

A New Method for Affine Registration of Images and Point Sets

Juho Kannala¹, Esa Rahtu¹, Janne Heikkilä¹, and Mikko Salo²

¹ Machine Vision Group
Department of Electrical and Information Engineering
University of Oulu
P.O.Box 4500, 90014 University of Oulu, Finland
{jkannala, erahtu, jth}@ee.oulu.fi

² Rolf Nevanlinna Institute
Department of Mathematics and Statistics
University of Helsinki
P.O.Box 68, 00014 University of Helsinki, Finland
msa@rni.helsinki.fi

Abstract. In this paper we propose a novel method for affine registration of images and point patterns. The method is non-iterative and it directly utilizes the intensity distribution of the images or the spatial distribution of points in the patterns. The method can be used to align images of isolated objects or sets of 2D and 3D points. For Euclidean and similarity transformations the additional constraints can be easily embedded in the algorithm. The main advantage of the proposed method is its efficiency since the computational complexity is only linearly proportional to the number of pixels in the images (or to the number of points in the sets). In the experiments we have compared our method with some other non-feature-based registration methods and investigated its robustness. The experiments show that the proposed method is relatively robust so that it can be applied in practical circumstances.

1 Introduction

The traditional approach to image registration is to first extract some salient features from both images and then find correspondence between these features. Thereafter the feature correspondences are used to recover the geometric transformation that registers the images [1]. The problem with this approach is that it is not always possible to find a sufficient number of features which can be localized accurately from both images and matched reliably between them. In addition, when registering point patterns an essential part of the problem is that correspondences between the point sets are unknown *a-priori* [2].

A different approach to registration is represented by methods that directly utilize the intensity information of the images. Unfortunately, there are quite a few direct methods for affine registration. The maximization of mutual information [3, 4] is an intensity-based technique that can be applied also in the case

of affine transformations. Though being quite a robust method it is not computationally very simple and it requires iterative optimization techniques. Two other possible methods for affine registration are the cross-weighted moments [5] and affine moment descriptors [6]. However, the cross-weighted moment method is somewhat cumbersome for large images due to its computational complexity. All these three methods, the maximization of mutual information, cross-weighted moments and affine moment descriptors can be used to register both images and point patterns [4–6].

In this paper we introduce an entirely new intensity-based method for affine registration. The method is based on a novel image transform which was recently developed to produce global affine invariant features directly from gray-scale images [7]. A slight modification of the transform makes it possible to compute such descriptor values that allow to recover the affine transformation parameters between the images. The method can be additionally extended to registration of point sets. The structure of the paper is as follows. In Section 2 we give background for the proposed registration method by first describing the affine invariant image transform behind it. The actual registration method is then described in Section 3 and the implementation is covered in Section 4. The experiments and results are presented and discussed in Sections 5 and 6.

2 Background

Recently a new affine invariant image transform was proposed for efficient computation of global affine invariant features from grayscale images [7]. This transform is the basis for our affine registration method, and hence, it is described in the following.

Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a compactly supported image function, whose centroid is $\boldsymbol{\mu}(f) = \int_{\mathbb{R}^2} \mathbf{x}f(\mathbf{x}) d\mathbf{x} / \|f\|_{L^1}$. We define the normalized function $\tilde{f}(\mathbf{x}) = f(\mathbf{x} + \boldsymbol{\mu}(f))$ by shifting the coordinate origin to the centroid of the image. Clearly, the normalized representation is invariant to translations of the image.

Then the image transform [7] is defined as follows and its affine invariance is proven below.

Definition 1 Let $f \in L^\infty(\mathbb{R}^2)$ be compactly supported. For $\alpha, \beta \in \mathbb{R}$ define

$$If(\alpha, \beta) = \frac{1}{\|f\|_{L^1}} \int_{\mathbb{R}^2} \tilde{f}(\mathbf{x}) \tilde{f}(\alpha\mathbf{x}) \tilde{f}(\beta\mathbf{x}) d\mathbf{x}. \quad (1)$$

Proposition 1 $I(f \circ \mathcal{A}^{-1}) = If$ for any affine transformation \mathcal{A} .

Proof. We first show that I is translation invariant. If $g(\mathbf{x}) = f(\mathbf{x} - \mathbf{x}_0)$ then $\|g\|_{L^1} = \|f\|_{L^1}$ and $\boldsymbol{\mu}(g) = \boldsymbol{\mu}(f) + \mathbf{x}_0$. Thus $\tilde{g}(\mathbf{x}) = g(\mathbf{x} + \boldsymbol{\mu}(g)) = \tilde{f}(\mathbf{x})$, and consequently $Ig = If$. Also, if \mathbf{A} is a nonsingular matrix let $g(\mathbf{x}) = f(\mathbf{A}^{-1}\mathbf{x})$. Since $\boldsymbol{\mu}(g) = \mathbf{A}\boldsymbol{\mu}(f)$ one has $\tilde{g}(\mathbf{x}) = f(\mathbf{A}^{-1}(\mathbf{x} + \mathbf{A}\boldsymbol{\mu}(f))) = \tilde{f}(\mathbf{A}^{-1}\mathbf{x})$ and

$$Ig(\alpha, \beta) = \frac{1}{|\det \mathbf{A}| \|f\|_{L^1}} \int_{\mathbb{R}^2} \tilde{f}(\mathbf{A}^{-1}\mathbf{x}) \tilde{f}(\alpha\mathbf{A}^{-1}\mathbf{x}) \tilde{f}(\beta\mathbf{A}^{-1}\mathbf{x}) d\mathbf{x}.$$

The change of variables $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ gives that $Ig = If$.

Remark 1. One obtains more general global affine invariants in the following way. Let $H : \mathbb{R}^k \rightarrow \mathbb{R}$ be a measurable function, and define

$$\hat{I}f(\alpha_1, \dots, \alpha_k) = \frac{1}{\|f\|_{L^1}} \int_{\mathbb{R}^2} H(\tilde{f}(\alpha_1 \mathbf{x}), \dots, \tilde{f}(\alpha_k \mathbf{x})) d\mathbf{x}. \quad (2)$$

If H satisfies some conditions (e.g. H continuous and $H(\mathbf{0}) = 0$), then the expression is well defined for compactly supported $f \in L^\infty(\mathbb{R}^2)$. The proof of Proposition 1 shows that \hat{I} is also affine invariant. By changing variables $\mathbf{x} \mapsto \alpha_1^{-1} \mathbf{x}$ we see that no information is lost if we normalize $\alpha_1 = 1$. In (1) we chose $k = 3$ and $H(x, y, z) = xyz$ and made the normalization $\alpha_1 = 1$.

The formula (1) provides a method for producing affine invariant features from a grayscale image and in [7] these features were used to classify affine transformed images which were additionally corrupted by noise and occlusion. It was observed that already a small subset of the features was enough for successful classification.

3 Registration Method

The above approach may also be utilized in affine registration after a slight modification. In the following we propose a transform which produces such descriptors in \mathbb{R}^2 that allow to recover the transformation between an image and its affine transformed version.

Definition 2 Let $f \in L^\infty(\mathbb{R}^2)$ be compactly supported. For $\alpha, \beta \in \mathbb{R}$ define

$$\mathbf{J}f(\alpha, \beta) = \frac{1}{\|f\|_{L^1}} \int_{\mathbb{R}^2} \mathbf{x} \tilde{f}(\mathbf{x}) \tilde{f}(\alpha \mathbf{x}) \tilde{f}(\beta \mathbf{x}) d\mathbf{x}. \quad (3)$$

Proposition 2 Let $\mathcal{A}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{t}$ be an affine transformation, where \mathbf{A} is a nonsingular matrix. Let f be an image function and f' the \mathcal{A} transformed version of f , i.e., $f'(\mathbf{x}) = f(\mathcal{A}^{-1}(\mathbf{x}))$. Then $\mathbf{J}f' = \mathbf{A}\mathbf{J}f$.

Proof. We get $f'(\mathbf{x}) = (f \circ \mathcal{A}^{-1})(\mathbf{x}) = f(\mathbf{A}^{-1}\mathbf{x} - \mathbf{A}^{-1}\mathbf{t})$. Then, since $\boldsymbol{\mu}(f') = \mathbf{A}\boldsymbol{\mu}(f) + \mathbf{t}$ and $\|f'\|_{L^1} = |\det \mathbf{A}| \|f\|_{L^1}$, we have $\tilde{f}'(\mathbf{x}) = f'(\mathbf{x} + \boldsymbol{\mu}(f')) = f(\mathbf{A}^{-1}(\mathbf{x} + \boldsymbol{\mu}(f')) - \mathbf{A}^{-1}\mathbf{t}) = f(\mathbf{A}^{-1}\mathbf{x} + \boldsymbol{\mu}(f)) = \tilde{f}(\mathbf{A}^{-1}\mathbf{x})$ and

$$\mathbf{J}f'(\alpha, \beta) = \frac{1}{|\det \mathbf{A}| \|f\|_{L^1}} \int_{\mathbb{R}^2} \mathbf{x} \tilde{f}(\mathbf{A}^{-1}\mathbf{x}) \tilde{f}(\alpha \mathbf{A}^{-1}\mathbf{x}) \tilde{f}(\beta \mathbf{A}^{-1}\mathbf{x}) d\mathbf{x}.$$

The change of variables $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ gives that $\mathbf{J}f' = \mathbf{A}(\mathbf{J}f)$.

The above proposition implies that by computing the transform values $\mathbf{J}f'(\alpha_i, \beta_i)$ and $\mathbf{J}f(\alpha_i, \beta_i)$ for at least two different pairs (α_i, β_i) one obtains a set of linear equations from which the transformation matrix \mathbf{A} may be solved. If more than two pairs are used, a linear least-squares solution of

$$\min_{\mathbf{A}} \sum_i \|\mathbf{J}f'(\alpha_i, \beta_i) - \mathbf{A}\mathbf{J}f(\alpha_i, \beta_i)\|^2 \quad (4)$$

can be computed. Thereafter the translation may be solved by $\mathbf{t} = \boldsymbol{\mu}(f') - \mathbf{A}\boldsymbol{\mu}(f)$.

In practice, with real images, it is better to use the normalized features $\mathbf{J}f'/If'$ and $\mathbf{J}f/If$ which are still related by the matrix \mathbf{A} since $If' = If$. Then it is sufficient that $f'(\mathbf{x}) = sf(\mathcal{A}^{-1}(\mathbf{x}))$, i.e., the intensity values in the affine transformed image may have different scale than in the original image. This leads to a better robustness against different illumination conditions.

The transform $f \mapsto \mathbf{J}f$ has the following symmetries, which are obtained from (3) by changes of variables.

- Proposition 3** (a) $\mathbf{J}f(\alpha, \beta) = \mathbf{J}f(\beta, \alpha)$,
(b) $\mathbf{J}f(\alpha, \beta) = \alpha^{-3}\mathbf{J}f(1/\alpha, \beta/\alpha)$ for $\alpha \neq 0$,
(c) $\mathbf{J}f(\alpha, \beta) = \beta^{-3}\mathbf{J}f(1/\beta, \alpha/\beta)$ for $\beta \neq 0$.

The symmetries may be used to show that it is enough to compute $\mathbf{J}f(\alpha, \beta)$ in the triangle T defined by the vertices $\{(-1, -1), (-1, 1), (1, 1)\}$. The values $\mathbf{J}f(\alpha, \beta)$ outside T may be computed from the values in T using the symmetries. The symmetries should be taken into account when choosing the points (α_i, β_i) in order to avoid redundancy.

Analogously to Remark 1 also the transform (3) may be generalized so that the transformation property stated in Proposition 3 still holds. Moreover, we may generalize it for functions defined in \mathbb{R}^d . Especially interesting is the case $d = 3$ since this allows to extend the above registration method to volume images. This is summarized in the following.

Proposition 4 Let $\mathcal{A} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ be an affine transformation so that $\mathcal{A}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{t}$, where \mathbf{A} is a nonsingular $d \times d$ matrix. Let $f \in L^\infty(\mathbb{R}^d)$ be compactly supported. The centroid of f is $\boldsymbol{\mu}(f) = \int_{\mathbb{R}^d} \mathbf{x}f(\mathbf{x}) d\mathbf{x} / \|f\|_{L^1}$ and the normalized version is defined by $\tilde{f}(\mathbf{x}) = f(\mathbf{x} + \boldsymbol{\mu}(f))$. Let $H : \mathbb{R}^k \rightarrow \mathbb{R}$ be a continuous function so that $H(\mathbf{0}) = 0$, and define

$$\hat{\mathbf{J}}f(\alpha_1, \dots, \alpha_k) = \frac{1}{\|f\|_{L^1}} \int_{\mathbb{R}^d} \mathbf{x}H(\tilde{f}(\alpha_1\mathbf{x}), \dots, \tilde{f}(\alpha_k\mathbf{x})) d\mathbf{x}. \quad (5)$$

Then $\hat{\mathbf{J}}(f \circ \mathcal{A}^{-1}) = \mathbf{A}(\hat{\mathbf{J}}f)$.

3.1 Euclidean and similarity transformations

Euclidean and similarity transformations are subgroups of affine transformations. In 2D, a general affine transformation has six degrees of freedom while a similarity transformation has four and a Euclidean transformation only three degrees of freedom. Hence, in such cases the matrix \mathbf{A} has to satisfy additional constraints and instead of a general nonsingular matrix we should estimate a transformation of the form

$$\mathbf{A} = s\mathbf{R}, \quad (6)$$

where \mathbf{R} is a rotation matrix and s is a scaling factor, which is equal to 1 for Euclidean transformations.

Nevertheless, the constraint (6) may be easily embedded to the registration method described above. There we estimated the affine transformation matrix \mathbf{A} by first computing a set of correspondences $\mathbf{J}f'(\alpha_i, \beta_i) \leftrightarrow \mathbf{J}f(\alpha_i, \beta_i)$ and then solving the resulting linear least-squares problem (4). However, the least-squares solution of (4) is possible in closed form also when the constraint (6) is forced. An algorithm for finding the least-squares solution when \mathbf{A} is forced to be a rotation matrix, i.e. $s = 1$, was described in [8] and it is based on the singular value decomposition. In [9] Umeyama made a correction to this method and presented a complete algorithm for the least-squares solution of similarity transformation parameters between two sets of d -dimensional points when the correspondences between the sets are known. Though we assume in (4) that there is no translation Umeyama's method can still be straightforwardly applied. Hence, we obtain the solution $\mathbf{A} = s\mathbf{R}$ and the translation is then again solved by $\mathbf{t} = \boldsymbol{\mu}(f') - \mathbf{A}\boldsymbol{\mu}(f)$.

3.2 Registration of point sets

Assume that instead of an image we have a set of points, \mathbf{x}_i , and an affine transformed version of it, $\mathbf{x}'_i = \mathcal{A}(\mathbf{x}_i) = \mathbf{A}\mathbf{x}_i + \mathbf{t}$, and we would like to solve the transformation between the point patterns without knowing the point correspondences.

We consider that the points \mathbf{x}_i are random samples of a random variable \mathbf{X} with some probability density p . Then the points \mathbf{x}'_i are samples of \mathbf{X}' which has density $p' = (p \circ \mathcal{A}^{-1}) / \|p \circ \mathcal{A}^{-1}\|_{L^1}$. We denote the mean and covariance of \mathbf{X} by $\boldsymbol{\mu}$ and \mathbf{C} , and those of \mathbf{X}' by $\boldsymbol{\mu}'$ and \mathbf{C}' . We define the normalized random variables by $\tilde{\mathbf{X}} = \mathbf{X} - \boldsymbol{\mu}$ and $\tilde{\mathbf{X}}' = \mathbf{X}' - \boldsymbol{\mu}'$ and the corresponding densities are denoted by \tilde{p} and \tilde{p}' . Then it holds that $\tilde{p}'(\mathbf{x}) = \tilde{p}(\mathbf{A}^{-1}\mathbf{x}) / |\det \mathbf{A}|$.

Furthermore, assume for the present that two functions, g and g' , are given so that $g'(\mathbf{x}) = sg(\mathbf{A}^{-1}\mathbf{x})$ where s is an arbitrary nonzero scale. Now, instead of (3) we consider descriptors defined for $\gamma \in \mathbb{R}$ as follows

$$\mathbf{H}(\gamma) = \frac{\int_{\mathbb{R}^2} \mathbf{x}g(\gamma\mathbf{x})\tilde{p}(\mathbf{x})d\mathbf{x}}{\int_{\mathbb{R}^2} g(\gamma\mathbf{x})\tilde{p}(\mathbf{x})d\mathbf{x}} = \frac{E[\tilde{\mathbf{X}}g(\gamma\tilde{\mathbf{X}})]}{E[g(\gamma\tilde{\mathbf{X}})]}, \quad (7)$$

where the integrals are interpreted as expectation values. In practice, as we do not know the function \tilde{p} we must compute (7) by using the sample means computed from the samples $\tilde{\mathbf{x}}_i = \mathbf{x}_i - \bar{\mathbf{x}}$, where $\bar{\mathbf{x}}$ is the centroid of the original points \mathbf{x}_i . The corresponding descriptors for the transformed pattern are $\mathbf{H}'(\gamma) = E[\tilde{\mathbf{X}}'g'(\gamma\tilde{\mathbf{X}}')]/E[g'(\gamma\tilde{\mathbf{X}}')]$ and it holds that $\mathbf{H}'(\gamma) = \mathbf{A}\mathbf{H}(\gamma)$. Thus, as above, a set of correspondences $\mathbf{H}'(\gamma_i) \leftrightarrow \mathbf{H}(\gamma_i)$ could be used to solve \mathbf{A} if we only knew suitable functions g and g' . Thereafter, the translation can be solved from $\mathbf{t} = \bar{\mathbf{x}}' - \mathbf{A}\bar{\mathbf{x}}$.

The problem in choosing g and g' is that we do not know \mathbf{A} beforehand. However, we may estimate the covariances \mathbf{C} and \mathbf{C}' as sample covariances from

the point sets $\{\mathbf{x}_i\}$ and $\{\mathbf{x}'_i\}$ and set

$$g(\mathbf{x}) = N(\mathbf{0}, \mathbf{C}), \quad g'(\mathbf{x}) = N(\mathbf{0}, \mathbf{C}'), \quad (8)$$

where $N(\mathbf{0}, \mathbf{C})$ is the zero mean Gaussian distribution with covariance \mathbf{C} . Then we have $g'(\mathbf{x}) = g(\mathbf{A}^{-1}\mathbf{x})/|\det \mathbf{A}|$ as required.

Thus, we have sketched an algorithm for solving the affine transformation between two point patterns. It should be noticed that in the noiseless case, $\mathbf{x}'_i = \mathbf{A}\mathbf{x}_i + \mathbf{t}$, the affine transformation is recovered exactly, up to a numeric round off error, but the above probabilistic framework provides a justification of the method also when there is noise in the point coordinates or some points in the other set have no counterparts in the other.

4 Implementation

4.1 Intensity-based method

In order to apply the proposed registration method for digital images we have to discretize the integral (3). This results in

$$\mathbf{J}f(\alpha, \beta) = \frac{1}{\sum_k f(\mathbf{x}_k)} \sum_k \mathbf{x}_k \tilde{f}(\mathbf{x}_k) \tilde{f}(\alpha \mathbf{x}_k) \tilde{f}(\beta \mathbf{x}_k), \quad (9)$$

where we need samples of \tilde{f} on three grids centered at the origin and having sample intervals $\{1, \alpha, \beta\}$. The sample values are interpolated from the original image f since we recall that \tilde{f} is a shifted version of f . In the experiments we used bilinear interpolation.

The descriptor values (9) are computed for some set of pairs (α_i, β_i) . It should be noted that with a suitable choice of these pairs one may select a single interpolation grid so that it will give samples needed for all used (α_i, β_i) . Hence, only one interpolation is needed to compute several descriptor values. The symmetries in Proposition 4 indicate that we would only need to take (α, β) s from the triangle $\{(-1, -1), (-1, 1), (1, 1)\}$, and we chose to take uniform sampling with sample interval 0.25 resulting in 45 different pairs (α_i, β_i) inside the triangle. In this case, the interpolation grid should also have the sample interval 0.25.

From (9) one can directly see that the computational complexity of the registration method is $O(N^2)$ for an $N \times N$ image. Finding samples of \tilde{f} requires interpolation to $N \times N$ points and the summation in (9) has N^2 terms which results in overall complexity $O(N^2)$.

4.2 Point-based method

Implementation of the point-based registration method of Section 3.2 is quite straightforward. We choose a set of values γ_i and compute the correspondences $\mathbf{H}'(\gamma_i) \leftrightarrow \mathbf{H}(\gamma_i)$ from which the affine transformation matrix is solved by least-squares. In the experiments we used values $\gamma_i = \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$. The algorithm is summarized in the following:

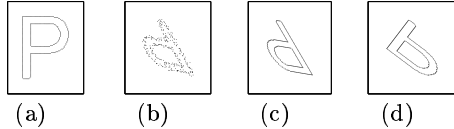


Fig. 1: Pattern matching: (a) original, (b) transformed and noise added ($\lambda = 0.06$), (c) recovered transformation with the new method ($\epsilon=0.12$), (d) with MD ($\epsilon=1.59$).

Table 1: Registration results for the point set in Fig. 1. The average values of the matching error ϵ at different levels of noise.

λ	0	0.02	0.04	0.06	0.08	0.10
MD	0.00	0.01	0.07	0.14	0.24	0.31
CW	0.00	0.05	0.10	0.16	0.23	0.30
H	0.00	0.04	0.09	0.13	0.18	0.23

1. Given two sets of points $\{\mathbf{x}_j\}$ and $\{\mathbf{x}'_j\}$ compute the corresponding sample means, $\bar{\mathbf{x}}$ and $\bar{\mathbf{x}}'$, and sample covariances, \mathbf{C} and \mathbf{C}' .
2. Compute the normalized points $\tilde{\mathbf{x}}_j = \mathbf{x}_j - \bar{\mathbf{x}}$ and $\tilde{\mathbf{x}}'_j = \mathbf{x}'_j - \bar{\mathbf{x}}'$. Set $g = N(\mathbf{0}, \mathbf{C})$ and $g' = N(\mathbf{0}, \mathbf{C}')$.
3. Compute the points $\mathbf{H}(\gamma_i)$ and $\mathbf{H}'(\gamma_i)$. The expectation values in (7) are computed as sample means by using the samples $\tilde{\mathbf{x}}_j$ and $\tilde{\mathbf{x}}'_j$.
4. Solve the affine transformation matrix \mathbf{A} from point correspondences $\mathbf{H}'(\gamma_i) \leftrightarrow \mathbf{H}(\gamma_i)$ by least-squares. If a Euclidean or a similarity transformation is wanted the additional constraints may be forced as described in Section 3.1. Finally, solve $\mathbf{t} = \bar{\mathbf{x}}' - \mathbf{A}\bar{\mathbf{x}}$.

It can be seen that the method is computationally quite simple. The computational complexity is only $O(n)$, where n is the number of points in the sets. For example, this is significantly better than the complexity of the cross-weighted moment method [5] which is $O(n^2)$.

5 Experiments

5.1 Point patterns

The first experiment was performed with the point pattern shown in Fig. 1(a). The points were transformed with random affine transformations and isotropic Gaussian noise was added to the coordinates before matching, cf. Fig. 1(b). The random transformation matrices were chosen according to

$$\mathbf{A} = \begin{pmatrix} \cos(\omega) & -\sin(\omega) \\ \sin(\omega) & \cos(\omega) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \kappa \end{pmatrix} \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix},$$

where $\omega, \phi \in [0, 2\pi]$ and $\kappa \in [0.3, 1]$ are uniformly distributed random variables. The standard deviation σ of the Gaussian noise was chosen to be proportional to the standard deviation of the x -coordinates of the original data points, i.e., $\sigma = \lambda\sigma_x$, where values $\lambda \in [0, 0.1]$ were used.

Patterns were matched with three different methods: affine moment descriptors (MD) [6], cross-weighted moments (CW) [5] and the new method (H). The cross-weighted moment method was implemented according to [5] with the parameter values $s_1 = 0.95$ and $s_2 = 0.99$ used by the authors in their experiments.

For each estimated transformation matrix $\hat{\mathbf{A}}$ we evaluated the distance ϵ to the true matrix \mathbf{A} by defining the points $\mathbf{p}_1 = (1, 0)^\top$ and $\mathbf{p}_2 = (0, 1)^\top$ and computing

$$\epsilon = \frac{1}{2} \sum_{i=1}^2 \frac{\|(\mathbf{A} - \hat{\mathbf{A}})\mathbf{p}_i\|}{\|\mathbf{A}\mathbf{p}_i\|}. \quad (10)$$

This is the measure we used to assess the matching result.

In Table 1 we have tabulated the average values of (10) among 1000 estimated transformations at different levels of noise. The results show that the new method seems to be most tolerant to noise. Although the moment descriptor method often gives a good result it sometimes badly fails as illustrated in Fig. 1(d). The new method and the cross-weighted moment method seem to behave more steadily.

5.2 Binary images

The third experiment was quite similar to the previous one but here we used the binary image shown in Fig. 2(a). The random transformations were obtained as above. The noise added to the transformed images was uniformly distributed binary noise with the noise level P indicating the probability of a single pixel to change its value. After adding the noise we removed the separated noise pixels from the background, cf. Fig. 2(b).

We did 500 random affine transformations and the average errors ϵ at different noise levels are shown in Table 2. Here the point-based variant of the new method (H) was applied so that the white pixels of the images were considered as 2D points. The method (J) is based on (9) and was implemented as described in Section 4.1.

The results show that the moment descriptor method works badly with this data. It works fine for some transformations but now it fails so often that also the average error is quite high. The intensity-based implementation of the new method does not work very well either. Perhaps one reason for this is that the intensity function of a binary image does not contain very much information that could be utilized in registration. Interestingly, the simple point-based implementation works best. In most cases the estimated transformation is reasonably close to the true one, although the errors are still notable.

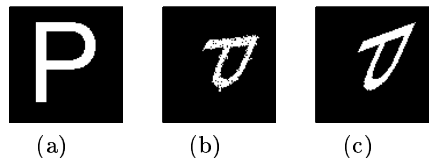


Fig. 2: Binary images: (a) original, (b) transformed and noise added ($P = 0.08$), (c) recovered transformation ($\epsilon = 0.26$).

Table 2: Results for the binary image with different levels of noise.

P	0	0.02	0.04	0.06	0.08	0.10
MD	0.47	0.62	0.67	0.74	0.74	0.79
CW	0.11	0.17	0.21	0.28	0.29	0.34
J	0.33	0.37	0.39	0.42	0.43	0.45
H	0.09	0.14	0.18	0.22	0.26	0.29

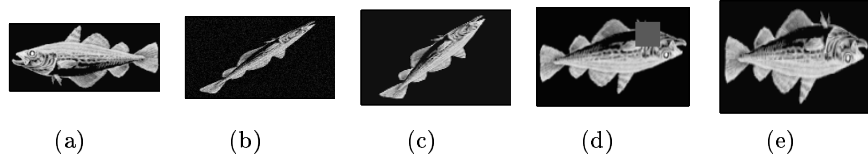


Fig. 3: Image registration: (a) original, (b) transformed and noise added (6%), (c) recovered transformation ($\epsilon = 0.19$), (d) transformed and occluded ($L=25$), (e) recovered transformation ($\epsilon = 0.13$).

5.3 Grayscale images

Next we tested our method by registering the grayscale image in Fig. 3(a) with its affine transformed versions, which were additionally corrupted by Gaussian noise (Fig. 3(b)) or square-shaped occlusion (Fig. 3(d)). The standard deviation of the Gaussian noise was chosen to be proportional to the maximum intensity of the image, i.e., $\sigma = \delta I_{\max}/100$, where values $\delta \in [0, 6]$ were used. The noise was added also to the black background of the transformed images, which may distort the intensity distribution and hence also the centroid. In the occlusion experiment gray squares with varying size were randomly placed in the image.

We did again 500 random affine transformations and the results are shown in Tables 3 and 4. Here we did not use the cross-weighted moment method since the number of pixels is large and the computation would have been very time consuming. It can be seen from Table 3 that the noise immediately causes the moment descriptor method (MD) to fail. This does not come as a surprise since the second- and third-order statistics are very sensitive to noise in the background [6]. The new method (J) seems to perform much better. Table 4 shows that in the occlusion experiment both methods manage to recover a transformation reasonably close to the true one when the size of the occlusion is not too large.

Finally, we experimented also real images. In Figs. 4(a) and 4(b) we have two views of a book. They were obtained by taking two digital photographs of the book on a dark background so that the book was rotated and translated between the exposures. The original images were decreased in size and segmented so that the background was set to have zero intensity value. The resulting 400×600 images are shown in Figs. 4(a) and 4(b). We estimated the similarity transformation between the two images with the new method, implemented as described in Sec. 4.1, and the registration result is illustrated in Fig. 4(c). We can see that an approximately correct transformation is recovered.

Table 3: Registration of noisy grayscale images. The average values of the matching error ϵ among 500 estimated transformations at six different levels of noise.

δ	0	0.5	1	2	4	6
MD	0.03	2.40	2.40	2.40	2.43	2.45
J	0.02	0.04	0.06	0.11	0.20	0.29

Table 4: Registration of occluded grayscale images. Parameter L denotes the side length of the occluding square.

L	5	9	15	19	25	31
MD	0.04	0.08	0.21	0.32	0.52	0.77
J	0.04	0.11	0.25	0.39	0.65	0.97

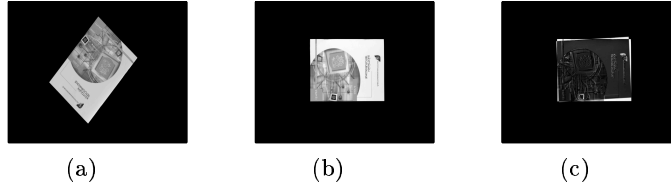


Fig. 4: Real images: (a) Image 1, (b) Image 2, notice the different illumination, (c) registration result.

6 Conclusions

We have proposed a novel method for affine registration of images and point patterns. The method is direct so that it does not require any separate feature extraction step and correspondence search. In addition, the method is non-iterative and efficient since the computational complexity is only linearly proportional to the number of pixels in the images or to the number of points in the patterns.

The particular way of utilizing the entire intensity information also implies that the method has some limitations. Ideally, both images should have similar content, and hence, images that have only partial overlap should be first segmented. However, the performed experiments showed that the method is relatively robust and gives an approximately correct transformation also when this assumption is slightly violated. Therefore, in the first place, the method might be used as an efficient way of computing an initial registration which is then refined by using other more complex registration methods, such as [2] or [3], for example.

References

1. Zitová, B., Flusser, J.: Image registration methods: a survey. *Image Vis. Comput.* **21** (2003) 977–1000
2. Fitzgibbon, A.W.: Robust registration of 2D and 3D point sets. In: *Proc. British Machine Vision Conference.* (2001)
3. Viola, P., Wells, W.M.: Alignment by maximization of mutual information. *International Journal of Computer Vision* (1997) 137–154
4. Rangarajan, A., Chui, H., Duncan, J.S.: Rigid point feature registration using mutual information. *Medical Image Analysis* **3** (1999) 425–440
5. Yang, Z., Cohen, F.: Cross-weighted moments and affine invariants for image registration and matching. *IEEE Trans. Pattern Anal. Mach. Intell.* **21** (1999) 804–814
6. Heikkilä, J.: Pattern matching with affine moment descriptors. *Pattern Recognition* **37** (2004) 1825–1834
7. Rahtu, E., Salo, M., Heikkilä, J.: A new efficient method for producing global affine invariants. Submitted (2005)
8. Arun, K.S., Huang, T.S., Blostein, S.D.: Least-squares fitting of two 3-D point sets. *IEEE Trans. Pattern Anal. Mach. Intell.* **9** (1987) 698–700
9. Umeyama, S.: Least-squares estimation of transformation parameters between two point patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **13** (1991) 376–380