**Noname manuscript No.**
(will be inserted by the editor)

# Homography and Fundamental Matrix Estimation Using an Affine Error Metric

**Jacob Bentolila · Joseph M. Francos**

Received: date / Accepted: date

**Abstract** Matching a pair of affine invariant regions between images results in estimation of the affine transformation between the regions. However, the parameters of the affine transformations are rarely used directly for matching images, mainly due to the lack of an appropriate error metric of the distance between them.

In this paper we derive a novel metric for measuring the distance between affine transformations: Given an image region, we show that minimization of this metric is equivalent to the minimization of the mean squared distance between affine transformations of a point, sampled uniformly on the image region. Moreover, the metric of the distance between affine transformations is equivalent to the $l_2$ norm of a linear transformation of the difference between the six parameters of the affine transformations. We employ the metric for estimating homographies and for estimating the fundamental matrix between images. We show that both homography estimation and fundamental matrix estimation methods, based on the proposed metric, are superior to current linear estimation methods as they provide better accuracy without increasing the computational complexity.

Jacob Bentolila
Ben Gurion University
Electrical & Computer Engineering Department
P.O.B 653 Beer-Sheva 84105, Israel Tel.: +972-8-6461518
Fax: +972-8-6472949
E-mail: bentolil@ee.bgu.ac.il

Joseph M. Francos
Ben Gurion University
Electrical & Computer Engineering Department
P.O.B 653 Beer-Sheva 84105, Israel Tel.: +972-8-6461518
Fax: +972-8-6472949
E-mail: francos@ee.bgu.ac.il

**Keywords** Epipolar geometry · Affine invariant regions · Fundamental matrix · Homographies

## 1 Introduction

Local image analysis through the investigation of the relations between collections of small areas has gained popularity in many computer vision tasks such as: Object recognition, Camera localization, Automatic navigation and 3D modeling. It offers robustness to changes in geometric appearance (changes in point of view / in the object) by employing the detected area for geometric normalization. The degree of normalization varies from a simple invariance to translation [16] to invariance to rotation and scale [11] or a more general affine normalization of the image patches.

Over the last 10 years, many affine invariant feature detectors have been proposed (such as [12],[13],[22]). A review of local feature detectors is given in [23]. In contrast to point wise matches (such as SIFT [11], and Harris corner detector [6]) that provide a match between a pair of image coordinates, the matching of affine invariant features also provides an approximation of the local derivatives of the deformation between the images at the point of correspondence: $A = \frac{d\boldsymbol{\mu}'}{d\boldsymbol{\mu}}$ [15]. We denote such a correspondence of points along with the derivatives of the deformation between the images at the point of correspondence by the name *affine correspondence*.

Although region matching methods based on affine normalization of the regions exist for a long time, the local affine approximation of the deformation between the regions is rarely used for estimation of a global relation between the images (such as an homography or a fundamental matrix). One of the major reasons that the

parameters of the affine transformation are not directly used for estimation of the relation between images is the lack of an appropriate error metric that provides a geometric interpretation to the distance between the parameters of the affine transformations. The aim of this paper is to derive such an error metric and to employ the metric for calculating global relations between images.

Some methods employ the local affine approximation for propagating the affine relation by searching the surrounding of existing matches for new matches [5],[10],[2]. Propagation methods result in a high number of affine relations between the images. However, the parameters of the affine transformations are not used for estimation of a global relation between the images.

In the absence of an appropriate metric between affine transformation, methods that do employ information from affine correspondences for calculation of a global model try to convert the affine correspondence into point correspondences. Generally, since an affine transformation is determined by 3 points, it seems that it is roughly equivalent to the matching of 3 points between images. In [3],[18],[19],[17] local affine approximation is used for generation of additional points. However, affine correspondences and point correspondences are different mathematical entities. Furthermore, there is no direct relation between the error in matching the simulated points to the average geometric error between the matched regions.

In this paper we propose an error metric for the distance between affine transformations. The error metric measures the mean squared distance between points in a region, subjected to two different affine transformations. We show that the error metric is equivalent to measuring the $l_2$ norm between linear transformations of the 6 parameters of the affine transformations and develop both linear solutions that minimize the algebraic error, and iterative solutions for exact minimization of the proposed metric. We further show that minimization of the error metric produces better results in the estimation of homographies and fundamental matrices [4] between images than points based methods, without increasing the computational load.

The structure of the paper is as follows: In Section 2 we derive a method for measuring the geometric error between affine transformations. Section 3 formulates the problem of finding a global transformation that minimizes the mean squared error between the measured regions and regions projected by the transformation. Linear solutions for calculation of homography and the fundamental matrix based on the affine error metric are described in Section 4 and non linear solutions that minimize the affine error metric are described in Section 5. Finally, a comparison with point based methods that employ the affine relation is presented in Section 6.

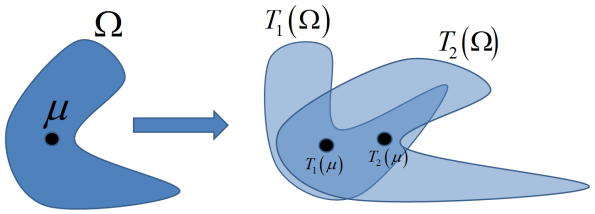


**Fig. 1** A region $\Omega$ subjected to two affine transformations $T_1$ and $T_2$.

## 2 Affine Error Metric

In this section we develop a metric measuring the average distance between points that undergo two different affine transformation.

Let $\mathbf{p} \in \mathbb{R}^2$ be a random point, sampled from a uniform distribution on some image region $\Omega$. We wish to measure the mean squared distance between the images of $\mathbf{p}$ under two different affine transformations $T_1$ and $T_2$. The configuration is illustrated in Figure 1.

Let $\boldsymbol{\mu}$ be the mean of $\mathbf{p}$ and $\boldsymbol{\Sigma}$ the covariance matrix of $\mathbf{p}$. Also let $T_1(\mathbf{p}) = \mathbf{A}_1\mathbf{p}+\mathbf{b}_1$ and $T_2(\mathbf{p}) = \mathbf{A}_2\mathbf{p}+\mathbf{b}_2$ be two affine transformations of $\mathbf{p}$. The vector between $T_1(\mathbf{p})$ and $T_2(\mathbf{p})$ is therefore defined as

$$\begin{aligned}\boldsymbol{\epsilon} &= T_1(\mathbf{p}) - T_2(\mathbf{p}) \\ &= \mathbf{A}_1(\mathbf{p}-\boldsymbol{\mu}) - \mathbf{A}_2(\mathbf{p}-\boldsymbol{\mu}) + T_1(\boldsymbol{\mu}) - T_2(\boldsymbol{\mu})\end{aligned} \tag{1}$$

As the two transformations are affine, the second order moments of $\boldsymbol{\epsilon}$ are linearly dependent on $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ and

$$\begin{aligned}E[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T] &= (\mathbf{A}_1-\mathbf{A}_2)\boldsymbol{\Sigma}(\mathbf{A}_1-\mathbf{A}_2)^T \\ &+ (T_1(\boldsymbol{\mu})-T_2(\boldsymbol{\mu}))(T_1(\boldsymbol{\mu})-T_2(\boldsymbol{\mu}))^T\end{aligned} \tag{2}$$

Let the Cholesky decomposition of $\boldsymbol{\Sigma}$ be $\boldsymbol{\Sigma} = \mathbf{D}\mathbf{D}^T$; then, the second order moments matrix of $\boldsymbol{\epsilon}$ takes the form

$$\begin{aligned}E[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T] &= (\mathbf{A}_1\mathbf{D}-\mathbf{A}_2\mathbf{D})(\mathbf{A}_1\mathbf{D}-\mathbf{A}_2\mathbf{D})^T \\ &+ (T_1(\boldsymbol{\mu})-T_2(\boldsymbol{\mu}))(T_1(\boldsymbol{\mu})-T_2(\boldsymbol{\mu}))^T\end{aligned} \tag{3}$$

The matrix $\mathbf{D}$ serves as a normalization matrix for the terms of the matrices $\mathbf{A}_1$ and $\mathbf{A}_2$; it translates the difference between the matrices to the mean distance between the image coordinates. The mean squared distance $E[\|\boldsymbol{\epsilon}\|^2]$ is the trace of (3).

We denote the vectorization of the normalized terms $\mathbf{A}_1\mathbf{D}$ and $\mathbf{A}_2\mathbf{D}$ as $\mathbf{s}_1$ and $\mathbf{s}_2$ respectively. Using this

notation, the mean squared distance between $T_1(\mathbf{p})$ and $T_2(\mathbf{p})$ becomes

$$E[\|\boldsymbol{\epsilon}\|^2] = \|\mathbf{s}_1 - \mathbf{s}_2\|^2 + \|T_1(\boldsymbol{\mu}) - T_2(\boldsymbol{\mu})\|^2 \quad (4)$$

The mean squared distance between the transformations of a point is therefore equivalent to the $l_2$ norm of the difference between two vectors: $\left(\mathbf{s}_1^T, T_1(\boldsymbol{\mu})^T\right)^T$ and $\left(\mathbf{s}_2^T, T_2(\boldsymbol{\mu})^T\right)^T$. Furthermore, the vectors are linearly dependent on the terms of the affine transformation. This brings a great computational advantage, as the problem of minimizing the mean squared error between the affine transformations of image regions translates into finding a least squares solution between vectors in $\mathbb{R}^6$.

Note that, as $\boldsymbol{\epsilon}$ is the error of a random, uniformly sampled point in $\Omega$, minimization of $E[\|\boldsymbol{\epsilon}\|^2]$ is equivalent to minimization of the average mean squared error between the affine transformations of points in $\Omega$.

## 3 Calculation of a Global Transformation from a Set of Matched Regions

The affine error metric derived in Section 2 can be employed for calculating a **global** transformation, $T(\mathbf{p})$, between two images from a set of affine correspondences.

Let $\boldsymbol{\mu}_i, \boldsymbol{\mu}'_i$ be the centers of mass of two corresponding regions, $\Omega_i$ and $\Omega'_i$. Also let $\mathbf{p}$ be a random variable, uniformly distributed in $\Omega_i$. The mean of $\mathbf{p}$ is therefore $\boldsymbol{\mu}_i$ and the covariance matrix of $\mathbf{p}$ is the second order centralized moment matrix of $\Omega_i$:

$$\boldsymbol{\Sigma}_i = \frac{\int_{\Omega_i} (\boldsymbol{\alpha} - \boldsymbol{\mu}_i)(\boldsymbol{\alpha} - \boldsymbol{\mu}_i)^T d\boldsymbol{\alpha}}{\int_{\Omega_i} d\boldsymbol{\alpha}} \quad (5)$$

The matched regions are first employed to extract a set of affine correspondences: It is shown in [15] that a coordinate system, $\mathbf{q}$, invariant to affine transformations, can be determined from the second order moment matrix of the region, such that

$$\mathbf{q} = \mathbf{R}\mathbf{D}_i^{-1}(\mathbf{p} - \boldsymbol{\mu}_i) \quad (6)$$

where $\mathbf{R}$ is an unknown rotation matrix and $\mathbf{D}_i\mathbf{D}_i^T$ is the Cholesky factorization of $\boldsymbol{\Sigma}_i$. We use the SIFT descriptor [11] to detect the unknown rotation (identified as the direction of the dominant gradient of the image at $\Omega$). Let $\mathbf{N}_i = \mathbf{R}\mathbf{D}_i^{-1}$ be the affine normalization matrix for $\Omega_i$. In a similar manner let $\mathbf{N}'_i$ be the affine normalization matrix of $\Omega'_i$. The affine normalizations yield an estimation of an affine relation between the regions (as in [1]):

$$T_i(\mathbf{p}) = \boldsymbol{\mu}'_i + (\mathbf{N}'_i)^{-1}\mathbf{N}_i(\mathbf{p} - \boldsymbol{\mu}_i) \quad (7)$$

Note that the affine normalization matrix is also related to the affine error metric as $\mathbf{N}_i^{-1}$ is a factorization of $\boldsymbol{\Sigma}_i$. It can therefore be used instead of $\mathbf{D}_i$ in (4) for normalization of the affine transformation matrix.

We next employ the affine approximation of the global transformation $T$ at $\boldsymbol{\mu}_i$ for calculation of $T$. The mean squared error between the transformation of $\Omega_i$ by $T$ to the transformation of $\Omega_i$ by $T_i$ is given by (4):

$$E[\|\boldsymbol{\epsilon}_i\|^2] = \|\mathbf{s}_i - \mathbf{t}_i\|^2 + \|\boldsymbol{\mu}'_i - T(\boldsymbol{\mu}_i)\|^2 \quad (8)$$

where $\mathbf{s}_i$ is a vectorization of $\mathbf{A}_i\mathbf{N}_i^{-1}$ and $\mathbf{t}_i$ is a vectorization of $\frac{dT}{d\mathbf{p}}|_{\boldsymbol{\mu}_i}\mathbf{N}_i^{-1}$.

Given $n$ matched regions, and assuming equal probability of a point $\mathbf{p}$ to originate from each matched region, minimization of (8) jointly for all $i = 1...n$ is equivalent to minimizing the $l_2$ norm of the vector

$$\begin{pmatrix} T(\boldsymbol{\mu}_1) \\ \mathbf{t}_1 \\ \vdots \\ T(\boldsymbol{\mu}_n) \\ \mathbf{t}_n \end{pmatrix} - \begin{pmatrix} \boldsymbol{\mu}'_1 \\ \mathbf{s}_1 \\ \vdots \\ \boldsymbol{\mu}'_n \\ \mathbf{s}_n \end{pmatrix} \quad (9)$$

Note that in cases where $\mathbf{t}_i$ and $T(\boldsymbol{\mu}_i)$ are linearly dependent on the parameters of the global transformation $T$ (such as a case where $T$ is an affine transformation) the problem can be easily solved as a least squares solution for the terms of $T$.

## 4 Linear Solutions for Homography and Fundamental Matrix Estimation

When estimating the parameters of an homography or of the fundamental matrix, minimization of the affine norm (9) between the measured regions is non linear. However, we can utilize (9) to derive a linear minimization scheme of the algebraic error between the images by using homogeneous coordinates. In the following section we describe linear solutions for estimating the homography between a pair of images and for estimating the fundamental matrix between images.

### 4.1 Using the Affine Metric for Linear Calculation of Homographies

Finding an homography between a pair of images is a common task in computer vision as it describes the relation between different views of the same world plane by pinhole cameras [9]. In this section we derive equations for calculating the homography by minimizing the affine error metric. The problem is more complex than solving for an affine transformation as both the

point correspondence and the local derivatives of the deformation are not linearly dependent on the terms of the homography. However, with a proper normalization, minimization of the algebraic error using homogeneous coordinates provides a good approximation of the optimal result.

We first evaluate (9) for the case where the transformation is an homography. Let $\mathbf{A}, \boldsymbol{\mu}, \boldsymbol{\mu}'$ be the parameters of an affine transformation that locally approximates the homography. We denote the terms of $\mathbf{A}$ as

$$\mathbf{A} = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \tag{10}$$

and the terms of the homography, $H$, expanded in homogeneous coordinates, as

$$\mathbf{H} = \begin{pmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{pmatrix} \tag{11}$$

where

$$\mathbf{H}(\boldsymbol{\mu}) = \left( \frac{h_1\mu_x + h_2\mu_y + h_3}{h_7\mu_x + h_8\mu_y + h_9}, \frac{h_4\mu_x + h_5\mu_y + h_6}{h_7\mu_x + h_8\mu_y + h_9} \right)^T \tag{12}$$

and $\mu_x$,$\mu_y$ are the first and second coordinates of $\boldsymbol{\mu}$. By denoting $k = h_7\mu_x + h_8\mu_y + h_9$, the derivatives of $\mathbf{H}$ are given by

$$\frac{d\mathbf{H}}{d\boldsymbol{\mu}} = \frac{1}{k} \begin{pmatrix} h_1 - h_7\mathbf{H}(\boldsymbol{\mu})_x & h_2 - h_8\mathbf{H}(\boldsymbol{\mu})_x \\ h_4 - h_7\mathbf{H}(\boldsymbol{\mu})_y & h_5 - h_8\mathbf{H}(\boldsymbol{\mu})_y \end{pmatrix} \tag{13}$$

Since $\boldsymbol{\mu}'$ is a noisy observation on $H(\boldsymbol{\mu})$, then following (13):

$$\frac{d\mathbf{H}}{d\boldsymbol{\mu}} = \frac{1}{k} \begin{pmatrix} h_1 - h_7\boldsymbol{\mu}'_x & h_2 - h_8\boldsymbol{\mu}'_x \\ h_4 - h_7\boldsymbol{\mu}'_y & h_5 - h_8\boldsymbol{\mu}'_y \end{pmatrix} \tag{14}$$

The minimum mean squared distance between the global homography and the local affine measured approximations can be expressed by substituting (12) and (14) to (9). Note that the minimization of the mean geometric error in (9) does not result in a linear set of equations. However, given $n$ matched regions, by multiplying the terms of the vectors by the appropriate $k_i$, we convert the problem into the minimization of

$$\left\| \begin{pmatrix} k_1 H(\boldsymbol{\mu}_1) \\ k_1\mathbf{t}_1 \\ \vdots \\ k_N H(\boldsymbol{\mu}_n) \\ k_N\mathbf{t}_n \end{pmatrix} - \begin{pmatrix} k_1\boldsymbol{\mu}'_1 \\ k_1\mathbf{s}_1 \\ \vdots \\ k_N\boldsymbol{\mu}'_n \\ k_N\mathbf{s}_n \end{pmatrix} \right\|^2 \tag{15}$$

where $\boldsymbol{\mu}_i$, $\boldsymbol{\mu}'_i$, $\mathbf{s}_i$ and $\mathbf{t}_i$, $i = 1, ..., N$, are known from the affine correspondences and $h_1, ..., h_9$ are the unknowns.

The minimization problem forms a set of $6N$ linear homogeneous equations. The equations can be expressed in the form $\mathbf{Lh} = 0$ where $\mathbf{h}$ is a vectorization of $\mathbf{H}$ and $\mathbf{L}$ depends only on the matched regions. We then solve the least squares problem given that $\|\mathbf{h}\| = 1$. This problem is solved by setting $\mathbf{h}$ to be the last column of $\mathbf{V}$, where $\mathbf{L} = \mathbf{U}\Sigma\mathbf{V}^T$ is the SVD of $\mathbf{L}$ (see for example [9]).

Note that this set of equations minimizes the algebraic error and not the geometric error. A pre normalization of the coordinates so that the mean value of $\{\boldsymbol{\mu}_i\}_{i=1,...,n}$ and $\{\boldsymbol{\mu}'_i\}_{i=1,...,n}$ is zero and the average distance from the origin is $\sqrt{2}$ (as in [8]) is therefore required to ensure that the $k_i$'s do not change drastically between the equations.

### 4.2 Using the Affine Metric for Calculation of the Fundamental Matrix

In this section we employ the affine normalization for calculating the fundamental matrix between a pair of images. In contrast with the homography calculation, the fundamental matrix is not a global transformation of points between images as it relates points in one image to the corresponding epipolar lines in the other image. However, the affine error metric can be employed to correctly normalize the constraints on the affine transformation from an affine relation (a non-linear solution for exact minimization of the affine error metric is described in Section 5.3).

We denote the terms of the fundamental matrix as

$$\mathbf{F} = \begin{pmatrix} f_1 & f_2 & f_3 \\ f_4 & f_5 & f_6 \\ f_7 & f_8 & f_9 \end{pmatrix} \tag{16}$$

The relation between the fundamental matrix and pointwise correspondences between a pair of images is given by

$$\begin{pmatrix} \mu'_x & \mu'_y & 1 \end{pmatrix} \mathbf{F} \begin{pmatrix} \mu_x \\ \mu_y \\ 1 \end{pmatrix} = 0 \tag{17}$$

For simplicity, we first consider the case where $\boldsymbol{\mu}' = \boldsymbol{\mu} = 0$. In this case the pointwise relation implies that

$$f_9 = 0 \tag{18}$$

Since an affine correspondence also provides the derivatives at the point of correspondence, we obtain two additional linear constraints by taking the derivatives of (17) with respect to $\boldsymbol{\mu}$ and letting $\frac{d\boldsymbol{\mu}'}{d\boldsymbol{\mu}} = \mathbf{A}$ and $f_9 = 0$. The resulting equations are

$$\begin{pmatrix} f_7 \\ f_8 \end{pmatrix} + \mathbf{A}^T \begin{pmatrix} f_3 \\ f_6 \end{pmatrix} = 0 \tag{19}$$

Equation (18) represents a pointwise relation and equation (19) represents a relation dependent on the derivatives in the proximity of the correspondence point. In Section 4.1 we showed that proper normalization of the terms in $\mathbf{A}$ is achieved by multiplication with $\mathbf{N}^{-1}$ (as $\mathbf{N}^{-1}\mathbf{N}^{-T} = \boldsymbol{\Sigma}$). We therefore multiply (19) on the left by $\mathbf{N}^{-T}$ to yield

$$\mathbf{N}^{-T}\begin{pmatrix} f_7 \\ f_8 \end{pmatrix} + (\mathbf{A}\mathbf{N}^{-1})^T \begin{pmatrix} f_3 \\ f_6 \end{pmatrix} = 0 \tag{20}$$

We can further simplify the equation since $\mathbf{A}$ can be expressed, (see, (7)) using the normalization matrices of the matched regions as $\mathbf{A} = \mathbf{N}'^{-1}\mathbf{N}$, therefore

$$\mathbf{N}^{-T}\begin{pmatrix} f_7 \\ f_8 \end{pmatrix} + \mathbf{N}'^{-T} \begin{pmatrix} f_3 \\ f_6 \end{pmatrix} = 0 \tag{21}$$

Each affine correspondence therefore results in 3 linear homogeneous constraints on the fundamental matrix.

The extension to the general case where $\boldsymbol{\mu}$ and $\boldsymbol{\mu}'$ can take any value, is obtained by shifting the coordinate systems. Let $\mathbf{S}_{\boldsymbol{\mu}}$ be a matrix that shifts the coordinate system by $\boldsymbol{\mu}$,

$$\mathbf{S}_{\boldsymbol{\mu}} = \begin{pmatrix} 1 & 0 & \mu_x \\ 0 & 1 & \mu_y \\ 0 & 0 & 1 \end{pmatrix} \tag{22}$$

Then, by using the constraints (18) and (21) on the terms of $\widetilde{\mathbf{F}} = \mathbf{S}_{\boldsymbol{\mu}'}\mathbf{F}\mathbf{S}_{\boldsymbol{\mu}}$, we form a set of 3 homogeneous linear equations as the shift of coordinates, shifts $\boldsymbol{\mu}$ and $\boldsymbol{\mu}'$ to 0. As in the homography estimation, a pre normalization of the coordinates is required, so that the mean value of $\{\boldsymbol{\mu}_i\}_{i=1,\ldots,n}$ and $\{\boldsymbol{\mu}'_i\}_{i=1,\ldots,n}$ is zero and the average distance from the origin is $\sqrt{2}$.

## 5 Non Linear Minimization of the Affine Geometric Error in Both Images

The minimization of the $l_2$ norm in (9) detailed in Section 3 only takes into consideration errors between regions in the second image and the transformations of the corresponding regions from the first image. Hence, the errors are only measured in the second image and do not reflect errors in the first image (measured by the distance between regions in the first image and the transformations of the corresponding regions in the second image). Such an objective function is convenient for the derivation of linear solutions such as those described in Section 4. However, if we omit the requirement of a linear solution, the error in both images can be taken into account in order to improve the estimate of the parameters of the model.

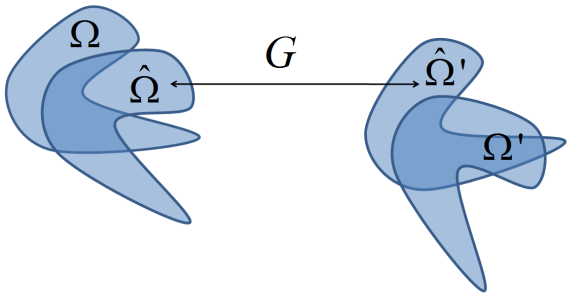


**Fig. 2** Two measured regions $\Omega$,$\Omega'$ and the rectified regions $\widehat{\Omega}$,$\widehat{\Omega}'$. The affine relation between $\widehat{\Omega}$ and $\widehat{\Omega}'$ perfectly agrees with the global model, $G$. The non linear solution minimizes the mean squared error between the measured and the rectified regions.

Given a global model, $G$, that represents the relation between two observations of a scene (a homography or an epipolar relation), the affine relation between $\Omega$ and $\Omega'$ does not necessarily comply to the model. We therefore wish to create new, rectified regions, $\widehat{\Omega}$ and $\widehat{\Omega}'$ that do agree with the model. The rectified regions should be as close as possible to the original, observed regions. The configuration is illustrated in Figure 2. By minimizing the affine norm we can guarantee that the average mean squared distance between each point in the rectified regions and its matching point in the original regions will be minimal.

We express the affine transformations relating the observed and rectified regions by their affine normalizations using (7). Let $\widehat{\mathbf{N}}$ be the affine normalization matrix of $\widehat{\Omega}$ and $\widehat{\boldsymbol{\mu}}$ the center of mass of the region. Also let $\widehat{\mathbf{N}}'$ and $\widehat{\boldsymbol{\mu}}'$ be the corresponding parameters of $\widehat{\Omega}'$ . The affine relations between the rectified and measured regions are then expressed as

$$\widehat{\mathbf{p}} = \widehat{\mathbf{N}}^{-1}\mathbf{N}(\mathbf{p} - \boldsymbol{\mu}) + \widehat{\boldsymbol{\mu}} \tag{23}$$

$$\widehat{\mathbf{p}}' = \widehat{\mathbf{N}'}^{-1}\mathbf{N}'(\mathbf{p}' - \boldsymbol{\mu}') + \widehat{\boldsymbol{\mu}}' \tag{24}$$

where $\widehat{\mathbf{p}}$ and $\mathbf{p}$ are the corresponding coordinates of the rectified and the observed regions respectively. Note that as our analysis depends only on a first-order approximation of the deformation between regions, the relation between the rectified regions and the measured regions is affine. According to Section 2, minimization of the mean squared distance between the rectified regions and the observed regions is equivalent to minimization of (4) between the identity transformation (that leaves $\Omega$ and $\Omega'$ unchanged) and the affine transformations in (23) and (24).

We therefore express the mean squared distance (4) between $\Omega$ and $\widehat{\Omega}$ as the affine error metric between two affine transformations. The first transformation, $T_1$, represents the rectifying transformation: $\mathbf{A}_1 = \widehat{\mathbf{N}}^{-1}\mathbf{N}$,

$T_1(\boldsymbol{\mu}) = \widehat{\boldsymbol{\mu}}$. The second transformation, $T_2$, represents the identity transformation: $\mathbf{A}_2 = \mathbf{I}$ , $T_2(\boldsymbol{\mu}) = \boldsymbol{\mu}$. Recall from the definition of $\mathbf{N}$ in Section 3 and from the discussion that follows the derivation of (7), that $\mathbf{N}$ is the affine normalization matrix of $\Omega$, and that $\mathbf{N}^{-1}$ is a factorization of $\boldsymbol{\Sigma}$. Thus, normalizing both $\mathbf{A}_1$ and $\mathbf{I}$ by $\mathbf{N}^{-1}$ and substituting the terms into (4) we have:

$$E[\|\boldsymbol{\epsilon}\|^2] = \left\|\widehat{\mathbf{N}}^{-1} - \mathbf{N}^{-1}\right\|_{HS}^2 + \|\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}\|^2 \tag{25}$$

where $HS$ stands for the Hilbert-Schmidt norm. The same derivation applies to $\Omega'$.

Equation (25) gives us a simple way to formalize the mean squared difference between the measured regions and the rectified regions. Let $\mathbf{m}$ be the vectorization of the measured quantities: $\boldsymbol{\mu}, \mathbf{N}^{-1}, \boldsymbol{\mu}', \mathbf{N}'^{-1}$. Also let $\widehat{\mathbf{m}}$ be the vectorization of the rectified quantities: $\widehat{\boldsymbol{\mu}}, \widehat{\mathbf{N}}^{-1}, \widehat{\boldsymbol{\mu}}', \widehat{\mathbf{N}}'^{-1}$. Then, the mean squared difference between the measured regions and the estimated regions is given by

$$\|\mathbf{m} - \widehat{\mathbf{m}}\|^2 \tag{26}$$

Note that the problem of minimizing the mean squared distance between the measured and rectified regions is expressed by (26) as minimizing the Euclidian distance between the parameters of the affine normalization of the regions. This relation is employed next for estimating global relations between the images.

### 5.1 Iterative Minimization of the Geometric Error

The minimization of the mean squared distance between the measured and the rectified regions (26) can be employed for estimation both of homographies and the epipolar relation between uncalibrated cameras. In both cases the solution is iterative as the dependence of (26) on the parameters of the models is non-linear. In this section we describe the non-linear optimization process.

Let $\mathbf{U}$ be the parameter vector of the relevant model (a homography or a parametrization of a camera matrix) and $\mathbf{V}_i$ ,$i = 1...n$ be the parameters defining rectified regions (so that given $\mathbf{U}$ and $\mathbf{V}_i$ we can calculate $\widehat{\Omega}_i$ and $\widehat{\Omega}'_i$). Then, given a set of $n$ affine correspondences (defined by $\boldsymbol{\mu}_i$,$\boldsymbol{\mu}'_i$,$\mathbf{N}_i^{-1}$,$\mathbf{N}_i'^{-1}$ , $i = 1...n$), we wish to find $\mathbf{U}$, and $\mathbf{V}_i$ ,$i = 1...n$ that minimize the sum of the squared Euclidean distance, (26), over all pairs of matched regions. We initialize the iterative process by the linear methods described in Section 4 and employ a Levenberg-Marquardt optimization process to estimate both the parameters of the model and the parameters of the rectified regions.

The details of the iterative process are as follows: Let

$$\mathbf{L} = [\mathbf{U}^T, \mathbf{V}_1^T, ..., \mathbf{V}_n^T]^T \tag{27}$$

be the vector of the parameters to be estimated. Also let $\widehat{\mathbf{m}}_i$ be a vectorization of the parameters of the rectified regions $\widehat{\boldsymbol{\mu}}_i, \widehat{N}_i^{-1}, \widehat{\boldsymbol{\mu}}'_i, \widehat{\mathbf{N}}_i'^{-1}$. We denote the vector representation of all the rectified regions as

$$\widehat{\mathbf{M}} = (\widehat{\mathbf{m}}_1^T, ..., \widehat{\mathbf{m}}_n^T)^T \tag{28}$$

Last, we define the measurement vector $M$ as

$$\mathbf{M} = (\mathbf{m}_1^T, ..., \mathbf{m}_n^T)^T \tag{29}$$

where $\mathbf{m}_i$ is a vectorization of $\boldsymbol{\mu}_i$, $\mathbf{N}_i^{-1}$, $\boldsymbol{\mu}'_i$ and $\mathbf{N}_i'^{-1}$.

Thus, minimization of the mean squared displacement between points in the matched regions turns into finding an optimal set of parameters $L$ that minimizes the $l_2$ distance

$$\left\|\mathbf{M} - \widehat{\mathbf{M}}\right\| \tag{30}$$

As indicated above, we solve the minimization problem by a Levenberg-Marquardt optimization process: given a set of $n$ matched regions, we perform a minimization over $|\mathbf{U}|+n|\mathbf{V}_i|$ parameters. The minimization problem exhibits a sparse structure as given the model parameters $\mathbf{U}$, there is no dependency between parameters of different regions. Therefore $\frac{d\widehat{\mathbf{m}}_i}{d\mathbf{V}_j} = 0$ in cases where $i \neq j$ and the Jacobian $\frac{d\widehat{\mathbf{M}}}{d\mathbf{L}}$ exhibits a sparse structure.

We employ the sparse Levenberg-Marquardt algorithm given in [9] for the iterative solution. Due to the sparse structure of the Jacobian matrix, the complexity of the optimization problem is linear in the number of matched regions instead of quadratic as in the non-sparse case.

### 5.2 Optimal Homography Calculation

In this section we detail on using the optimization process described in Section 5.1 for estimation of an optimal Homography, $\mathbf{H}$. We define the vector of model parameters $\mathbf{U}$, the parameters defining the rectified regions $\mathbf{V}_i$ ,$i = 1...n$ and their relation to the parameters of the rectified regions $\widehat{\mathbf{M}}$.

We set $\mathbf{U}$ to be the vectorization of $\mathbf{H}$. Since by their definition, the rectified regions are related by a homography, it is sufficient to parameterize the regions in only one image (as parametrization in the other image provides no additional constraints). Therefore, $\mathbf{V}_i$ is defined to be the vectorization of $\widehat{\boldsymbol{\mu}}_i^T$ and $\widehat{\mathbf{N}}_i^{-1}$. The rectified regions in the second image are calculated from the regions in the first image using the homography.

As the affine relation between $\widehat{\Omega}$ and $\widehat{\Omega}'$ is known by (7), the affine normalization matrix of a rectified region from the first image, $\widehat{\mathbf{N}}$, is related to $\widehat{\mathbf{N}}'$ by

$$\frac{d\widehat{\mathbf{p}}'}{d\widehat{\mathbf{p}}} = \widehat{\mathbf{N}}'^{-1}\widehat{\mathbf{N}} \tag{31}$$

where $\frac{d\widehat{\mathbf{p}}'}{d\widehat{\mathbf{p}}}$ is defined by the homography. Therefore

$$\widehat{\mathbf{N}}'^{-1} = \frac{d\widehat{\mathbf{p}}'}{d\widehat{\mathbf{p}}}\widehat{\mathbf{N}}^{-1} \tag{32}$$

The estimated center of mass of a region in the second image $\widehat{\boldsymbol{\mu}}'$ can also be directly calculated by the homography $H$ and $\widehat{\boldsymbol{\mu}}$.

Given a set of $n$ measured affine correspondences (denoted as $\mathbf{m}_i$, $i = 1...n$) , the estimated affine correspondences (28) are therefore fully parameterized by $L$. As described in Section 5.1 ,the problem is then solved by a Levenberg-Marquardt iterative solution.

### 5.3 Optimal Fundamental Matrix Calculation

In this section we detail on using the optimization process described in Section 5.1 for estimating the camera matrices relating two views. We define the vector of model parameters $\mathbf{U}$, the parameters defining the rectified regions $\mathbf{V}_i$ ,$i = 1...n$ and their relation to the parameters of the rectified regions $\widehat{\mathbf{M}}$.

In the case of pointwise matches, estimation of the relations between a pair of cameras also produces an estimation of the world points that project those matches. In a similar manner, we show that a set of world planes can be estimated together with the camera relations.

Given a pair of matching rectified regions $\widehat{\Omega}$ and $\widehat{\Omega'}$, we employ affine normalized coordinate systems to estimate the world plane that is viewed by the regions. Recall that as the rectified regions are related by an affine transformation, an affine normalized coordinate, $\mathbf{q}$, defines corresponding points in both regions. As the rectified regions fully agree with the camera configuration, there exist $x \in R^3$ that is projected to those points.

Let $\boldsymbol{\xi} \in R^3$ denote the center of mass of the world region projected to both rectified regions. We denote the local relation between the world coordinates around $\boldsymbol{\xi}$ and the affine normalized coordinate, $\mathbf{q}$, as $\mathbf{Q} = \frac{dx}{d\mathbf{q}} \in R^{3\times 2}$. The configuration is described in Figure 3. Note that the differential relation, $\mathbf{Q}$, and the point, $\boldsymbol{\xi}$, define a world plane containing $\boldsymbol{\xi}$.

The rectified affine normalization matrices $\widehat{\mathbf{N}}$ and $\widehat{\mathbf{N}}'$ are related to $\mathbf{Q}$ by the affine normalized coordinate

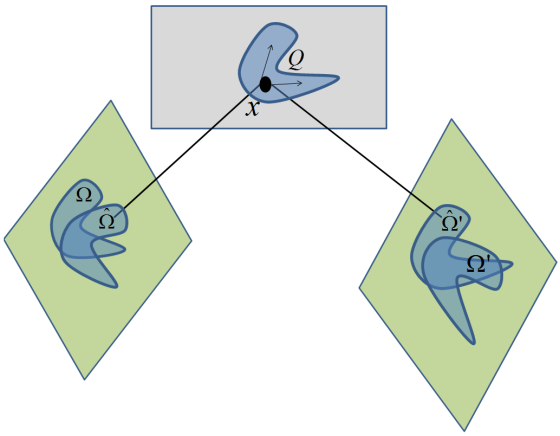


**Fig. 3** The estimated region, viewed by the two cameras, is represented by the coordinates of its center of mass, $\boldsymbol{\xi}$, and by the differential relation between movements in the affine normalized coordinate systems to movements around $x = \boldsymbol{\xi}$: $Q = \frac{dx}{d\mathbf{q}}$, evaluated at the center of mass .

system. Let $\widehat{\mathbf{p}}$ and $\widehat{\mathbf{p}}'$ be the projection of $x$ on the two views, then

$$\widehat{\mathbf{N}}^{-1} = \frac{d\widehat{\mathbf{p}}}{d\mathbf{q}} = \frac{d\widehat{\mathbf{p}}}{dx}\frac{dx}{d\mathbf{q}} = \frac{d\widehat{\mathbf{p}}}{dx}\mathbf{Q} \tag{33}$$

$$\widehat{\mathbf{N}'}^{-1} = \frac{d\widehat{\mathbf{p}}'}{d\mathbf{q}} = \frac{d\widehat{\mathbf{p}}'}{dx}\frac{dx}{d\mathbf{q}} = \frac{d\widehat{\mathbf{p}}'}{dx}\mathbf{Q} \tag{34}$$

As in the gold-standard calculation of the fundamental matrix in the pointwise case [9], we set the first camera to be fixed: $P = (I|0)$ and estimate the parameters of the second camera $P'$. Note that $\frac{d\widehat{\mathbf{p}}}{dx}$, $\frac{d\widehat{\mathbf{p}}'}{dx}$ and the projection of $\boldsymbol{\xi}$ on the two views ($\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\mu}}'$) can be calculated from $P$ and $P'$.

Given a set of $n$ measured affine correspondences (denoted by $\mathbf{m}_i$, $i = 1...n$), the rectified regions (28) are therefore fully parameterized by $P'$ ,$\boldsymbol{\xi}_i$ and $\mathbf{Q}_i$, $i = 1...n$. We set $\mathbf{U}$ to be the vectorization of $P'$ and $\mathbf{V}_i$ to be the vectorization of $\boldsymbol{\xi}_i$ and $\mathbf{Q}_i$. As described in Section 5.1 ,the problem is then solved by a Levenberg-Marquardt iterative solution.

## 6 Experiments

In the following section we compare the accuracy of the proposed methods for homography estimation and fundamental matrix estimation with two other estimation methods. The first method discards the affine relation and uses only the center of mass of the regions for matching. The second method employs the affine relation to extract three corresponding points from each matched region. As a local affine invariant coordinate system can be extracted from each matched region [15],

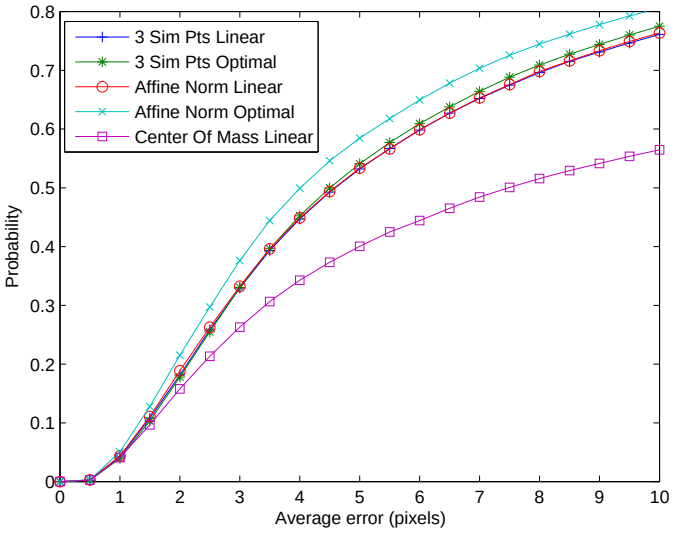


**Fig. 4** Cumulative distribution function of the mean squared distance in both images (for each experiment) from the center of mass of the regions in the second image to the transformation of the centers of mass of the region in the first image by the calculated homography. 1000 experiments were carried out for each pair of images. In each experiment, 4 matched regions were chosen at random; an homography was then calculated from the chosen 4 regions using each of the homography estimation methods.

two more points of correspondence are the transformations of unit vectors in the directions of the $x$ axis and the $y$ axis in the normalized coordinate systems to the coordinate systems of the images.

For each of the methods we first perform linear estimation using homogeneous coordinates, we then refine the solution by an iterative method as described in Section 5 to minimize the error in both images. We present both the results of the linear methods and of the iterative solutions.

We use MSER [12] regions, matched between pairs of images prior to the experiments to generate a set of inlier matches between the images. Each pair of matched regions is then used to generate a local estimation of the affine deformation between the images.

### 6.1 Estimation of homographies

The "Covariant Features Dataset" [14] is used for comparison of the accuracy of the homography estimation methods as the images in the dataset are related by homographies. The dataset contains series of images (named as: bikes, graf, bark, boat, leuven, trees, ubc and wall) with viewpoint changes, illumination changes, blur and compression errors. Each series consists of 6 images. The first image serves as a reference image; the rest of the images show an increasing amount of geometric and radiometric changes from the reference image (caused by changes of the viewpoint, zoom and illumination).

In each series, we compare the reference image to the rest of the images. We perform 1000 experiments for each pair of images. In each experiment, 4 matched regions are chosen at random; an homography is then calculated from the chosen 4 regions. As we have no knowledge on the true location of the regions, we measure the error in both images on all the matched regions as the root of the mean squared distance between the measured center of mass of the regions to the transformation of the centers of mass from the other image by the homography:

$$\left(\frac{1}{2n}\sum_{i=1}^{n} d(\boldsymbol{\mu}_i', H(\boldsymbol{\mu}_i))^2 + d(\boldsymbol{\mu}_i, H^{-1}(\boldsymbol{\mu}_i'))^2\right)^{\frac{1}{2}} \tag{35}$$

The linear pointwise estimations are performed by using the DLT method and the iterative methods by the gold-standard algorithm for homography estimation [7]. In Figure 4 we depict the cumulative distribution function of the root mean squared error for all of the compared methods. Note that in this case the iterative solution and the linear solution for estimating the homography from 4 points are the same as this is the minimal configuration for homography estimation. An example of the regions transformed from the first image to the second image by homographies calculated using each of the estimation methods is shown in Figure 5.

The results clearly show that discarding the information in the affine matches greatly reduce the accuracy of the homography estimation. In addition, the results show that minimization of the mean squared error between the regions results in better accuracy than minimization of the error between 3 simulated points from each region both when using linear estimation methods and by the iterative solution. As the computational complexity of the 3 points based method and of the affine normalized method is the same (linear complexity in the number of regions) the superiority of the affine normalized method proposed in this paper is clear.

### 6.2 Fundamental matrix estimation

A test for the accuracy in estimating the fundamental matrix is performed using the "Daisy dataset" [20],[21]. The dataset consist of pairs of images taken with a calibrated stereo rig. As in the previous experiment, we perform 1000 experiments for each pair of images. In

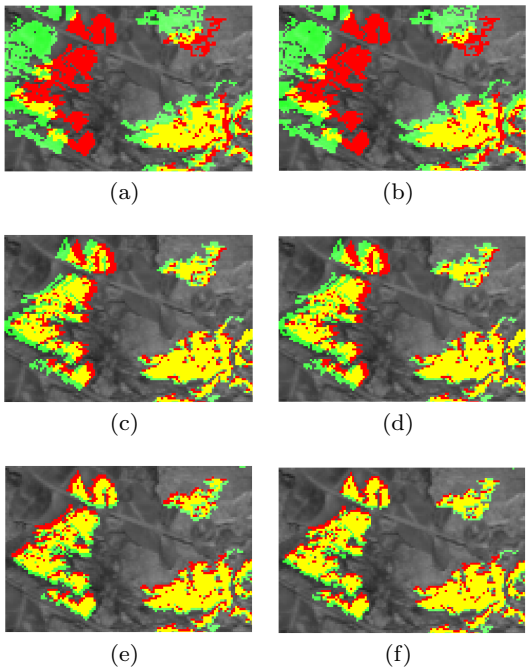

**Fig. 5** The transfer of regions from the first image to the second by the calculated homographies. The homographies were calculated using six methods: 5(a) - linear pointwise calculation using only the center of mass of the regions, 5(b) - iterative pointwise calculation using only the center of mass of the regions, 5(c) - linear calculation from 3 points of each region, 5(d) - iterative calculation from 3 points of each region, 5(e) - linear calculation using the proposed method, 5(f) - iterative calculation using the proposed method . The original MSER regions detected in the second image are marked in red, the transferred regions by the homographies are marked in green and the intersections of the regions are marked in yellow.

each experiment, 8 matched regions are chosen at random; a fundamental matrix is then calculated from the chosen 8 regions. We measure the error in both images on all the matched regions as the root of the mean squared distance between the measured center of mass of the regions to the epipolar line, calculated from the centers of mass of the region on the other image:

$$\left(\frac{1}{2n}\sum_{i=1}^{n} d(\boldsymbol{\mu}_i', F\left[\boldsymbol{\mu}_i,1\right]^T)^2 + d(\boldsymbol{\mu}_i, F^T\left[\boldsymbol{\mu}_i',1\right]^T)^2\right)^{\frac{1}{2}} \tag{36}$$

The linear pointwise estimations are performed by using the normalized 8 points method and the iterative pointwise solutions are obtained by the gold-standard algorithm for calculating the fundamental matrix [8]. Figure 6 depicts the cumulative distribution function of the average distance. An example of epipolar lines, calculated using the methods compared in this paper, is shown in Figure 7.

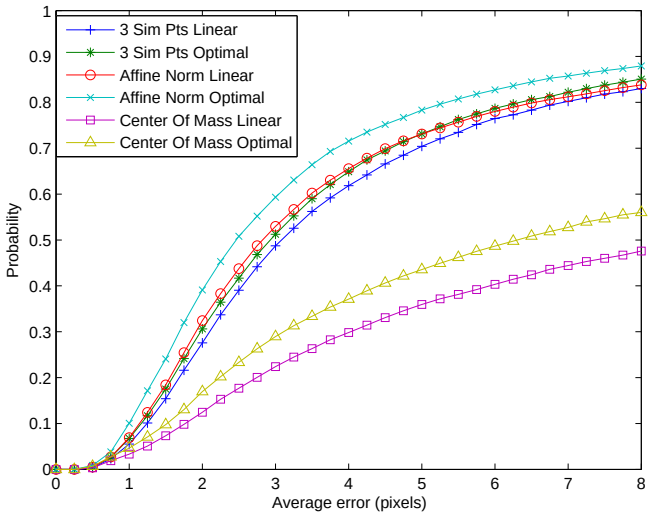


**Fig. 6** Cumulative distribution function of the mean squared distance (for each experiment) from the center of mass of the regions in each image to epipolar lines, generated by the calculated fundamental matrices and the centers of mass of the regions in the other image. 1000 experiments were carried out for each pair of images. In each experiment, 8 matched regions were chosen at random; a fundamental matrix was then calculated from the chosen 8 regions using each of the fundamental matrix estimation methods.

As in the estimation of homographies, the results clearly show that discarding the information in the affine matches greatly reduces the accuracy of the fundamental matrix estimation. Once again, it is shown that the proposed method is more accurate then choosing 3 points from each region. Moreover, as the computational complexity of these methods is the same, it is clearly better to minimize the affine normalized error than to minimize the error between simulated points.

## 7 Discussion - the Error Measured by Simulated Points

We showed in Section 5, equation (25), that minimization of the mean squared error between an estimated region and the measured region is equivalent to minimization of

$$\|\widehat{n}_1 - n_1\|^2 + \|\widehat{n}_2 - n_2\|^2 + \|\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}\|^2 \tag{37}$$

where $\widehat{n}_1, \widehat{n}_2$ are the columns of $\widehat{N}^{-1}$ and $n_1, n_2$ are the columns of $N^{-1}$. We now turn to look at what error is being minimized in the case where the transformation is represented by simulated points. In such case, the original region is represented by $\boldsymbol{\mu}$,$\boldsymbol{\mu} + n_1$ and $\boldsymbol{\mu} + n_2$. The corresponding points in the estimated region are

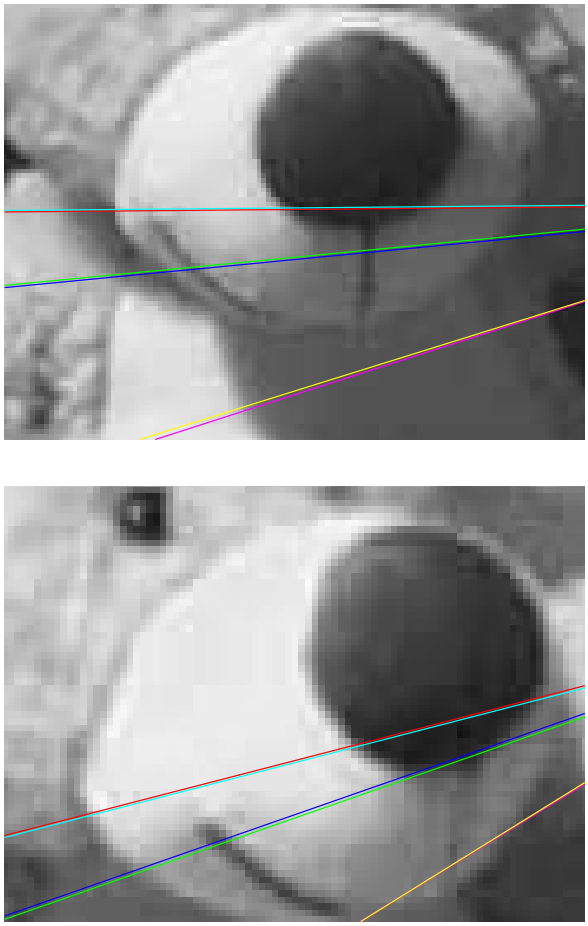

**Fig. 7** Example of epipolar lines calculated by the methods compared in this paper. The top and bottom images are corresponding sections from the two images used for calculating the fundamental matrix. Each color is associated with a different method: cyan - the proposed iterative affine normalized method, red - the proposed linear affine normalized method, green - iterative minimization using 3 points, blue - linear estimation using 3 points, yellow - iterative minimization using only the center of mass of the regions, magenta - linear estimation using only the center of mass of the regions. Note that the epipolar lines calculated by the proposed method are in the best agreement among all the tested methods.

$\widehat{\boldsymbol{\mu}}$,$\widehat{\boldsymbol{\mu}}+\widehat{n}_1$ and $\widehat{\boldsymbol{\mu}}+\widehat{n}_2$ and the measured error is

$$\|\widehat{n}_1 - n_1 + \widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}\|^2 + \|\widehat{n}_2 - n_2 + \widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}\|^2 + \|\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}\|^2 \tag{38}$$

By comparing (38) to the mean squared error between the measured and estimated regions (37), it is clear that the points based method is overly biased towards minimization of the error of the center of the region. Therefore, there is no direct relation between the mean squared error between the regions and the error measured from simulated points.

## 8 Conclusion

In this paper we provide a geometric meaning to the distance between the parameters of affine transformations. We show that a proper normalization of the affine transformation parameters results in equivalence of the proposed metric with the $l_2$ norm between the normalized affine parameters. We employ the normalization to derive methods for estimating homographies between images and for estimating the fundamental matrix between images.

To the best of our knowledge, this is the first method that provides a geometric meaning to the distance between the parameters of an affine transformation. A direct treatment of the affine parameters as a vector in $\mathbb{R}^6$ does not have a geometric meaning as it does not take into account the size and shape of the region. In addition, as we show in this paper, the mean squared error between different transformations of a region is not represented accurately by simulating points from the region.

The experiments carried out in this paper show that the proposed method is preferable to the current approach that simulates points according to the affine normalization of a region, as it yields more accurate results with the same computational load. The experiments further verify that discarding the affine transformation by using only the center of mass of the regions greatly degrades the accuracy of the estimates.

## References


1. Bentolila, J., Francos, J.M.: Affine consistency graphs for image representation and elastic matching. In: International Conference on Image Processing (2012)
2. Cho, M., Lee, J., Lee, K.M.: Feature correspondence and deformable object matching via agglomerative correspondence clustering. In: Computer Vision, 2009 IEEE 12th International Conference on, pp. 1280 –1287 (2009)
3. Chum, O., Matas, J., Stepan, O.: Epipolar geometry from three correspondences. In: Computer Vision Winter Workshop (2003)
4. Faugeras, O.: Three-Dimensional Computer Vision: A Geometric Viewpoint. Artificial Intelligence. Mit Press (1993)
5. Ferrari, V., Tuytelaars, T., Van Gool, L.: Simultaneous Object Recognition and Segmentation from Single or Multiple Model Views. International Journal of Computer Vision **67**(2), 159–188 (2006)
6. Harris, C., Stephens, M.: A combined corner and edge detection. In: Proceedings of The Fourth Alvey Vision Conference, pp. 147–151 (1988)
7. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision, 2 edn. Cambridge University Press, New York, NY, USA (2003)
8. Hartley, R.I.: In defence of the 8-point algorithm. In: Proceedings of the Fifth International Conference on Computer Vision, ICCV '95, pp. 1064–. IEEE Computer Society, Washington, DC, USA (1995)

9. Hartley, R.I., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press, ISBN: 0521623049 (2000)
10. Kannala, J., Brandt, S.: Quasi-dense wide baseline matching using match propagation. In: Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on, pp. 1–8 (2007)
11. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision **60**, 91–110 (2004)
12. Matas, J.: Robust wide-baseline stereo from maximally stable extremal regions. Image and Vision Computing **22**(10), 761–767 (2004)
13. Mikolajczyk, K., Schmid, C.: An affine invariant interest point detector. In: Proc. European Conf. Computer Vision, pp. 128–142. Springer Verlag (2002)
14. Mikolajczyk, K., Schmid, C.: Scale & affine invariant interest point detectors. Int. J. Comput. Vision **60**(1), 63–86 (2004)
15. Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T., Gool, L.V.: A comparison of affine region detectors. International Journal of Computer Vision **65**, 2005 (2005)
16. Moravec, H.: Obstacle avoidance and navigation in the real world by a seeing robot rover. In: tech. report CMU-RI-TR-80-03, Robotics Institute, Carnegie Mellon University, CMU-RI-TR-80-03 (1980)
17. Perdoch, M., Matas, J., Chum, O.: Epipolar geometry from two correspondences. In: Proceedings of the 18th International Conference on Pattern Recognition - Volume 04, ICPR '06, pp. 215–219. IEEE Computer Society, Washington, DC, USA (2006)
18. Riggi, F., Toews, M., Arbel, T.: Fundamental matrix estimation via tip - transfer of invariant parameters. In: Proceedings of the 18th International Conference on Pattern Recognition - Volume 02, ICPR '06, pp. 21–24. IEEE Computer Society, Washington, DC, USA (2006)
19. Szeliski, R., Torr, P.: Geometrically Constrained Structure from Motion: Points on Planes. In: 3D Structure from Multiple Images of Large-Scale Environments, pp. 171–186 (1998)
20. Tola, E., Lepetit, V., Fua, P.: A fast local descriptor for dense matching. Computer Vision and Pattern Recognition, IEEE Computer Society Conference on **0**, 1–8 (2008)
21. Tola, E., Lepetit, V., Fua, P.: Daisy: An efficient dense descriptor applied to wide-baseline stereo. Pattern Analysis and Machine Intelligence, IEEE Transactions on **32**(5), 815 –830 (2010)
22. Tuytelaars, T., Gool, L.V.: Wide baseline stereo matching based on local, affinely invariant regions. In: Proc. BMVC, pp. 412–425 (2000)
23. Tuytelaars, T., Mikolajczyk, K.: Local invariant feature detectors: A survey. Foundations and Trends in Computer Graphics and Vision **3**(3), 177–280 (2008)