an invariant manner to the geometric transformation. The computa-

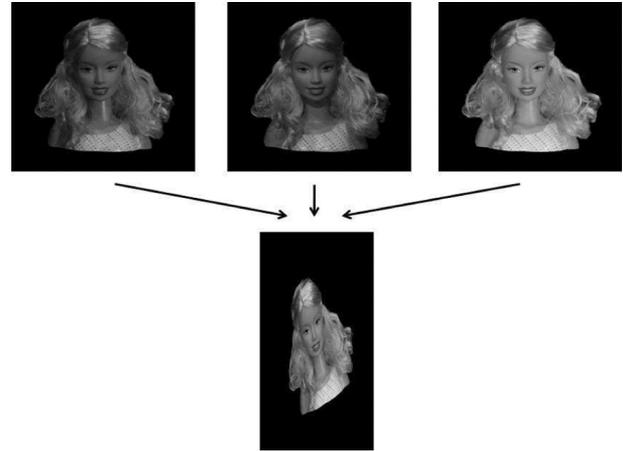


Fig. 1. Images of differently illuminated object and an affine-transformed linear combination thereof.

tional complexity of the method is low as it requires no more than solving a *linear* set of equations.

The paper is organized as follows. Section II defines the problem of jointly estimating the illumination and affine geometric changes. Section III elaborates on the search for linear and affine invariant transformations, and presents a solution to the problem. Section IV introduces linear solutions to the illumination estimation problem. Sections V–VII describe experiments that test the performance and robustness of the illumination estimation, and the effect of prior illumination estimation on the accuracy of registration.

## II. PROBLEM DEFINITION

This section introduces the mathematical model used to describe the problem of object registration in the presence of varying illumination.

We begin by defining the geometric estimation problem. Let  $f : R^2 \rightarrow R$  be an integrable image function with bounded support, and  $\phi = \mathbf{Ax} + \mathbf{b}$  an invertible affine geometric deformation that acts on  $f$  and produces an observation  $h$  such that  $h = f \circ \phi = f(\mathbf{Ax} + \mathbf{b})$ . Thus, in the absence of illumination variation, the registration problem is formulated as follows: given the functions (images)  $h$  and  $f$ , such that  $h = f \circ \phi$ , find  $\phi$ .

The radiometric changes (when both the camera and object are fixed) are modeled by linear combinations of a number of basis images. Let  $f_i$ ,  $i = 1, \dots, n$  be a set of basis functions, then  $f = \sum_{i=1}^n a_i f_i$  describes the image of an object viewed at a certain illumination state.

The combined model of an affine geometric transformation and a spatially varying radiometric deformation discussed in this paper is therefore defined by

$$h = \left( \sum_{i=1}^n a_i f_i \right) \circ \phi. \quad (1)$$

Thus, given only the observation  $h$  and the basis images  $f_i$ ,  $i = 1, \dots, n$ , we wish to find the coefficients  $a_i$  and the geometric transformation  $\phi$ . Fig. 1 displays an affine-transformed mixture of differently illuminated images of the object.

The same notations  $\phi$ ,  $f$ , and  $f_i$  are used throughout this paper to describe the affine transformation, the geometrically undeformed illuminated image, and the model illumination basis functions, respectively.

### III. SEARCH FOR AN AFFINE INVARIANT LINEAR TRANSFORM

Searching for a linear method to estimate the illumination coefficients  $a_i$  in (1) leads to a search for a transformation that is both invariant to the geometric changes and linear in the parameters of the illumination model. A linear transformation can be written as a kernel operator

$$T(f) = \int_{R^2} k(\mathbf{x})f(\mathbf{x})d\mathbf{x} \quad (2)$$

where  $f, k \in L_2$  (functions with finite energy). The affine invariance requirement is formulated as follows:

$$T(f \circ \phi) = T(f). \quad (3)$$

However, as the following theorem shows, a transformation that is both linear and affine invariant does not exist. The only such transformation is the null transformation ( $T(f) = 0$ ).

*Theorem 1:* A transformation  $T$  that obeys both (2) and (3) is the null transformation.

*Proof:* Since  $T(f \circ \phi) = T(f)$  then

$$\begin{aligned} T(f) &= \int_{R^2} k(\mathbf{x})f(\mathbf{x})d\mathbf{x} \\ &= \int_{R^2} k(\mathbf{x})f(\mathbf{A}\mathbf{x} + \mathbf{b})d\mathbf{x}. \end{aligned}$$

A change of variables  $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$  leads to

$$T(f) = \frac{1}{|\det(\mathbf{A})|} \int_{R^2} k(\mathbf{A}^{-1}(\mathbf{y} - \mathbf{b}))f(\mathbf{y})d\mathbf{y}.$$

By the Cauchy–Schwarz inequality

$$\begin{aligned} |T(f)|^2 &\leq \frac{1}{|\det(\mathbf{A})|^2} \int_{R^2} k^2(\mathbf{A}^{-1}(\mathbf{y} - \mathbf{b}))d\mathbf{y} \\ &\quad \times \int_{R^2} f^2(\mathbf{y})d\mathbf{y} \\ &= \frac{\|f\|_{L_2}}{|\det(\mathbf{A})|^2} \int_{R^2} k^2(\mathbf{A}^{-1}(\mathbf{y} - \mathbf{b}))d\mathbf{y}. \end{aligned}$$

Changing the variables again and letting  $\mathbf{x} = \mathbf{A}^{-1}(\mathbf{y} - \mathbf{b})$  leads to

$$\begin{aligned} |T(f)|^2 &\leq \frac{\|f\|_{L_2}}{|\det(\mathbf{A})|^2} \int_{R^2} |\det(\mathbf{A})|k^2(\mathbf{x})d\mathbf{x} \\ &= \frac{\|f\|_{L_2}\|k\|_{L_2}}{|\det(\mathbf{A})|}. \end{aligned}$$

Since  $T$  should be invariant to any  $\mathbf{A} \in Gl(2)$ , then  $|\det(\mathbf{A})|$  can be arbitrary large and

$$T(f) = 0.$$

Clearly, either the affine invariance or the linearity requirements needs to be relaxed. One possible approach, which is

taken here restricts the transformation to be linear only for images sharing the same support. Let  $\Omega_f \subset R^2$  be a bounded region calculated only from the support of  $f$ . Also let  $1_{\Omega_f}(\mathbf{x})$  be an indicator function for  $\Omega_f$

$$1_{\Omega_f}(\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{x} \in \Omega_f \\ 0, & \text{else.} \end{cases}$$

We define the functional

$$T(1_{\Omega_f}, f) = \frac{1}{\|1_{\Omega_f}\|_{L_1}} \int_{R^2} k(1_{\Omega_f}, \mathbf{x})f(\mathbf{x})d\mathbf{x} \quad (4)$$

where  $k$  is a kernel function that is chosen so that the affine invariance requirement is fulfilled, as shown in the following. Since  $\Omega_f$  depends only on the image support, the functional is linear for images (functions) sharing the same support.

Let  $h$  be an affine transformation of  $f$ , ( $h = f \circ \phi$ ). Also let  $\Omega_h$  be the corresponding region to  $\Omega_f$  under the affine transformation ( $\Omega_h = \{\mathbf{x}|\mathbf{A}^{-1}(\mathbf{x} - \mathbf{b}) \in \Omega_f\}$ ). The desired geometric invariance of the functional (4) is therefore formulated as follows:

$$T(1_{\Omega_f}, f) = T(1_{\Omega_h}, h). \quad (5)$$

Applying the functional to  $h$  leads to

$$\begin{aligned} T(1_{\Omega_h}, h) &= T(1_{\Omega_f} \circ \phi, h) \\ &= \frac{1}{\|1_{\Omega_f} \circ \phi\|_{L_1}} \int_{R^2} k(1_{\Omega_f} \circ \phi, \mathbf{x})f(\mathbf{A}\mathbf{x} + \mathbf{b})d\mathbf{x}. \end{aligned}$$

A change in the integration variable  $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$ ,  $d\mathbf{y} = |\det(\mathbf{A})|d\mathbf{x}$  leads to

$$\begin{aligned} T(1_{\Omega_h}, h) &= \frac{1}{\|1_{\Omega_f} \circ \phi\|_{L_1}|\det(\mathbf{A})|} \\ &\quad \times \int_{R^2} k(1_{\Omega_f} \circ \phi, \phi^{-1}(\mathbf{y}))f(\mathbf{y})d\mathbf{y}. \end{aligned}$$

Since the area of  $\Omega_f$  changes as the Jacobian of the affine transformation, we have that  $\|1_{\Omega_f} \circ \phi\|_{L_1}|\det(\mathbf{A})| = \|1_{\Omega_f}\|_{L_1}$ , therefore

$$T(1_{\Omega_h}, h) = \frac{1}{\|1_{\Omega_f}\|_{L_1}} \int_{R^2} k(1_{\Omega_f} \circ \phi, \phi^{-1}(\mathbf{y}))f(\mathbf{y})d\mathbf{y}. \quad (6)$$

Since the relation applies to every image  $f$ , the affine invariance requirement (5) on (4) and (6) translates to the following requirement on the kernel:

$$k(1_{\Omega_f}, \mathbf{x}) = k(1_{\Omega_f} \circ \phi, \phi^{-1}(\mathbf{x})) \quad (7)$$

#### A. Solution to the Kernel Equation

Each kernel that satisfies (7) provides a single linear affine invariant constraint on the illumination coefficients. We therefore look for parametric families of solutions to (7), leading to sets of linear constraints on the illumination coefficients. One possible solution to (7) is to choose

$$k(1_{\Omega_f}, \mathbf{x}) = 1_{\Omega_f}(\mathbf{x}) \quad (8)$$

which implies that (4) becomes the normalized sum ( $L_1$  norm) of the image intensities in the area defined by  $\Omega$ . Equation (7) holds for the kernel since

$$k(1_{\Omega_f} \circ \phi, \phi^{-1}(\mathbf{x})) = 1_{\Omega_f} \circ \phi \circ \phi^{-1}(\mathbf{x}) = 1_{\Omega_f}(\mathbf{x}). \quad (9)$$

Additional kernels may be derived by applying geometric transformations to  $\Omega_f$ . Such kernels can be expressed as a composition of functions

$$k(1_{\Omega_f}, \mathbf{x}) = 1_{\Omega_f} \circ \psi(\mathbf{x}) \quad (10)$$

where  $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is a geometric transformation.  $\psi$  may also depend on the properties of  $\Omega_f$ . Using the choice (10), (7) takes the form

$$1_{\Omega_f} \circ \psi(\mathbf{x}) = 1_{\Omega_f} \circ \phi \circ \psi \circ \phi^{-1}(\mathbf{x}). \quad (11)$$

Therefore, for (11) to hold for every  $\phi$  and  $\psi$ ,  $\phi$  and  $\psi$  are required to commute.

Let  $\Omega_f \subset \mathbb{R}^2$  be a region in an intensity image  $f$ . Its center of mass is defined as  $\mathbf{m}_{\Omega_f} = (m_{1\Omega_f}, m_{2\Omega_f})^T$ , where

$$\mathbf{m}_{\Omega_f} = \frac{1}{\|1_{\Omega_f}\|_{L_1}} \int_{\mathbb{R}^2} \mathbf{x} \cdot 1_{\Omega_f}(\mathbf{x}) d\mathbf{x}. \quad (12)$$

Let  $\alpha$  be a scale factor  $\alpha \in (0, 1]$  and let

$$\psi_{\Omega_f, \alpha}(\mathbf{x}) = \frac{1}{\alpha}(\mathbf{x} - \mathbf{m}_{\Omega_f}) + \mathbf{m}_{\Omega_f} \quad (13)$$

be a scaling of the coordinates around  $\mathbf{m}_{\Omega_f}$ . Around the center of mass, the scaling operator  $\psi$  and the affine transformation  $\phi$  commute and therefore (11) holds. Using  $\psi$ , we can finally define the affine invariant kernel as follows:

$$k_\alpha(1_{\Omega_f}, \mathbf{x}) = 1_{\Omega_f} \circ \psi_{\Omega_f, \alpha}(\mathbf{x}) = 1_{\Omega_{f, \alpha}}(\mathbf{x}) \quad (14)$$

where  $\Omega_{f, \alpha} = \{\mathbf{x} \mid \psi_{\Omega_f, \alpha}(\mathbf{x}) \in \Omega_f\}$ .

We remain with the problem of choosing corresponding regions  $\Omega_f$  and  $\Omega_h$ . A trivial choice would be the support of the images. Such an implementation was presented in previous conference papers [24], [25]. In this paper, we use elliptic regions calculated from the supports of the regions. The elliptic regions are less sensitive to small changes in the support. In addition, as will be shown in the following, the convex shape of the ellipse allows fast calculation of the transformation.

Given an image  $f$  and an affine transformation of the image  $h$ , ellipses with the same first- and second-order moments as the corresponding images also satisfy the same affine model [4]. The elliptic region can be calculated from the support of  $f$  by taking  $\Omega_f = \{\mathbf{x} \mid M^{-1/2}\mathbf{x} < s\}$ , where  $M$  is the second-order moments matrix of the support and  $s$  is a parameter that determines the size of the ellipse. We therefore choose  $\Omega_f$  and  $\Omega_h$  to be corresponding ellipses with the same first- and second-order moments as the supports of  $f$  and  $h$ , respectively. Example of affine-transformed images and the corresponding regions is shown in Fig. 2.

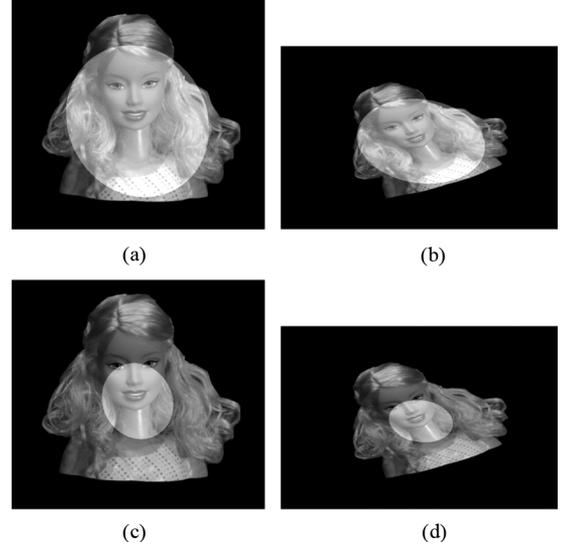


Fig. 2. Original image  $f$ , an affine transformation  $h$ , and the choice of  $\Omega$  with various scaling parameters. (a)  $\Omega_{f,0.8}$ . (b)  $\Omega_{h,0.8}$ . (c)  $\Omega_{f,0.4}$ . (d)  $\Omega_{h,0.4}$ .

#### IV. DERIVING LINEAR CONSTRAINTS FOR THE ILLUMINATION ESTIMATION

Let

$$\mu_0(f, \Omega_f, \alpha) = \frac{1}{\|1_{\Omega_f}\|_{L_1}} \int_{\mathbb{R}^2} f(\mathbf{x}) \cdot 1_{\Omega_{f, \alpha}}(\mathbf{x}) d\mathbf{x} \quad (15)$$

be the linear affine invariant transformation, as derived in Section III. Next, we show in the following that the affine invariance and linearity of the transformation lead to an explicit formulation of the illumination estimation problem using a linear set of equations.

Let  $h$  be an observation on an affine-transformed object, subject to an unknown illumination, as defined in (1). Notice that since  $f_i$  are differently illuminated observations of the object at the same pose, they share the same support. We define the ellipse calculated from their support as  $\Omega_f$ .

Due to the linearity and affine invariance of the transform derived in (4)–(7), we have that

$$\begin{aligned} \mu_0(h, \Omega_h, \alpha) &= \mu_0\left(\sum_{i=1}^n a_i \cdot f_i, \Omega_f, \alpha\right) \\ &= \sum_{i=1}^n a_i \cdot \mu_0(f_i, \Omega_f, \alpha). \end{aligned} \quad (16)$$

Thus, every value of  $\alpha$  yields a single linear constraint on the illumination coefficients  $a_i$ . Let  $\alpha_1, \dots, \alpha_m$  be a set of  $m$  different scaling parameters. Evaluating  $\mu_0$  for each of these scaling parameters results in the following set of linear equations in  $a_1, \dots, a_n$ :

$$\begin{pmatrix} T_0^{f_1}(\alpha_1) & \cdots & T_0^{f_n}(\alpha_1) \\ \vdots & \ddots & \vdots \\ T_0^{f_1}(\alpha_m) & \cdots & T_0^{f_n}(\alpha_m) \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} T_0^h(\alpha_1) \\ \vdots \\ T_0^h(\alpha_m) \end{pmatrix} \quad (17)$$

where  $T_0^h(\alpha) \triangleq \mu_0(h, \Omega_h, \alpha)$  and  $T_0^{f_i}(\alpha) \triangleq \mu_0(f_i, \Omega_f, \alpha)$ .

The illumination estimation problem is therefore reduced to the solution of a linear set of equations. The set of equations can only be singular in cases, where the intensity image does not change when increasing the size of the ellipse and the object is circular symmetric up to an affine transformation. However, such cases are isoteric.

In the absence of noise, using  $m = n$  equations provides a linear system with  $n$  equations and  $n$  unknowns that can be solved by matrix inversion. In the presence of noise, using  $m > n$  equations leads to a linear regression problem. Once  $a_1, \dots, a_n$  are known, the illumination changes can be compensated for and the problem is reduced to a *strictly geometric* form  $h = f \circ \phi$ , where  $f$  is known and is given by  $f = \sum_{i=1}^n a_i f_i$ . The remaining, strictly geometric, problem can then be solved by any global (eg., [12] and [24]) or local-feature-based (eg., [5] and [6]) affine registration method.

#### A. Linear Estimation Framework for the Illumination Coefficients

The quantities in (17) are subject to noise due to quantization, interpolation errors, and model mismatches. In addition, the illumination coefficients that represent the object's illumination share a joint statistics that represents the probable illuminations of the object. Hence, we adopt a Bayesian framework to derive a linear estimator for the illumination coefficients. Rewriting the noisy version of (17) in a matrix form, we have

$$\mathbf{B} \cdot \mathbf{a} = \mathbf{r} + \mathbf{w} \quad (18)$$

where  $\mathbf{B}$  and  $\mathbf{r}$  are the products of the transformation on the basis images and the observation,  $\mathbf{a}$  is a vector of the illumination coefficients, and  $\mathbf{w}$  is the "error" vector. Assuming that the noise is uncorrelated with the illumination coefficients, by [26], the best linear estimator for the illumination coefficients is given by

$$\hat{\mathbf{a}} = E[\mathbf{a}] + (\mathbf{C}_a^{-1} + \mathbf{B}^T \mathbf{C}_w \mathbf{B})^{-1} \cdot (\mathbf{B}^T \mathbf{C}_w) \cdot (\mathbf{r} - \mathbf{B} E[\mathbf{a}]) \quad (19)$$

where  $C_a$  and  $C_w$  are the covariance matrices of the illumination coefficients  $\mathbf{a}$  and of the noise term  $\mathbf{w}$ , respectively. The statistics of  $\mathbf{a}$  are learned from the variation of the illumination coefficients over a training set, as explained in the experiments in the following. The statistics of  $\mathbf{w}$  are learned by performing random affine transformations on the training set and measuring the difference  $\mathbf{r} - \mathbf{B}\mathbf{a}$ .

#### B. Implementation and Runtime of the Proposed Solution

The quantities described earlier can be calculated rapidly on a digital computer. Given a discrete image  $f(u, v)$ , discretization of (15) leads to the form

$$\mu_0(f, \Omega_f, \alpha) = \frac{1}{\|1_{\Omega_f}\|_{L_1}} \sum_{u,v} f(u, v) \cdot 1_{\Omega_f, \alpha}(u, v).$$

Assume that we use  $N$  scaling parameters  $\alpha$ . It seems that we need to perform  $N$  summations of the image. However, we can calculate the sums for all  $\alpha$  values in a single pass on the image.

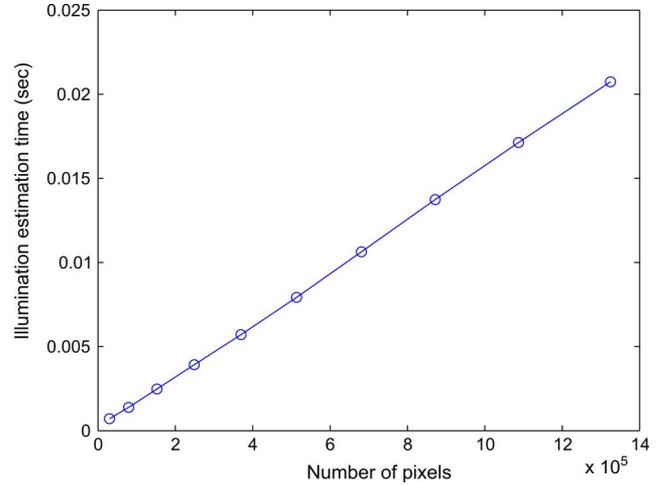


Fig. 3. Illumination estimation time as a function of the number of image pixels.

We set a fixed scaling increment  $\alpha_k = k/N$ . Let  $\Delta_k$  be the summation of  $f$  on  $\Omega_{f, \alpha_k} / \Omega_{f, \alpha_{k-1}}$ . The following is performed for each pixel  $(u, v)$ :

- 1) Calculate  $d$ , the minimum scaling of the ellipse containing  $(u, v)$ . Therefore,  $k = \lceil Nd \rceil$  is the smallest ellipse that contains  $(u, v)$ .
- 2) Increment  $\Delta_k$  by the value of the pixel- $f(u, v)$ .

Finally, the required quantities are obtained by  $\mu_0(f, \Omega_f, \alpha_k) = \sum_{l=1}^k \Delta_l$ .

The resulting implementation is linear time  $O(n)$ , where  $n$  is the number of pixels in the image and does not depend on the number of scaling parameters. Runtime results of full illumination estimation process are displayed in Fig. 3. The results are an average of 100 runs. The algorithm was implemented as a MATLAB script on an Intel Core 2 Duo, 2.54 Ghz laptop. The summation phase is implemented as a MEX file.

In the following examples, 50 values of  $\alpha$  are used for calculating the linear constraints. The acquired images are 8-b grayscale images with values in the range of 0–255.

#### V. EXPERIMENTS—ILLUMINATION ESTIMATION

In the following experiments, we test the performance of the proposed linear solution for the estimation of object's illumination in the presence of combined affine geometric deformation and spatially varying illumination changes. We use an array of 14 light sources to illuminate an object from 14 directions, as illustrated in Fig. 4.

A set of 75 training images taken at a fixed pose, using a random illumination power from each light source, is used to learn the possible illuminations of a doll head. A PCA [27], performed on the differently illuminated images, shows that 98% of the variation in the illumination can be represented by the mean image and the first five basis images. We note the mean image as  $f_m$  and the basis functions as  $f_1, \dots, f_5$ .

An additional set of 74 images illuminated randomly with the 14 light sources is used as a test set. Pseudorandom affine transformations are applied to the test images to produce 296 observations. The transformations consist of a uniformly distributed scale factor of 1 to 2, a rotation of  $0^\circ$ – $360^\circ$  and off-diagonal

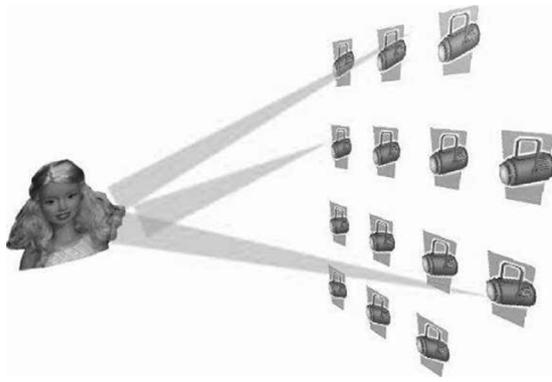


Fig. 4. Illustration of the illumination setting.

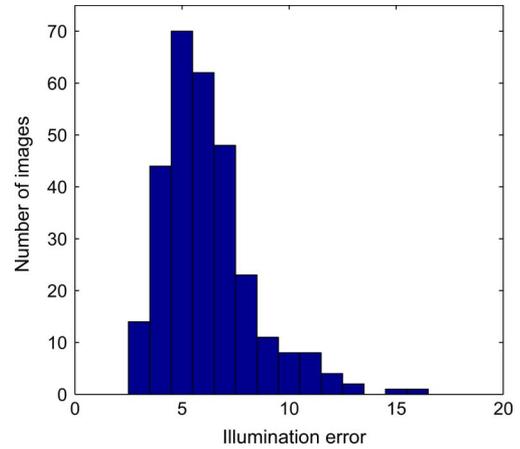


Fig. 6. Illumination estimation error between the unknown template and the estimated template from an affine-transformed observation.

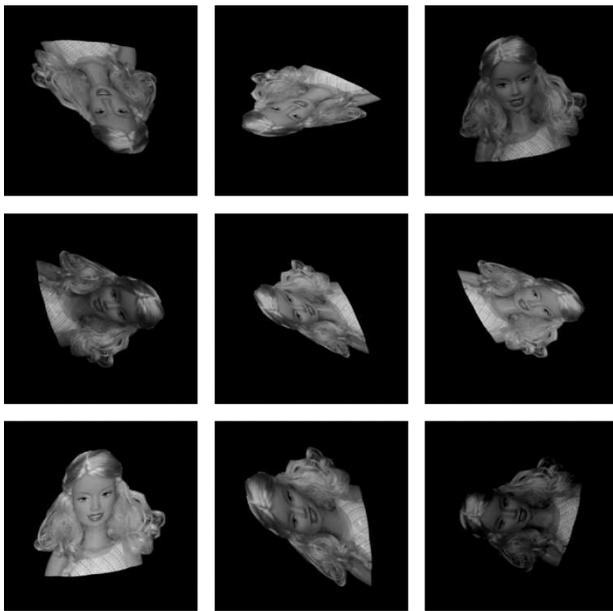


Fig. 5. Affine-transformed images with varying illumination from the test images.

skew terms of  $-0.3$  to  $0.3$ . Fig. 5 displays example images from the test set.

*A. Illumination Changes Compensation in the Presence of a Geometric Affine Transformation*

The experiment described here tests the accuracy of the illumination estimation in the presence of an affine transformation. Given only the affine-transformed image  $h = f \circ \phi$ , we wish to estimate the illumination coefficients  $a_1, \dots, a_5$  so that  $\hat{f} = f_m + \sum_{i=1}^5 a_i f_i$  will be similar to  $f$ . The distance measure of the illumination error is the square root of the mean square error between  $f$  and  $\hat{f}$ .

A histogram of the illumination estimation error is shown in Fig. 6. An example of the distance measure between different observations is shown in Fig. 7. The results show that illumination compensation using the proposed method is capable of estimating the change in illumination, and produces a template image that differs from the observation by the geometric change and low noise only. The average standard deviation on all images is 6.26 gray levels.

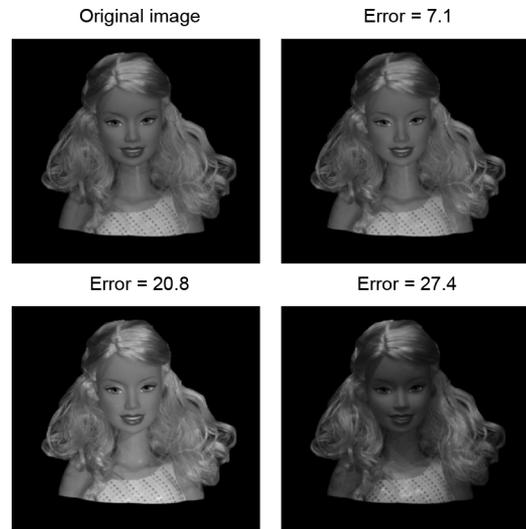


Fig. 7. Example—Illumination error between different observations.

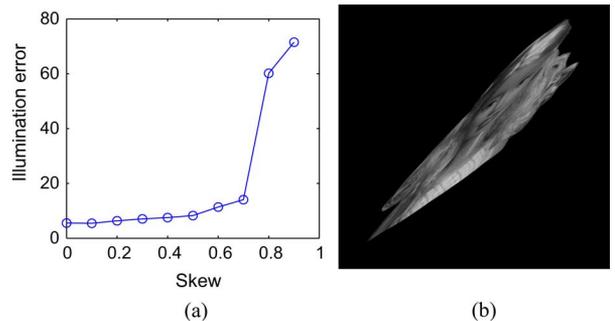


Fig. 8. Robustness to extreme affine transformations (a) Illumination estimation error with increasing skew. (b) Affine-transformed image with 0.7 skew.

*B. Robustness of the Illumination Estimation to Extreme Deformations*

Although the theoretical analysis shows that the deformation is invariant to affine deformations, quantization and interpolations error are not taken into account. We therefore test the amount of distortion that the method can handle. We measure the growth of the average illumination estimation error as a function of the skew (the relation between the off-diagonal

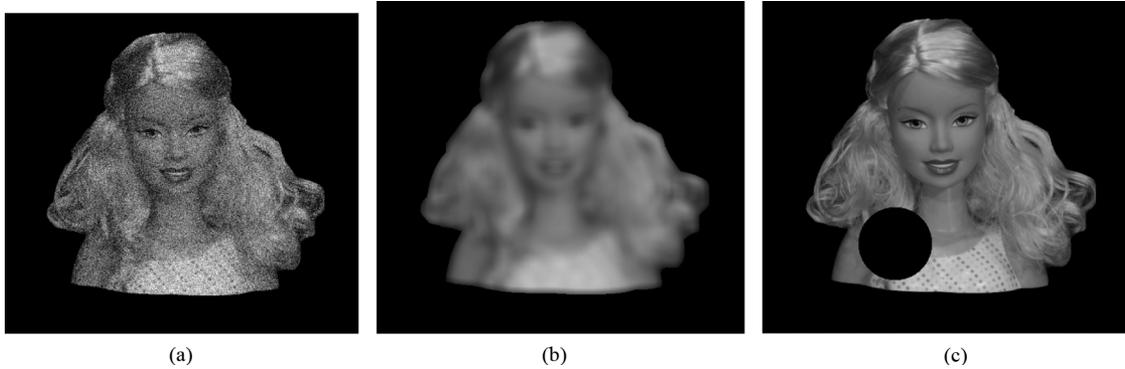


Fig. 9. Observations with noise, blur, and occlusion. (a) Added noise with standard deviation of 40 gray levels. (b) Blurred image, Gaussian kernel with standard deviation of 20. (c) Occlusion of 8%.

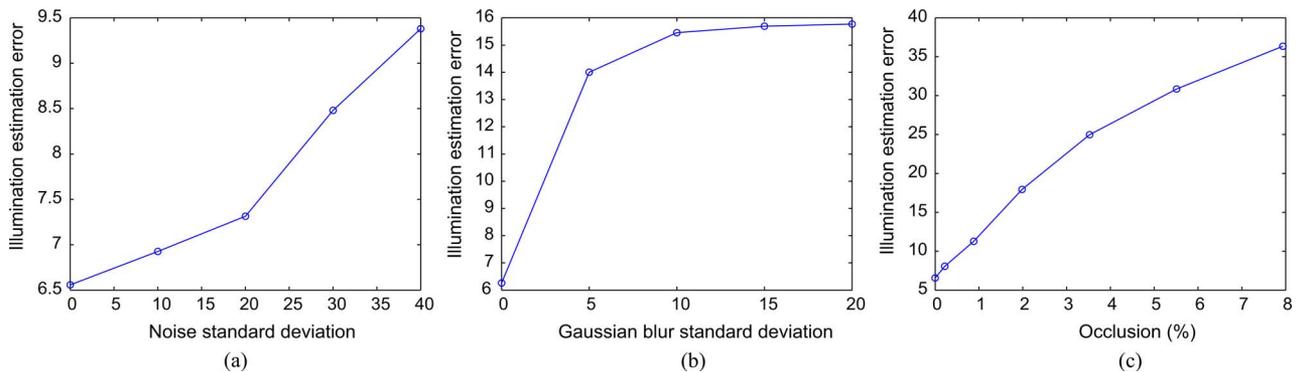


Fig. 10. Average illumination estimation error: with noise, blur, and occlusion. (a) Registration with additive noise. (b) Illumination estimation with blur. (c) Registration with occlusion.

terms and the diagonal terms) of the affine transformation employed. The matrix becomes singular when the skew is 1. The results are shown in Fig. 8. The method is stable for a skew factor of up to 0.7.

### C. Effect of Noise Blur and Occlusion on the Illumination Estimation

In this set of experiments, we test the effects of noise, blur, and occlusion on the registration result. We conduct the same illumination estimation experiment, as in Section V-A, with the exception that the observation  $h$  is replaced by an image  $\tilde{h}$  that reflects the phenomenon. The measured score is, as in the previous experiment, the standard deviation of the pixels difference image  $f - \hat{f}$  (notice that the unknown template  $f$  remains unchanged and only the affine-transformed observation is degraded).

1) *Effect of Noise*: To simulate the added noise, the observation  $h$  is added with a white Gaussian noise  $N$  so that  $\tilde{h} = h + N$ . Fig. 10(a) displays the average illumination estimation error as a function of the standard deviation of the noise. Fig. 9(a) displays an image with the maximal noise level that was tested in this experiment (added noise with standard deviation of 40 gray levels). The results demonstrate high robustness of the illumination estimation process to noise.

2) *Effect of Blur*: To simulate blur, we use a Gaussian convolution kernel  $K$  so that  $\tilde{h} = h * K$ . Fig. 10(b) displays the average illumination estimation error as a function of the standard

deviation (width) of the Gaussian kernel. Fig. 9(b) displays an image with the maximal blur. The results show that it is still possible to compensate for most of the illumination changes when the observation is highly blurred. However, the quality of the estimation degrades from the nonblurred case.

3) *Effect of Partial Occlusion*: An occlusion of the image is simulated by a black circle covering parts of the image. The location of the circle is selected randomly inside the object, as shown in Fig. 9(c). Fig. 10(c) displays the average illumination estimation error as a function of the percentage of the occluded image area. The method can handle small occlusions of the image; however, it fails for large occlusions since the employed operator is based on integration on the entire image.

### D. Illumination Estimation With a Small Training Set

In many applications, a large training set for illumination estimation does not exist. However, we show here that illumination estimation can be performed using a small number of images. We manually chose four images from the original training set. In each of the images, the light comes from different direction: frontal, right, left, and upward illumination. We trained the illumination estimator using only the 4 images and tested the illumination estimation on the test set of 296 images.

The estimation results and the four images are shown in Fig. 11. The average error is 10.1 gray levels. The results show that although there is a small degradation in the performance, the method can still compensate for most of the illumination

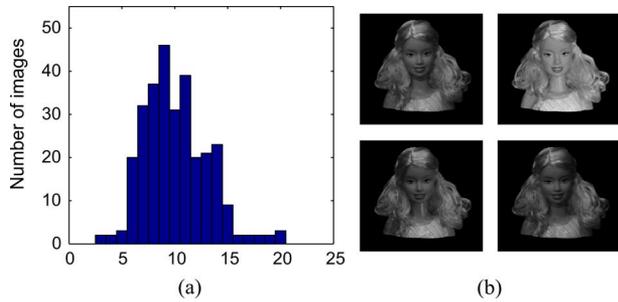


Fig. 11. Estimation with four images as a training set. (a) Illumination estimation error. (b) Training set.

TABLE I  
AVERAGE REGISTRATION ERROR (PIXELS) WITH AND WITHOUT PRIOR ILLUMINATION ESTIMATION

SIFT	2.61
SIFT + Illumination	1.37
MSER	2.35
MSER + Illumination	1.31

changes, even when only a very restricted set of training images is available.

## VI. MSER- AND SIFT-BASED REGISTRATION WITH ILLUMINATION CHANGES COMPENSATION

In the following experiment, we test the effect of decoupling the problems of estimating the geometric and radiometric transformations on the performance of commonly used methods for estimating geometric deformations. The experiment described here uses the same basis images and test images as in Section V.

Two experiments to estimate the affine transformation between an observation  $h = f \circ \phi$  and the (unknown) template  $f$  are conducted. The first experiment is performed without illumination estimation using the mean image  $f_m$  as a template. In the second experiment, the estimation of the affine transformation employs the proposed affine invariant illumination compensation method and performs the registration using  $\hat{f} = f_m + \sum_{i=1}^5 a_i f_i$  as a template.

In order to emphasize the significant gain in decoupling the problems of estimating the geometric and radiometric transformations and in employing the proposed method for affine invariant illumination estimation, estimation of the geometric deformation is performed using two local-feature-based affine registration methods: SIFT [5] + random sample consensus RANSAC [28], and MSER [6] + RANSAC. In theory, local-feature-based methods gain the lowest benefit from estimating the radiometric deformation as they normalize the illumination to local changes in a generic way. The accuracy of the registration is measured in pixels as an average distance from the location of each pixel, as determined by the known simulated affine transformation, to the estimated location. The average error of each method with and without illumination compensation is described in Table I. The average registration time (feature extraction + matching) using VLFeat [29] is about 0.8 s, whereas the average illumination estimation time using the proposed method is 0.012 s. Therefore, the prior illumination estimation did not result in a noticeable runtime increase.

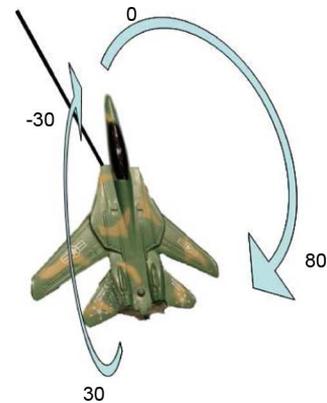


Fig. 12. Tilt and rotation of the plane.

The results show that a prior illumination compensation improves the registration accuracy by a factor of 2 and that even though local-feature-based methods normalize the illumination to local changes in a generic way, estimation of the illumination model considerably improves the accuracy of the registration.

## VII. LAB EXPERIMENT—REGISTRATION OF A MOVING OBJECT WITH PRIOR ILLUMINATION ESTIMATION

The final experiment tests the robustness of the illumination estimation process, as many of the assumptions made in the derivation are violated by the experimental setting. Since the object is not flat, the tilt of the object hides and unveils parts of the object to the camera. The finite distance between the object and the camera adds a perspective nature to the geometric transformation. The illumination model is also more complex since the object is not Lambertian.

The goal of the experiment is to register images of the moving object in the presence of varying illumination. A glossy toy airplane is connected to a two-axis motorized stage for rotation and tilt. We wish to register an airplane image taken without rotation and tilt to images taken at varying rotation, tilt, and illumination conditions.

### A. Experiment

We use the same illumination settings as in Section V. The training set consists of 200 images taken without rotation and tilt. PCA analysis of the training set shows that ten images represent 97.2% of the illumination variation of the object. The increase in the number of basis images from the previous experiment is caused by the non-Lambertian nature of the plastic object.

The test data set is obtained by rotating the airplane in angles of  $0^\circ$ – $80^\circ$  in increments of  $20^\circ$ , while the tilt of the plane changes in the range  $-30^\circ$  to  $30^\circ$  in increments of  $10^\circ$ . Ten images are taken at each pose of the airplane. The same illumination estimation process, as in Section V, is applied to produce the estimated template that differs from the observation by a geometric change only. We test the improvement to the registration accuracy of SIFT and MSER gained by performing a prior illumination compensation. A ground-truth homography is obtained by manually selecting very localized patterns on each pose and estimating the homography from the selected points.

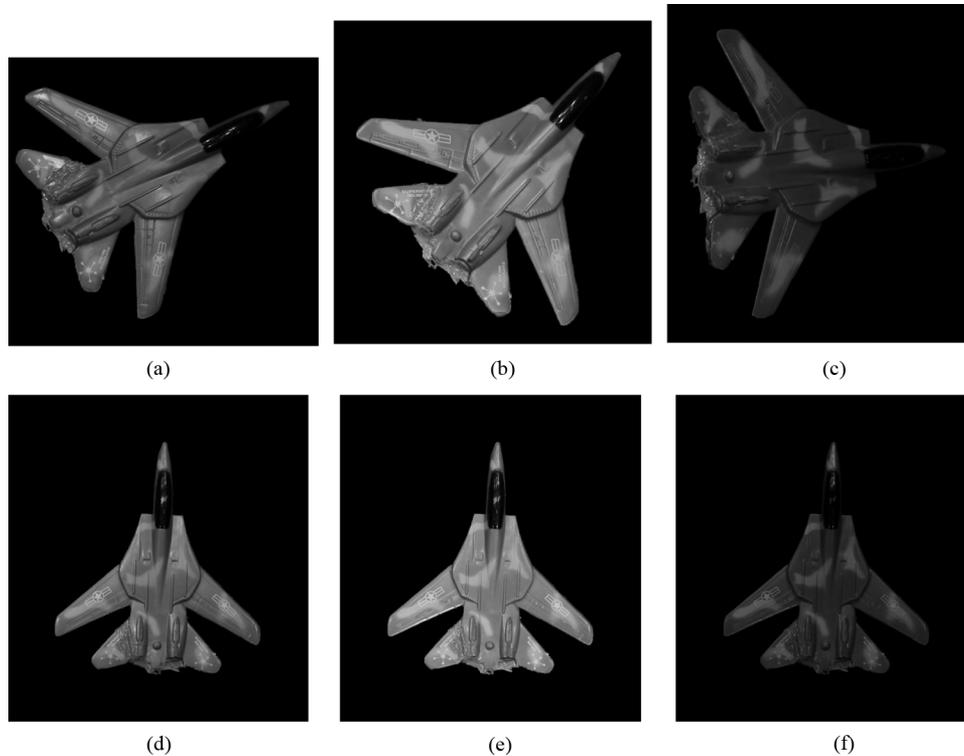


Fig. 13. Registration with and without illumination estimation. An observation and a template image compensated for the illumination change using the estimated illumination model are displayed in each column. The registration error in pixels using MSER and SIFT is given below each image in the form (registration error without illumination estimation/registration error using illumination estimation). (a) Observation—rotation 60, tilt 10. (b) Observation—rotation 40, tilt 20. (c) Observation—rotation 80, tilt 30. (d) Template with estimated illumination. MSER (6.79/4.75) and SIFT (3.62/2.93). (e) Template with estimated illumination. MSER (6.31/3.89) and SIFT (10.86/5.65). (f) Template with estimated illumination. MSER (12.87/4.30) and SIFT (13.38/5.11).

TABLE II  
REGISTRATION RESULTS ON THE AIRPLANE SET

Method	Accuracy	Number of failed registration
SIFT	5.31	14
SIFT + Illumination	4.54	0
MSER	5.45	6
MSER + Illumination	4.44	3

### B. Results

Examples of the observations and estimated templates in various poses is displayed in Fig. 13. In each column, the observation  $h$  is shown on the top row, while the estimated illumination-compensated template  $\hat{f}$  is shown below. Ideally, the difference between the two images is the geometric deformation alone. Table II displays the average error in pixels using MSER and SIFT, with and without prior illumination estimation. The proposed method yields a smaller improvement than in the case of the doll head; however, the prior illumination estimation significantly reduces the number of failed registrations (a registration with an error of more than 200 pixels).

## VIII. CONCLUSION

This paper presents a novel approach for handling a complex radiometric–geometric deformation model in registering two observations on the same object. More specifically, the observations differ due to a combined effect of an affine geometric transformation and a radiometric transformation that is location dependent. In this framework, we derive a method for obtaining

linear and affine invariant constraints for estimating the illumination model. Thus, the joint problem of illumination estimation and geometric transformation estimation is decoupled by this operation into two separate phases. On the first phase, the illumination model is estimated in an invariant manner to the geometric transformation. The problem is then reduced to a strictly geometric problem, which can be solved by any conventional affine registration method.

The proposed affine invariant illumination estimation method is shown to improve both the accuracy and the failure rate of local-feature-based registration methods. In cases, where the template and the observation are related by an affine geometric transformation, the accuracy of the registration improved by a factor of 2. The prior illumination estimation procedure also improved the failure rate of the feature-based registration by a factor of more than 2.

The global nature of the transform and the fact that all the statistics required in the solution are computed by integration on areas in the image provide robustness such that the illumination estimation process is highly insensitive to noise. The proposed method is very fast as it requires only summation of nested regions in the image and solving a linear set of equations.

## REFERENCES

- [1] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," *ACM Comput. Surv.*, vol. 38, Dec. 2006.
- [2] R. Haralick and J. Hyonam, "2d-3d pose estimation," in *Proc. 9th Int. Conf. Pattern Recognit.*, Nov. 1988, vol. 1, pp. 385–391.
- [3] R. Szeliski, "Image alignment and stitching: a tutorial," *Found. Trends. Comput. Graph. Vis.*, vol. 2, no. 1, pp. 1–104, Jan. 2006.

- [4] T. Tuytelaars and L. V. Gool, "Wide baseline stereo matching based on local, affinely invariant regions," in *Proc. Brit. Mach. Vis. Conf.*, 2000, pp. 412–425.
- [5] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, pp. 91–110, 2004.
- [6] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide baseline stereo from maximally stable extremal regions," in *Brit. Mach. Vis. Conf.*, 2002, vol. 1, pp. 384–393.
- [7] E. Rahtu, M. Salo, and J. Heikkilä, "Affine invariant pattern recognition using multiscale autoconvolution," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 6, pp. 908–918, Jun. 2005.
- [8] M. K. Hu, "Visual pattern recognition by moment invariants," *IEEE Trans. Inf. Theory*, vol. IT-8, no. 2, pp. 179–187, Feb. 1962.
- [9] Z. Yang and F. S. Cohen, "Cross-weighted moments and affine invariants for image registration and matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 8, pp. 804–814, Aug. 1999.
- [10] M. Petrou and A. Kadyrov, "Affine invariant features from the trace transform," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 1, pp. 30–44, Jan. 2004.
- [11] F. Mindru, T. Moons, and L. V. Gool, "Recognizing color patterns irrespective of viewpoint and illumination," in *Proc. Conf. Comput. Vis. Pattern Recognit.*, 1999, pp. 368–373.
- [12] J. M. Francos and R. Hagege, "Parametric estimation of affine transformations: An exact linear solution," *J. Math. Imag. Vis.*, vol. 37, pp. 1–16, 2010.
- [13] J. Heikkilä, "Pattern matching with affine moment descriptors," *Pattern Recognit.*, vol. 37, no. 9, pp. 1825–1834, 2004.
- [14] J. Flusser and T. Suk, *A moment-based approach to registration of images with affine geometric distortion*, vol. 32, no. 2, pp. 382–387, Mar. 1994.
- [15] C. Domokos and Z. Kato, "Parametric estimation of affine deformations of planar shapes," *Pattern Recognit.*, vol. 43, no. 3, pp. 569–578, 2010.
- [16] P. N. Belhumeur and D. J. Kriegman, "What is the set of images of an object under all possible lighting conditions," *Int. J. Comput. Vis.*, vol. 28, pp. 270–277, 1998.
- [17] A. Shashua, "On photometric issues in 3d visual recognition from a single 2d image," *Int. J. Comput. Vis.*, vol. 21, pp. 99–122, 1997.
- [18] P. Hallinan, "A low-dimensional representation of human faces for arbitrary lighting conditions," in *Proc. Conf. Comput. Vis. Pattern Recognit.*, 1994, pp. 995–999.
- [19] R. Basri and D. Jacobs, "Lambertian reflectance and linear subspaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 2, pp. 218–233, Feb. 2003.
- [20] A. Bartoli, "Groupwise geometric and photometric direct image registration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 12, pp. 2098–2108, Dec. 2008.
- [21] J. Jia and C.-K. Tang, "Image registration with global and local luminance alignment," presented at the 9th IEEE Int. Conf. Comput. Vis., Nice, France, 2003.
- [22] S. Z. Kovalsky, G. Cohen, R. Hagege, and J. M. Francos, "Decoupled linear estimation of affine geometric deformations and nonlinear intensity transformations of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 5, pp. 940–946, May 2010.
- [23] R. Frenkel and J. M. Francos, "Registration of geometric deformations in the presence of varying illumination," in *Proc. IEEE Int. Conf. Image Process.*, 2007, vol. 3, pp. 489–492.
- [24] K. Bentolila and J. M. Francos, "Joint affine and illumination estimation using scale manipulation features," in *Proc. 15th Int. Conf. Image Anal. Process.*, Berlin, Heidelberg, 2009, pp. 844–852.
- [25] K. Bentolila and J. M. Francos, "Object pose estimation in the presence of local illumination changes using scale manipulation transform," in *Proc. IEEE Int. Workshop Multimedia Signal Process.*, 2009, pp. 1–6.
- [26] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [27] I. T. Jolliffe, *Principal Component Analysis*, 2nd ed. New York: Springer-Verlag, 2002.
- [28] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981.
- [29] A. Vedaldi and B. Fulkerson, "VLFeat: An Open and Portable Library of Computer Vision Algorithms," in *Proc. Int. Conf. Multimedia*, Firenze, Italy, 2010, pp. 1469–1472.



**Jacob Bentolila** was born in Paris, France, on December 8, 1978. He received the B.Sc. (summa cum laude) degree in electrical engineering from the Technion-Israel Institute of Technology, Haifa, Israel, in 2007. He is currently working toward the Ph.D. degree at the Electrical and Computer Engineering Department, Ben-Gurion University, Beer Sheva, Israel.

His current research interests include computer vision, object recognition, and image processing.



**Joseph M. Francos** (S'89–M'90–SM'97) received the B.Sc. degree in computer engineering and the D.Sc. degree in electrical engineering from the Technion-Israel Institute of Technology, Haifa, Israel, in 1982 and 1991, respectively.

Since 1993, he has been with the Department of Electrical and Computer Engineering, Ben-Gurion University, Beer-Sheva, Israel, where he currently is a Professor, and also heads the Mathematical Imaging Group and the Signal Processing track. From 1982 to 1987, he was with the Signal Corps

Research Laboratories, Israeli Defense Forces. From 1991 to 1992, he was a Visiting Assistant Professor in the Department of Electrical Computer and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY. During 1993, he was with Signal Processing Technology, Palo Alto, CA. 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 include image registration, estimation of object deformations from images, parametric modeling and estimation of 2-D random fields, random fields theory, and texture analysis and synthesis.

Dr. Francos was an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1999 to 2001.