

Generalized affine moment invariants for object recognition

Esa Rahtu, Mikko Salo, Janne Heikkilä

Machine Vision Group

Dep. of Electrical and Information Engineering

P.O. Box 4500, 90014 Univ. of Oulu, Finland

{erahtu, msa, jth}@ee.oulu.fi

Jan Flusser

Institute of Information Theory and Automation

Academy of Sciences of the Czech Republic

182 08 Prague 8, Czech Republic

flusser@utia.cas.cz

Abstract

This paper introduces a new way of extracting affine invariant features from image functions. The presented approach is based on combining affine moment invariants (AMI) with multiscale invariants, in particular multiscale autoconvolution (MSA) and spatial multiscale affine invariants (SMA). Our approach includes all of these invariants as special cases, but also makes it possible to construct new ones. According to the performed experiments the introduced features provide discriminating information for affine invariant object classification, clearly outperforming standard AMI, MSA, and SMA.

1 Introduction

Affine invariant descriptors are important tools in object recognition problems. These techniques are commonly divided into two main categories according to how they make use of the image function. In so called local approaches the objects are segmented to smaller elements and invariants are computed separately for each of them. This class includes differential invariants [1] and patch based approaches [2]. The advantage with these methods is that the recognition can be carried out many times even if only a small part of the object is visible in the image. The drawback on the other hand is that the results become sensitive to the success of the segmentation step, and it is not always easy to find segments containing useful information about the image.

The second category consists of the global approaches, where the features are computed directly from the whole image intensity function. The advantage is that one does not have the problem of dividing the object into segments, and the features use information in the whole object instead of just small subparts. The drawbacks are that a large occlusion may disrupt the recognition and that the background of the objects must be eliminated, which may be difficult. Nevertheless, in many situations such as in industrial applications this can be done reliably. Also, it may be possible to apply the global methods in local patches, which could

improve the performance in the presence of occlusions.

Several global invariants have been introduced in the literature. The first method is the affine invariant moments (AMI) [3], [4], dating back to 1962. Later methods include cross-weighted moments [5], affine invariant spectral signatures [6], and trace transform [7]. However these appear to be either computationally expensive or difficult to implement. In addition to these, recent approaches include multiscale autoconvolution (MSA) [8] and spatial multiscale affine invariants (SMA) [9], where SMA is known to have some nonuniqueness.

In this paper we propose a new method for affine invariant feature extraction. In this approach we combine the basic ideas used in the AMI, MSA, and SMA methods, to form new invariants which provide more information than any of the methods separately. The presented formulation gives rise to numerous possibilities in constructing affine invariants. The performed experiments indicate that the new features enable reliable object recognition also under noisy conditions.

2 Pure moment invariants

Let $f \geq 0$ be an image intensity function on \mathbf{R}^2 . We want to consider affine invariants, which are functionals applied to image functions f and give the same value for f and any affine transformed version $f \circ \mathcal{A}$. Here $\mathcal{A}(x) = Ax + b$ is any affine transformation, where A is a 2×2 real nonsingular matrix and $b \in \mathbf{R}^2$.

We define $\tilde{x} = x - \mu(f)$ for $x = (x_1, x_2)^t \in \mathbf{R}^2$, where

$$\mu(f) = \frac{1}{\int_{\mathbf{R}^2} f(x) dx} \left(\int_{\mathbf{R}^2} x_1 f(x) dx, \int_{\mathbf{R}^2} x_2 f(x) dx \right)^t$$

is the image centroid. The central moments $m_{pq}(f)$ are given by

$$m_{pq}(f) = \int_{\mathbf{R}^2} \tilde{x}_1^p \tilde{x}_2^q f(x) dx.$$

We also define the cross product $C(x, y)$ for $x, y \in \mathbf{R}^2$ by

$$C(x, y) = x_1 y_2 - x_2 y_1.$$

Affine moment invariants are functionals of the form

$$If = \frac{1}{\|f\|_{L^1}^{w+N}} \int_{\mathbf{R}^{2N}} \prod_{1 \leq k < l \leq N} C(\tilde{x}_k, \tilde{x}_l)^{n_{kl}} \prod_{i=1}^N f(x_i) dx_i$$

where $N \geq 2$ and $n_{kl} \geq 0$ are integers, $w = \sum_{k,l} n_{kl}$, and there are N integrals over \mathbf{R}^2 . This is affine invariant because the cross products involve the shifted variables \tilde{x} , and because $C(Ax, Ay) = \det(A)C(x, y)$ for a nonsingular matrix A . If w is odd one needs to take the absolute value of If , but in this article w will always be even. Expanding the expressions $C(\tilde{x}_k, \tilde{x}_l)^{n_{kl}}$ shows that If is a polynomial of the central moments $m_{pq}(f)$. The simplest example is obtained when $N = 2$ and $n_{12} = 2$, so that If becomes

$$If = \frac{2}{\|f\|_{L^1}^4} (m_{20}m_{02} - m_{11}^2).$$

3 The new approach

Let $G(f, x)$ be a function, defined for admissible image functions f and points x in \mathbf{R}^2 , which satisfies the condition

$$G(f \circ \mathcal{A}, x) = G(f, \mathcal{A}(x)) \quad (1)$$

for any affine transformation \mathcal{A} . Given such a function G , the new invariants are defined as follows.

Definition. If f is an image function, define

$$If = \frac{1}{\|f\|_{L^1}^{w+N}} \int_{\mathbf{R}^{2N}} \prod_{1 \leq k < l \leq N} C(\tilde{x}_k, \tilde{x}_l)^{n_{kl}} \prod_{i=1}^N G(f, x_i) dx_i$$

where $N \geq 2$ and $n_{kl} \geq 0$ are integers, and $w = \sum_{k,l} n_{kl}$. If $N = 1$ the invariant is

$$If = \frac{1}{\|f\|_{L^1}} \int_{\mathbf{R}^2} G(f, x) dx.$$

Using the same argument as with affine invariant moments, we see that the new expression is affine invariant: $I(f \circ \mathcal{A}) = If$ for any affine transformation \mathcal{A} (again if w is odd then $|If|$ is invariant). Also, we define generalized central moments by

$$m_{pq}(G, f) = \int_{\mathbf{R}^2} \tilde{x}_1^p \tilde{x}_2^q G(f, x) dx.$$

It follows by expanding the expressions $C(\tilde{x}_k, \tilde{x}_l)^{n_{kl}}$ that If is a polynomial of the generalized central moments. This is very useful for the implementation of the new invariants, since one may compute several invariants from a finite set of generalized central moments. This also makes it possible to obtain invariants by replacing the central moments in AMIs by the generalized central moments, so one can use the well

developed theory for the AMIs as described in [10]. In fact, we will use the following five invariants corresponding to the ones in [10], where $p, q \in \{0, 1, 2, 3\}$. The m_{pq} will now be the generalized central moment $m_{pq}(G, f)$.

$$I_0 = m_{00}/\|f\|_{L^1},$$

$$I_1 = (m_{20}m_{02} - m_{11}^2)/\|f\|_{L^1}^4,$$

$$I_2 = (-m_{30}^2m_{03}^2 + 6m_{30}m_{21}m_{12}m_{03} - 4m_{30}m_{12}^3 - 4m_{21}^3m_{03} + 3m_{21}^2m_{12}^2)/\|f\|_{L^1}^{10},$$

$$I_3 = (m_{20}m_{21}m_{03} - m_{20}m_{12}^2 - m_{11}m_{30}m_{03} + m_{11}m_{21}m_{12} + m_{02}m_{30}m_{12} - m_{02}m_{21}^2)/\|f\|_{L^1}^7,$$

$$I_4 = (-m_{30}^3m_{03}^3 + 6m_{30}^2m_{11}m_{12}m_{03} - 3m_{30}^2m_{02}m_{12}^2 - 6m_{20}m_{11}^2m_{21}m_{03} - 6m_{20}m_{11}^2m_{12}^2 + 12m_{20}m_{11}m_{02}m_{21}m_{12} - 3m_{20}m_{02}^2m_{21}^2 + 2m_{31}^3m_{30}m_{03} + 6m_{31}^3m_{21}m_{12} - 6m_{11}^2m_{02}m_{30}m_{12} - 6m_{11}^2m_{02}m_{21}^2 + 6m_{11}m_{02}^2m_{30}m_{21} - m_{02}^3m_{30}^2)/\|f\|_{L^1}^{11}.$$

Next we discuss possible choices for the function G . The easiest choice is $G(f, x) = f(x)$, which leads to the standard AMIs.

Based on the multiscale approach and the results in [8], another promising choice is $G = M_{\alpha\beta}$ given by

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

where $\alpha, \beta \in \mathbf{R}$, $\gamma = 1 - \alpha - \beta$, and $f_a(x) = a^{-2}f(x/a)$ for $a \neq 0$ (and $f_a(x) = \|f\|_{L^1}\delta(x)$ if $a = 0$). This is related to the MSA transform considered in [8], where it is also shown that $M_{\alpha\beta}$ satisfies (1). The generalized moments are given by

$$m_{pq}(M_{\alpha\beta}, f) = \frac{1}{\|f\|_{L^1}^2} \int_{\mathbf{R}^2} \tilde{x}_1^p \tilde{x}_2^q f(x)(f_\alpha * f_\beta * f_\gamma)(x) dx.$$

This triple integral is computationally expensive to evaluate directly, but as in the case of MSA it is essential that we may write the convolution in terms of the Fourier transform $\hat{f}(\xi) = \int_{\mathbf{R}^2} e^{-2\pi i x \cdot \xi} f(x) dx$. Writing $h_{pq}(x) = \tilde{x}_1^p \tilde{x}_2^q f(x)$, the generalized moment will take the form

$$m_{pq}(M_{\alpha\beta}, f) = \frac{1}{\|f\|_{L^1}^2} \int_{\mathbf{R}^2} \hat{h}_{pq}(-\xi) \hat{f}(\alpha\xi) \hat{f}(\beta\xi) \hat{f}(\gamma\xi) d\xi.$$

The simplest invariant is

$$I_0 = \frac{m_{00}}{\|f\|_{L^1}} = \frac{1}{\|f\|_{L^1}^3} \int_{\mathbf{R}^2} \hat{f}(-\xi) \hat{f}(\alpha\xi) \hat{f}(\beta\xi) \hat{f}(\gamma\xi) d\xi$$

which is exactly the MSA transform.



Figure 1. (a) Samples of fish images. From left to right: original images, affine and Gaussian noise distorted images (std 10 % of maximum intensity), and affine distorted and occluded images (occlusion size 5 % of image side). (b) Samples of Coil-100 images at angle 0, 10, and 20 degrees.

The third choice corresponds to the spatial multiscale affine invariants (SMA) introduced in [9]. Here $G = S_{\alpha\beta}$ is defined for $\alpha, \beta \in \mathbf{R}$ by

$$S_{\alpha\beta}(f, x) = f(x)f(\alpha\tilde{x} + \mu(f))f(\beta\tilde{x} + \mu(f)).$$

It is easily seen that $S_{\alpha\beta}$ satisfies (1). The generalized central moments become

$$\begin{aligned} m_{pq}(S_{\alpha\beta}, f) &= \int_{\mathbf{R}^2} \tilde{x}_1^p \tilde{x}_2^q f(x)f(\alpha\tilde{x} + \mu(f))f(\beta\tilde{x} + \mu(f)) dx \\ &= \int_{\mathbf{R}^2} x_1^p x_2^q f(x + \mu(f))f(\alpha x + \mu(f))f(\beta x + \mu(f)). \end{aligned}$$

Again, $I_0 = m_{00}/\|f\|_{L^1}$ is just the SMA transform of f .

4 Experiments

Here we evaluate the proposed method in object classification tasks, comparing it to the traditional AMI, MSA, and SMA. Our goal is to show that the new features offer a good basis for affine invariant classification and that the new approach provides better performance than AMI, MSA, or SMA. We begin with implementational issues.

To evaluate the proposed features we need to construct the affine invariant polynomials and compute the moments $m_{pq}(M_{\alpha\beta}, f)$ and $m_{pq}(S_{\alpha\beta}, f)$ from the function f . The affine invariant polynomials are exactly the same as in the case of AMI, and in these experiments we have used the five polynomials I_0, I_1, I_2, I_3, I_4 . Also the implementation of $m_{pq}(M_{\alpha\beta}, f)$ and $m_{pq}(S_{\alpha\beta}, f)$ can be done directly following the same principles as in MSA [8] and SMA [9]. The implementation that we used can be retrieved from the URL: <http://www.ee.oulu.fi/research/imag/msa/>.

For the AMI we computed the features using 4 and 40 first invariant moment polynomials from [10] (in the case of AMIs I_0 does not carry any information and we do not use it). We will refer to these methods as AMI4 and AMI40, respectively. For the SMA we used the 36 (α, β) values given in [9] and for the MSA we took 37 (α, β) pairs from triangle $\{(1/3, 1/3), (1, 0), (1, 1)\}$ using uniform sampling.

We note that the chosen triangle is an equivalent to the one given in [8]. For the new invariants based on $m_{pq}(S_{\alpha\beta}, f)$ we used the same (α, β) pairs as in SMA case and similarly for the $m_{pq}(M_{\alpha\beta}, f)$ based invariants we used those given for MSA. This choice was done to see whether the new invariants really bring additional information compared to the original methods. We denote the new invariants as SMA moment and MSA moment, respectively.

As the first experiment we classified gray-scale images of 94 different fish. We used these original 94 images to train a nearest neighbor classifier for each method and then classified distorted versions of the same images. The distortions were obtained by first making a random affine transformation and then adding one of the following non-affine distortions: Gaussian noise, occlusion, and illumination distortion. In addition we also classified projective transformed versions of the images. The distortions and projective transformation were implemented as in [9], except that with Gaussian noise the resulting values were not limited between the minimum and maximum intensities of the original image. Some samples of the original images and the distorted versions are shown in Figure 1(a). The classification error rates with different distortions and distortion strengths are illustrated in Figure 2. These clearly demonstrate that the multiscale moment invariants offer discriminative information for reliable classification.

In addition to the experiment with fish images we assessed the methods in a recognition problem involving real 3D objects. For this purpose we took the Coil-100 image database of Columbia University as a test set. Coil-100 contains 100 different objects each viewed from 72 different angles, i.e. every 5 degrees. For the details of the database arrangements see the URL: <http://www.cs.columbia.edu>. In this classification task we trained again the nearest neighbor classifier using now the images taken at angle 0 and also four random affine transformed versions of them. The training set was made a bit larger in order to have some insight about the absolute performance. With more samples we were also able to use PCA decorrelation for the features.

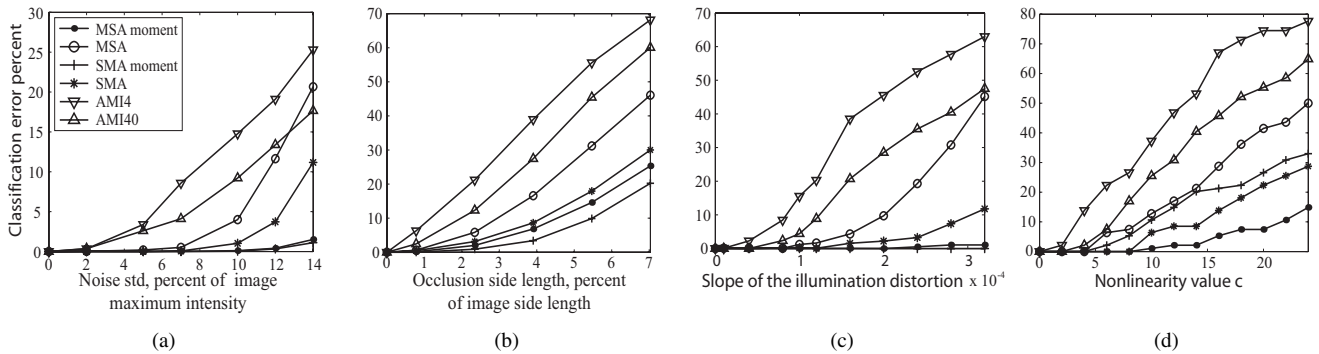


Figure 2. Classification error rates for (a) Gaussian noise, (b) occlusion, (c) illumination distortion, and (d) projective transformation.

From the resulting decorrelated features we took 30 most significant dimensions for MSA and MSA moments, 25 dimensions for SMA, SMA moments. The test set was taken to include all the objects at angles from -25° to $+25^\circ$. Samples of these images are given in Figure 1(b). Table 1 shows the error rates with different viewing angles. It can be observed that the new approach provides reliable results also with real view angle changes, clearly outperforming corresponding MSA, SMA, and AMI invariants.

5 Conclusions

In this paper we have proposed a new framework for constructing affine invariant features from images. The method combines multiscale representations with standard affine moment invariants, and as specific instances of this we constructed new invariants based on the MSA and SMA transforms. We assessed the invariants in object classification tasks and compared them to other similar methods. The results clearly indicate that the new techniques offer discriminating information also under strong nonaffine distortions and in the case of real three dimensional objects. In addition to the invariants considered in this paper, we expect that the presented framework can be used for finding also other new affine invariant constructions.

Table 1. Error rates for the Coil-100 classification with different view angles.

Angle	MSA		SMA		AMI4	AMI40
	mom.	MSA	mom.	SMA		
-25°	8 %	16 %	24 %	28 %	54 %	39 %
-20°	4 %	8 %	11 %	13 %	53 %	28 %
-15°	1 %	3 %	8 %	8 %	45 %	16 %
-10°	0 %	0 %	3 %	4 %	28 %	9 %
-5°	0 %	0 %	0 %	0 %	17 %	2 %
5°	0 %	0 %	0 %	0 %	20 %	3 %
10°	0 %	1 %	3 %	4 %	47 %	9 %
15°	3 %	5 %	8 %	10 %	55 %	16 %
20°	8 %	8 %	20 %	18 %	64 %	24 %
25°	12 %	17 %	32 %	34 %	68 %	30 %

6 Acknowledgements

This work was funded by the Academy of Finland project No. 110751. J. Flusser was supported by the Grant Agency of the Czech Republic under the project No. 102/04/0155.

References

- [1] I. Weiss. Projective invariants of shapes. *Proc. Image Understanding Workshop*, 1125–1134, 1988.
- [2] F. Rothganger, S. Lazebnik, C. Schmid, and J. Ponce. Object modeling and recognition using local affine-invariant image descriptors and multi-view spatial constraints. *Submitted to International Journal of Computer Vision*, 2005. Accepted.
- [3] M. Hu. Visual pattern recognition by moment invariants. *IEEE Trans. Information Theory*, 8:179–187, 1962.
- [4] J. Flusser and T. Suk. Pattern recognition by affine moment invariants. *Pattern Recognition*, 26(1):167–174, January 1993.
- [5] Z. Yang and F. Cohen. Cross-weighted moments and affine invariants for image registration and matching. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 21(8):804–814, August 1999.
- [6] J. Ben-Arie and Z. Wang. Pictorial recognition of objects employing affine invariance in the frequency domain. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 20(6):604–618, June 1998.
- [7] M. Petrou and A. Kadyrov. Affine invariant features from the trace transform. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 26(1):30–44, January 2004.
- [8] E. Rahtu, M. Salo, and J. Heikkilä. Affine invariant pattern recognition using multiscale autoconvolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(6):908–918, June 2005.
- [9] E. Rahtu, M. Salo, and J. Heikkilä. A new efficient method for producing global affine invariants. In *Proc. International Conference on Image Analysis and Processing*, pages 407–414, Cagliari, Italy, 2005.
- [10] T. Suk and J. Flusser. Graph method for generating affine moment invariants. In *Proc. International Conference on Pattern Recognition*, volume 2, pages 192–195, Cambridge, England, 2004.