AFFINE REGISTRATION USING MULTISCALE APPROACH

Esa Rahtu, Mikko Salo, and Janne Heikkilä

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

ABSTRACT

This paper introduces a multiscale framework for affine registration of image intensity functions. The idea in the introduced approach is to extend the given pattern to a set of affine covariant versions, each carrying slightly different information, and then to extract features for registration from each of them separately. The proposed approach is demonstrated by constructing several new affine registration methods by using different preprocessing techniques and combination schemes. The achieved registration methods are finally illustrated in two registration experiments involving real and synthetic images.

1. INTRODUCTION

In affine registration we are given two image intensity functions, which are known to be related by affine transformation. We are then asked to find the parameters of this mapping. In practice the affine relation does not necessarily hold exactly, but is a good approximation of the precise relation.

There are basically two main approaches to image registration, namely area and feature-based solutions [1]. The idea in area-based registration is to obtain the transformation parameters directly using standard optimization methods with a proper quality measure. This measure is usually computed by first mapping one pattern on the other, according to the current estimate, and then evaluating some function giving the alignment error. The choice of this function is maybe the most crucial step in the process and many different options have been proposed [1]. A simple example for such error function is the sum of absolute pointwise differences between the aligned patterns. Another possible example is a measure based on the mutual information between the patterns [2].

In the feature-based approach the transformation parameters are not estimated directly, but using salient features extracted from the patterns. The extraction of these features is the most important step here and many different ways to perform it have been introduced. The main difference in these methods is whether they find the features using some local information or the entire pattern at once.

Many examples of local feature extraction methods are listed by [1], where the possible features range from interest regions to different kinds of lines and points. The common characteristic to all these features is that they are extracted based on small local regions of the pattern, which means that they remain unaltered no matter what changes are made outside these neighborhoods. Another way to interpret the local feature extraction is to see it as a segmentation of distinguishable local characteristics.

In the context of the affine registration for images, the methods using the affine covariant neighborhoods [3] or interest points [4] are maybe the most frequently used. The registration using these techniques usually follow the three step algorithm:

1. Extract affine covariant neighborhoods or interest points from both images.
2. Find the correspondences between the extracted features.
3. Estimate the transformation parameters using the matched features.

The correspondence search in step 2 can be performed by applying any standard affine recognition technique, where the input patterns are now the extracted affine covariant areas or the neighborhoods around the interest points. The estimate of the transformation in the third step can then be done for example by minimizing the reprojection error between corresponding features [5]. This estimate is however not without problems, since there are usually false matches referred to as outliers in step 2. This may happen for example if the detected regions are small, or if the interest point matching is done using nonaffine covariant neighborhoods. The outliers can result in significant errors in the final estimate of the transformation parameters. One way to approach this problem is to apply the RANSAC algorithm [6], which has proved to be efficient in applications.

In general the local approaches perform well if the patterns have distinctive local characteristics which can be accurately extracted and reasonably well matched. These are normally fulfilled with standard photographs, but for example medical images usually lack such strong details [1].

Another option is to extract the features using the entire pattern. These so called global features have the characteristic that they do change if the pattern is altered at any place. Another common character is that using these techniques the correspondence matching, step 2 in the registration algorithm, is achieved automatically. This happens since each feature corresponds to a certain parameter configuration in the feature extraction method and hence we know that the features achieved with similar configurations in two patterns must be the corresponding ones. This gives the additional advantage that no outlier detection is required.

The simplest example of global approach is to use the image centroid as a feature. This would however only produce one point correspondence, which alone would be enough to determine the translation t. However, as it will be shown, we can use the multiscale framework to produce a set of images from which to compute enough corresponding centroids to solve the entire affine transformation. Other methods in this category include the cross weighted moments [7], the affine moment descriptors [8], the trace transform [9], and the affine invariant spectral signatures [10].

The main problems with the previously proposed global methods are the facts that they produce either only few reliable features or the evaluation of additional features is nontrivial to implement. The problem with the low number of features is most clearly illustrated in the case of image centroid, which does not even in theory produce enough information to recover the entire affine mapping. The problems with the implementation and evaluation are present with cross weighted moments, trace transform, and affine invariant spectral signatures.

In this paper we propose a simple way to extend previously proposed global affine registration methods to produce more features. The new approach is based on a multiscale framework, where the idea is to extend the given pattern to a set of affine covariant versions, each carrying slightly different information, and then to extract features for registration from each of them separately. The construction of the covariant set within the framework can be done in many different ways, offering a number of possibilities for variations. The proposed technique makes it possible to use even just the image centroid approach to estimate the affine transformations between image functions.

2. MULTISCALE FRAMEWORK

The basic idea in the multiscale approach can be presented as a three step algorithm:

1. Given an image f represent it in n different scales $f(\alpha_1 x), \ldots, f(\alpha_n x)$.
2. Combine the scaled images to a new image $Gf(x)$. The combination is required to be affine covariant, which means that for any affine transformation $\mathscr{A}$ one has
$$G(f \circ \mathscr{A}^{-1})(x) = (Gf)(\mathscr{A}^{-1}(x)),$$
where $\mathscr{A}(x) = Ax + t$, A is nonsingular 2×2 matrix, and $t \in \mathbf{R}^2$.
3. Perform affine registration, using f, Gf, and some known registration technique.

The advantage of the approach is that by varying the scales α_i and combinations G, we are able to generate a great variety of different descriptors. This is possible regardless of the inherent number of features offered by the descriptor used in the step 3.

The first step, scaling of images, is straightforward. The third step can also be quite simple depending on which methods are selected, but if we take the image centroids the implementation is not too difficult. The second step is often the most complicated one and is the key part of the whole approach. There one needs to take the scaled images $f(\alpha_i x)$ and combine them to a new image $Gf(x)$ so that Gf and $G(f \circ \mathscr{A}^{-1})$ are related with an affine transformation $\mathscr{A}$. The reason for this requirement is quite obvious, since we need to maintain the affine relationship of the new images $G_i f$ in order to get the methods in the step three to work.

With some combination schemes the translation component t of the affine transformation $\mathscr{A}$ can cause problems. In such case t can be estimated separately using the centroids of the original images, as follows. The examined images f and $g = f \circ \mathscr{A}^{-1}$ are normalized to produce $\tilde{f}(x) = f(x+\mu(f))$ and $\tilde{g}(x) = g(x+\mu(g))$, where $\mu(f)$ is the centroid of image f and $\mu(g)$ is the centroid of image g. These normalized versions can be shown to be connected as $\tilde{g}(x) = \tilde{f}(A^{-1}x)$. The multiscale methods are then applied using $\tilde{f}$ and $\tilde{g}$ where the estimated affine transformation is just the matrix A. The estimate $\hat{t}$ of the translation component is finally computed as $\hat{t} = \mu(g) - \hat{A}\mu(f)$, where the $\hat{A}$ is the estimate computed for matrix A.

3. MULTISCALE AFFINE REGISTRATION

In the previous section we defined the general ideas in the multiscale framework for registration. In this section we demonstrate the proposed framework though several examples. Since the main point in these examples is to present different combination schemes, the methods are constructed using the image centroid approach at the final step of the multiscale algorithm. Other approaches are straightforward to construct just by replacing the centroid computation with some other technique, such as cross weighted moments, the affine moment descriptors, the trace transform, and the affine invariant spectral signatures.

3.1. The first example

In this first example we present one of the simplest ways of using the multiscale framework in image registration. In this construction we take the combination method in the second step of the multiscale algorithm to be a pointwise product of f and two scaled representations of it. This method results in the same construction as in [15], but here the approach follows directly from the multiscale framework instead of the invariants presented in [11].

The pointwise product behaves well under affine transformation except for translation part, which must be considered separately using the technique presented in Section 2. Then in terms of normalized images $\tilde{f}$ and $\tilde{g}$, where $g = f \circ \mathscr{A}^{-1}$, we can write

$$G_{\alpha,\beta} f(x) = \tilde{f}(x)\tilde{f}(\alpha x)\tilde{f}(\beta x) \tag{1}$$

and

$$G_{\alpha,\beta} g(x) = \tilde{g}(x)\tilde{g}(\alpha x)\tilde{g}(\beta x)$$

where the scaling parameters $\alpha, \beta \in \mathbf{R}$. It can be shown that $G_{\alpha,\beta} g(x) = G_{\alpha,\beta} f(A^{-1}x)$ as required in the multiscale framework.

The matrix A can be then achieved using the centroids

$$\mu(G_{\alpha,\beta} f) = \frac{1}{\|f\|_{L^1}} \int_{\mathbf{R}^2} x G_{\alpha,\beta} f(x) dx$$

and

$$\mu(G_{\alpha,\beta} g) = \frac{1}{\|f\|_{L^1}} \int_{\mathbf{R}^2} x G_{\alpha,\beta} g(x) dx$$

with two α, β pairs. This is possible, since for both α, β pairs we have

$$\mu(G_{\alpha,\beta} g) = A\mu(G_{\alpha,\beta} f)$$

and the four terms in A can be solved from the resulting system of equations. We can also use more than two α, β pairs in which case the solution for A is achieved as a least squares solution.

The translation component is then solved as $\hat{t} = \mu(g) - \hat{A}\mu(f)$, where the $\hat{A}$ is the estimate computed for matrix A.

3.2. The second example

In the previous example we used pointwise product as a combination scheme in the multiscale framework. This resulted in simple construction, but the translation part had to be considered separately. Using convolutions in the combination as in the probabilistic approaches in [16] and [12] we can however avoid this.

Take the combination of the scaled images to be

$$G_{\alpha,\beta} f(x) = \frac{1}{\|f\|_{L^1}^2} f(x)(f_\alpha * f_\beta * f_\gamma)(x), \tag{2}$$

where $\alpha, \beta \in \mathbf{R}$, $\|f\|_{L^1}$ is the L^1 norm of f, $\gamma = 1 - \alpha - \beta$, $f_a(x) = a^{-2} f(x/a)$, and $*$ denotes convolution.

Making the same constructions using g and computing the centroids from both of them we have the relation

$$\mu(G_{\alpha,\beta} g) = A\mu(G_{\alpha,\beta} f) + t\frac{1}{\|f\|_{L^1}} \int_{\mathbf{R}^2} G_{\alpha,\beta} f(x) dx.$$

Taking three or more α, β pairs we have a linear system of equations, from which the A and t can be solved as a least squares solution. The evaluation can be done efficiently using fast Fourier transform as in [12].

3.3. Other combination schemes

In the last two sections we illustrated the multiscale framework in two different examples. The main difference in these approaches was the selected way to combine the scaled representations in the second step of multiscale algorithm. It may be observed that by changing only this part resulted in two quite different registration methods, which illustrates the flexibility of the proposed framework.

The presented two combination schemes are, however, not the only possibilities that can be constructed. In the following we present two more ways of performing the combinations using both nonlinear operations and adding an image transformation before the scaling. The first one of these uses similar constructions as the invariants in [13], where we replace the pointwise products in the previous examples by comparison operation

$$X(a, b) = \begin{cases} 1 & \text{if } a > b, \\ 0 & \text{otherwise.} \end{cases}$$

This results in two new formulations for Gf as

$$G_\alpha f(x) = X(f(x), f(\alpha x + (1 - \alpha)\mu(f))), \tag{3}$$

and

$$G_\alpha f(x) = X(f(x), \frac{1}{\|f\|_{L^1}}(f_\alpha * f_{1-\alpha})(x)), \tag{4}$$

where the scaling parameters $\alpha \in \mathbf{R} \setminus \{0\}$, and $f_a(x) = a^{-2} f(x/a)$. In these methods we used only one scaled representation instead of two in the other examples, since it

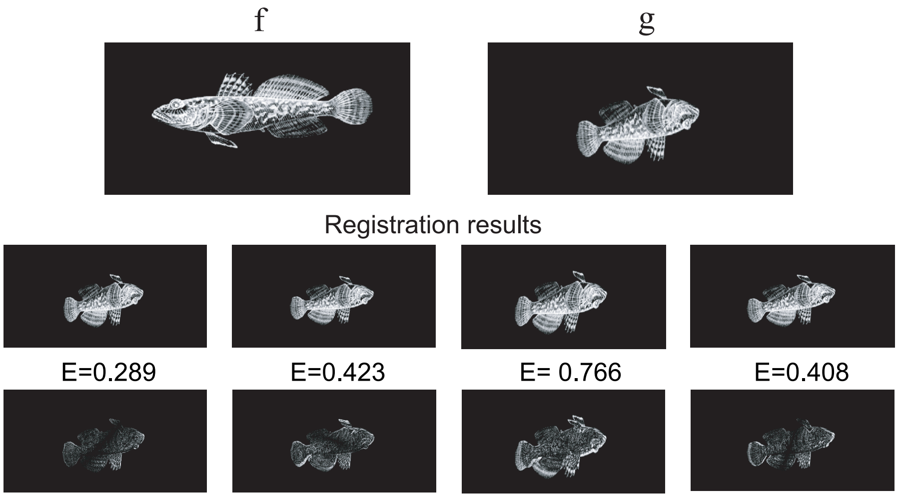

Fig. 1. The registration results with synthetic data. Top row has the original image f and the affine transformed version $g = f \circ \mathscr{A}^{-1}$. The second row has the images of f mapped using the estimated transformations. The third row illustrates the absolute pointwise differences between the results in row two and g. The relative sum of this difference is also written on top of each of these images. The Gfs used are from left to right (1), (2), (3), and (6).

was more natural to define the comparison operation only between two functions. Comparison, as a nonlinear operation, is expected to provide larger variations in the centroids computed from $G_\alpha f$ resulting in more stable registration.

In the basic form of the multiscale framework we are using scaled versions of the original image f. It is also possible to extend the approach by making a transform $\mathscr{T}$ to the image function f before applying the algorithm. Some extra caution must be paid here, since the relation between these new images is not likely to be the original affine transformation. However if we can find an inverse mapping from the resulting relation to $\mathscr{A}$, also the registration can be performed. Our next constructions give two examples of this, where the first one applies Fourier transform $\mathscr{F}$ as $\mathscr{T}$.

The Fourier transformed versions of f and $g = f \circ \mathscr{A}^{-1}$ are connected as

$$\hat{g}(\xi) = |det(A)|e^{-j2\pi t\cdot\xi}\hat{f}(A^t\xi), \tag{5}$$

where $\hat{g} = \mathscr{F}(g)$ and $\hat{f} = \mathscr{F}(f)$. From (5) one can see that if $t \neq 0$ and $det(A) \neq 1$ the resulting relationship is not affine and the application of the standard multiscale framework would fail.

In order to make this approach work we need to do something about these problems. The issue with the $det(A)$ is easily fixed by defining a normalized Fourier transform as

$$\hat{f}_n(\xi) = \frac{1}{\hat{f}(0)}\hat{f}(\xi).$$

The problem with the translation is however slightly more complicated. We could of course normalize the translation by computing the image centroid, but a better choice is to choose the combination G so that this effect is eliminated. This is done by defining

$$Gf_{\alpha_1,\ldots,\alpha_N}(\xi) = \hat{f}_n(\alpha_1\xi)\hat{f}_n(\alpha_2\xi)\ldots\hat{f}_n(\alpha_N\xi), \tag{6}$$

where we require that $\alpha_1 + \alpha_2 + \ldots + \alpha_N = 0$. Now it is easy to show that

$$G_{\alpha_1,\ldots,\alpha_N}g(\xi) = G_{\alpha_1,\ldots,\alpha_N}f(A^t\xi)$$

and evaluating moments with two or more sets of scales $\alpha_1, \ldots, \alpha_N$ results in enough equations to solve A^t and consequently A. The moments used in this estimation must be normalized differently. Define these moments as

$$\mu'(G_{\alpha,\beta}f) = \|f\|_{L^1}\int_{\mathbf{R}^2} xG_{\alpha,\beta}f(x)dx.$$

The translation part has to be again estimated as $\hat{t} = \mu(g) - \hat{A}\mu(f)$.

Another possibility besides the Fourier transformation is to use a Ridgelet transformation as $\mathscr{T}$, similarly as in [14]. The Ridgelet transformation is defined for image f as

$$Rf(\xi) = \int_{\mathbf{R}^2} f(x)\psi(x\cdot\xi)dx, \tag{7}$$

where $\xi \in \mathbf{R}^2$ and ψ is a wavelet. (7) behaves similarly as the Fourier transform under the affine transformation except for the translation part, which must be estimated separately.

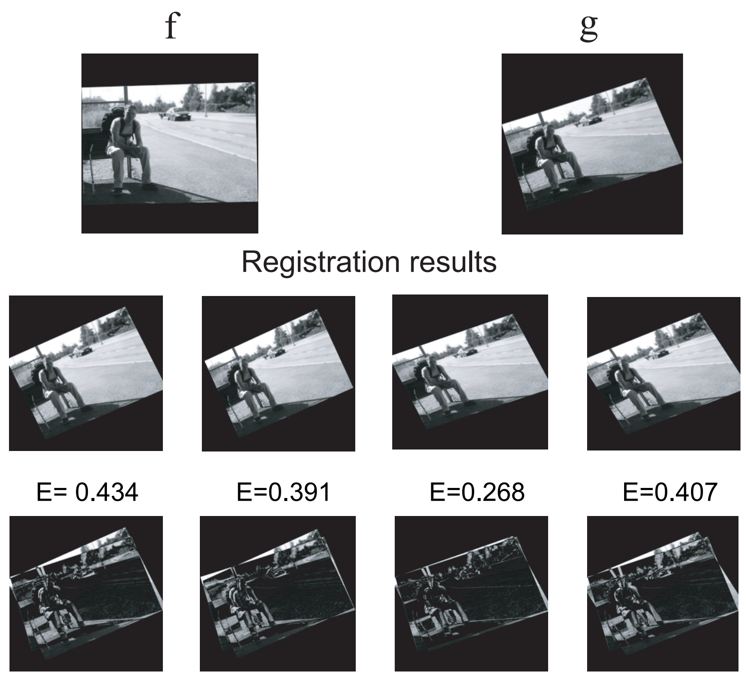

Fig. 2. The registration results with real data. Top row has the original images f and g taken from different viewing angles. The second row has the images of f mapped using the estimated transformations. The third row illustrates the absolute pointwise differences between the results in row two and g. The relative sum of this difference is also written on top of each of these images. The Gfs used are from left to right (1), (2), (3), and (6).

Using again the normalized image $\tilde{f}$ we can construct the corresponding Gf as

$$G_{\alpha,\beta}f(\xi) = \frac{1}{\|f\|_{L^1}^3} R\tilde{f}(\xi) R\tilde{f}(\alpha\xi) R\tilde{f}(\beta\xi), \qquad (8)$$

where the scaling parameters $\alpha, \beta \in \mathbf{R}$. Similarly as in the case of Fourier transform we have the relation

$$G_{\alpha,\beta}g(\xi) = G_{\alpha,\beta}f(A^t\xi).$$

so the determination of A and t proceeds in the same way.

These additional combination schemes illustrated two more approaches to apply the multiscale framework, which again result in very different registration methods compared to the first two examples.

4. EXPERIMENTS

In this section we present experiments indicating that the multiscale methods are applicable in image registration tasks. In the experiments we use both real and synthetic images and multiscale constructions based on Gfs in (1), (2), (3), and (6). More experiments involving similar constructions can be also found in [15, 16].

In the first experiment we take a gray scale image of fish, which we then affinely transform using a linear interpolation. Then we register the two images using each method and use the achieved estimate to transform the original image again. Figure 1 shows the original image and the registration results. One can observe that the registration is fairly accurate.

In the second experiment we have two segmented images of the same planar object taken from two different viewing angles. The registration task in this case becomes more complicated, since the transformation between the images is not exactly affine, but a projective one. Also due to photographing conditions there are several distortions in the images, including noise, illumination differences, radial distortion, etc. Figure 2 illustrates the registration results with different methods. It can be observed that also in this case the multiscale methods perform well.

5. CONCLUSIONS

In this paper we proposed a simple and efficient way of extending the existing affine registration methods to produce more features for registration. The new approach was based on the idea of a multiscale framework, where the original image is extended to an affine covariant set to which the standard registration methods can be applied. The affine covariant set makes it possible to use even simple methods like image centroids to fully recover the parameters of affine transformation. The key part of the multiscale framework is the construction of the covariant set, and several approaches for this were discussed in this paper. These constructions were inspired by similar constructions in affine invariant recognition methods. The results achieved in the experiments performed indicate that the multiscale approach is applicable in real registration tasks. The methods can also be used to obtain an initial guess for computationally more demanding iterative methods, in the cases where more accuracy is needed.

6. ACKNOWLEDGMENTS

This work was supported by the Academy of Finland and Infotech Oulu graduate school.

7. REFERENCES

[1] B Zitova and J Flusser, "Image registration methods: a survey," *Image and vision computing*, vol. 21, pp. 977–1000, 2003.

[2] P Viola and W M Wells, "Alignment by maximization of mutual information," *International journal of computer vision*, vol. 24, pp. 137–154, 1997.

[3] K Mikolajczyk, T Tuytelaars, C Schmid, A Zisserman, J Matas, F Schaffalitzky, T Kadir, and L Van Gool, "A comparison of affine region detectors," *International Journal of Computer Vision*, vol. 65, no. 1, pp. 43–72, 2005.

[4] K Mikolajczyk and C Schmid, "Scale & affine invariant interest point detectors," *International Journal of Computer Vision*, vol. 60, no. 1, 2004.

[5] R Hartley and A Zisserman, *Multiple view geometry in computer vision, second edition*, Cambridge university press, United Kindom, 2003.

[6] M A Fischler and R C Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of ACM*, vol. 24, no. 6, 1981.

[7] Z Yang and F S Cohen, "Cross-weighted moments and affine invariants for image registration and matching," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 21, no. 8, pp. 804–814, August 1999.

[8] J Heikkilä, "Pattern matching with affine moment descriptors," *Pattern recognition*, vol. 37, no. 9, 2004.

[9] A Kadyrov and M Petrou, "Affine parameter estimation from the trace transform," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 28, pp. 1631 – 1645, October 2006.

[10] J Ben-Arie and Z Wang, "Pictorial recognition of objects employing affine invariance in the frequency domain," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 6, pp. 604–618, June 1998.

[11] E Rahtu, M Salo, and J Heikkilä, "A new efficient method for producing global affine invariants," in *Proceedings of 3th International Conference on Image Analysis and Processing*, 2005, pp. 407–414.

[12] E Rahtu, M Salo, and J Heikkilä, "A new affine invariant image transform based on ridgelets," in *Proceedings of the 16th British Machine Vision Conference*, 2006, vol. 3, pp. 1059–1068.

[13] E Rahtu, M Salo, and J Heikkilä, "Nonlinear functionals in the construction of multiscale affine invariants," in *Proceedings of the 15th Scandinavian Conference on Image Analysis*, 2007, pp. 482–491.

[14] E Rahtu, M Salo and J Heikkilä, "Affine invariant pattern recognition using multi-scale autoconvolution," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 6, pp. 908–918, June 2005.

[15] J Kannala, E Rahtu, M Salo, and J Heikkilä, "A new method for affine registration of images and point sets," in *Proc. 14th Scandinavian Conference on Image Analysis*, Joensuu, Finland, 2005, pp. 224–234.

[16] J Kannala, E Rahtu, and J Heikkilä, "Affine registration with multi-scale autoconvolution," in *Proc. International Conference on Image Processing*, Genoa, Italy, 2005, vol. 3, pp. 1064–1067.