

Robust recognition of planar shapes under affine transforms using Principal Component Analysis

Georgios Tzimiropoulos, Nikolaos Mitianoudis and Tania Stathaki

Abstract—A scheme, based on Principal Component Analysis (PCA), is proposed that can be used for the recognition of 2D planar shapes under affine transformations. A PCA step is first used to map the object boundary to its canonical form, reducing the problem of the non-uniform sampling of the object contour introduced by the affine transformation. Then, a PCA-based scheme is employed to train a set of basis functions on the signals extracted from the objects' boundaries. The derived bases are used to analyze the boundary locally. Based on the theory of invariants and local boundary analysis, a novel invariant function is constructed. The performance of the proposed framework is compared with a standard wavelet-based approach with promising results.

Index Terms—Principal Component Analysis, affine transformation, invariants, shape recognition

EDICS Category: IMD-STOR

I. INTRODUCTION

ASSUME that a collection of objects of interest is stored in a database and that we desire to design an algorithm which incorporates shape information, as extracted from the objects' boundaries, with the goal of identification in an unknown environment. Obviously, a desired property of a robust recognition system is *invariance* to the object shape deformation, caused by the arbitrary camera viewpoint positions. To tackle the problem, a common simplification made by the computer vision community is to approximate shape variation with an affine transformation. Let $\mathbf{c}(t) = [c_x(t), c_y(t)]^T$ denote a parametric closed curve in 2D space, representing the boundary of an object, where t is an arbitrary parameter used for curve parameterization (e.g. the arc length). An affine transformation models scaling, rotation, shearing and translation of the object boundary as follows:

$$\mathbf{c}'(t') = A\mathbf{c}(t) + \mathbf{b} \quad (1)$$

where $\mathbf{c}'(t') = [c'_x(t'), c'_y(t')]^T$ is the affine transformed curve, A is a 2×2 nonsingular matrix and \mathbf{b} is a 2×1 vector. The matrix A can be decomposed as follows:

$$A = s \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix} \quad (2)$$

where $s \in R^+$ models global scaling, $\theta \in [0, 2\pi)$ models rotation and $\alpha \in R$ is the shearing parameter. The vector \mathbf{b} represents translation. The transformed parameter t' , which is, in general, a function of t , reflects the problem of sampling appropriately the object contour.

Suppose we compute a quantity I from \mathbf{c} and the corresponding quantity I' from \mathbf{c}' . If I and I' are related as follows:

$$I' = \mu I \quad (3)$$

where $\mu \neq 0$ is a constant, then I is an affine invariant. If $\mu = 1$, I is called an *absolute invariant*, otherwise I is called a *relative invariant*.

Signal processing techniques, used to derive features from the boundary representation, that remain invariant under affine transformations, have played a key role in shape-based recognition algorithms. Classical shape analysis includes Fourier Descriptors [1], moments [2] and matched filtering [3]. Recently, methods based on the dyadic wavelet transform, have been proposed, which are reported to achieve state-of-the-art performance. The object boundary is analyzed at different scales, yielding to the approximation and the detail signals, which are then used for the construction of affine invariant functions. The choice of the signals, the number of decomposition levels and the wavelet functions used, have all resulted in a number of different approaches [4], [5], [6], [7], [8]. Unfortunately, the performance of the above methods is strongly affected by the non-uniform sampling introduced by the boundary extraction process and the parameter transformation problem described above.

In this work, we propose a scheme, based on Principal Component Analysis (PCA), that can be used for object identification under large deformations and noise. Using PCA, the affine transformed boundary is first mapped to its canonical form. This step preserves object shape information and, at the same time, provides an efficient way to tackle the problem of the non-uniform sampling of the object contour. Then, the object boundary is analyzed locally, similarly to the wavelet analysis, using basis functions, derived from a PCA-based training scheme and the boundaries of the objects of interest. Based on the theory of determinants, an invariant function is constructed, which can be used for object identification under affine transforms.

II. SHAPE RECOGNITION USING PCA

A. Robust boundary encoding using PCA

In this section, the problem of an appropriate parameterization of the object boundary is considered. As mentioned earlier, to preserve the linearity of the affine transformation, the object contour must be parameterized using a parameter that transforms linearly under affine transformation. In this case, a one-to-one correspondence between equally spaced, in terms of the affine invariant parameter, points on the original

contour and points on its affine transformed version can be established [4]. The most popular parameter used is the enclosed area defined as [1]:

$$\sigma = \int_a^b |x\dot{y} - y\dot{x}|dt \quad (4)$$

where \dot{x} , \dot{y} are the first-order derivatives of $x(t)$ and $y(t)$ with respect to t and the integration interval $[a, b]$ denotes a segment along the curve. It transforms linearly under affine transformation only if $b = 0$. For this purpose, the origin of the coordinate system is set to the object area center [1]. Unfortunately, the computation of both the enclosed area and the area center is very sensitive to large deformations and noise, resulting in deterioration of the performance of the matching process.

To reduce the effect of the parameterization problem, we use Principal Component Analysis to map the affine transformed shape to its canonical form. For this purpose, \mathbf{c}' is first normalized to zero mean and, then, its covariance matrix, defined as $R_{c'} = E\{\mathbf{c}'\mathbf{c}'^T\}$, is estimated. Let H be the matrix containing the eigenvectors of $R_{c'}$ and D the diagonal matrix containing the eigenvalues of $R_{c'}$, such that the i -th diagonal element corresponds to the i -th column of H . Then, the canonical form \mathbf{z}' of the affine distorted shape \mathbf{c}' is given by:

$$\mathbf{z}' = W\mathbf{c}' \quad (5)$$

where W is the whitening matrix defined as $W = D^{-1/2}H^T$. It can be shown that \mathbf{z}' is uncorrelated and of unit variance, that is $E\{\mathbf{z}'\mathbf{z}'^T\} = I$. Using the same principle, suppose that we compute the canonical form \mathbf{z} of the original contour \mathbf{c} . Then, it can be shown that \mathbf{z}' and \mathbf{z} are related as follows:

$$\mathbf{z}' = \Theta\mathbf{z} \quad (6)$$

where Θ is a 2×2 orthogonal matrix representing rotation or mirror rotation and reflection [9]. Note that the whitening transform preserves all the shape information of \mathbf{c}' . This is simply due to the fact that Eq.(5) can be seen as an affine transformation itself. We further mention that the accurate computation of the second moment matrix $R_{c'}$ also requires an appropriate affine invariant curve parameterization. However, from our experiments, we have observed that the shape principal axes can be restored successfully, under large deformations and noise, without employing the enclosed area parameter.

Once the object has been brought to its canonical form, we can encode the object boundary using a set of equally spaced points, in terms of the arc length, whose computation is simpler and more accurate. In this way, the stability and robustness of the invariant features, derived from the object boundary, are much less affected by the parameter transformation problem.

B. An invariant function using PCA bases

We have considered, so far, a PCA-based preprocessing step, which can be used to reduce the effect of the non-uniform sampling of the object contour, introduced by random affine transformations. We will now describe a methodology to derive kernels that can be employed for shape analysis, based

on Principal Component Analysis and the boundaries of the object of interest.

We assume that all objects are normalized to their canonical form and represented by a set of N_c equally spaced points, in terms of the arc length parameter. Let us denote by $\mathbf{s}(n) = [s_x(n), s_y(n)]^T$ any curve segment of length N_s along the boundary of any of our objects and let also \mathbf{s}_{xy} be a $N_s \times 1$ vector representing either $s_x(n)$ or $s_y(n)$. Our target is to estimate a set of appropriate basis functions to represent \mathbf{s}_{xy} . Principal Component Analysis can be used to train a set of uncorrelated bases on the curve segments extracted from the objects' boundaries, by optimizing an energy compaction mechanism. It has been used for feature extraction and dimensionality reduction in a wide range of applications [10], [11]. Since it operates on the signal second-order statistics, it implicitly assumes a Gaussian distribution of the data. However, this Gaussian profile imposed by PCA seems quite reasonable due to the low-pass nature of the signals obtained from the objects' boundaries.

Curve segments from all objects are placed in a matrix S as follows:

$$S = [\mathbf{s}_{xy1}, \mathbf{s}_{xy2}, \dots, \mathbf{s}_{xyK}] \quad (7)$$

where K is the total number of curve segments. We note that, when constructing the matrix S , to ensure signal's stationarity and capture all local structure, the contour of each object should be divided into a number of overlapping segments. In the usual way, we compute the eigenvalue decomposition of the covariance matrix of S , defined as $R_S = SS^T/K$. If U is the matrix containing the eigenvectors of R_S and Y is the diagonal matrix containing the eigenvalues of R_S , such that the i -th diagonal element corresponds to the i -th column of U , then, the PCA bases are defined as the rows of the following matrix V :

$$V = Y^{-1/2}U^T \quad (8)$$

Overall, a total of N_s basis functions is provided by the scheme described above. However, one can easily form a reduced set of bases that can be used for an efficient data representation with only a little loss of information. This is because the significance of the derived bases is indicated by the corresponding eigenvalues. Therefore, this information can be used, through selecting the most significant basis functions, to retain most of the signals structure and, at the same time, to perform considerable noise reduction. In fact, this is an advantage of our approach over the popular wavelet-based methods, which, depending on the application, entail a thorough investigation of the wavelet decomposition tree in order to identify the wavelet and pick the levels which yield the best possible performance [4]. In contrary, the PCA analysis functions are directly estimated from the actual data, along with a measure of significance, which simplifies the selection process. The eight most important PCA bases, for the application of aircraft silhouette identification (considered in the following section), along with their frequency content can be seen in Fig. 1. It can be observed that the boundary is analyzed locally using a series of bandpass filters, similarly to the wavelet transform. We note that the function corresponding to the largest eigenvalue is not presented, since it represents a

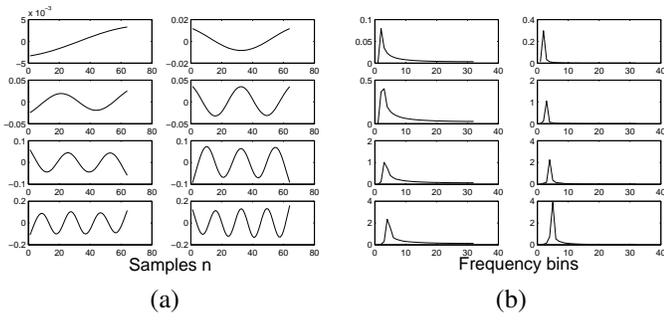


Fig. 1. (a) The derived PCA bases and (b) their frequency content.

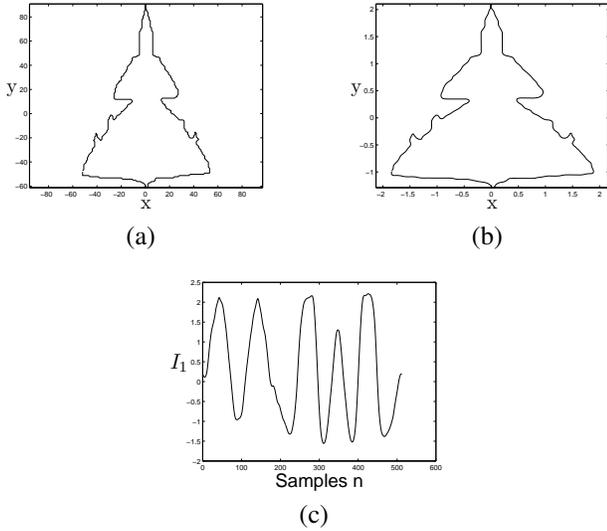


Fig. 2. (a) An aircraft model, (b) its canonical form and (c) the corresponding invariant function I_1 .

DC component, i.e. a change in signal level, DC , and, therefore, is of no significance.

Let us denote by $P_i\varphi$ the set of coefficients obtained by projecting the signal φ to the i -th basis function, using a sliding window approach. By applying the i -th and the j -th kernel to the object boundary, one can form:

$$\begin{bmatrix} P_i z'_x & P_j z'_x \\ P_i z'_y & P_j z'_y \end{bmatrix} = \Theta \begin{bmatrix} P_i z_x & P_j z_x \\ P_i z_y & P_j z_y \end{bmatrix} \quad (9)$$

Taking the determinants in both parts yields:

$$\begin{aligned} I'_1(i, j) &= P_i z'_x P_j z'_y - P_j z'_x P_i z'_y \\ &= \det(\Theta)(P_i z_x P_j z_y - P_j z_x P_i z_y) = \pm I_1(i, j) \end{aligned} \quad (10)$$

since $\det(\Theta) = \pm 1$. Therefore, the function I_1 is a relative invariant with $\mu = \pm 1$. Note that the above derivation has been used to provide invariants for the case of affine transformations, while in our case, the matrix R represents only rotation. However, it stills provides an efficient way to embody the transform coefficients, obtained by applying two different kernels to the object boundary, in a single function. Fig. 2 shows the boundary extracted from an aircraft model, its canonical form and the corresponding derived invariant function I_1 .

The degree of similarity between two objects κ and λ can be measured as the maximum value of the normalized circular cross-correlation [6]:

$$C_{1,\kappa,\lambda}(m) = \frac{\sum_m \sum_n I_{1,\kappa}(n) I_{1,\lambda}(n-m)}{\sqrt{\sum_n I_{1,\kappa}(n)^2 \sum_n I_{1,\lambda}(n)^2}} \quad (11)$$

where $I_{1,\kappa}$ and $I_{1,\lambda}$ are the invariant functions derived from the objects κ and λ respectively and $I_{1,\lambda}$ is circularly shifted. If the maximum absolute value is used, then the sign ambiguity introduced by Eq. (10) will be removed and the function I_1 can be considered as an absolute invariant. The circular cross-correlation is used to reduce the effect of the unknown shift between the starting points of the two contours [5]. This is because both $c(t)$ and $c(t+t_0)$, where t_0 is any shift in the origin of the object boundary, essentially represent the same object.

III. RESULTS

To evaluate the performance of our approach, the method is applied to the within-class object recognition application of aircraft silhouette identification. Fig. 3 shows the contours of $K = 20$ aircraft models which have also been used in [6] to test the discrimination power of the invariant function and its ability to capture small variations. To model a 128×128 image resolution, the resolution of all model images in Fig. 3 is such that each aircraft approximately fits in a $d_1 \times d_2$ rectangular grid, where $d_1 d_2 = 128^2$. From these images, the aircraft contours are extracted, using a simple 8-point connectivity algorithm [12]. Each model is then transformed to its canonical form and represented by a set of $N_c = 512$ equally spaced points, in terms of the arc length parameter. The length of the curve segment N_s is chosen to be equal to 64 points. The PCA basis functions are extracted, and, for each model, the invariant function I_1 is computed using the two most significant PCA bases, as described in the previous section. For each aircraft model, a set of test images is generated, which depict the same model under large deformations. We have used the following affine parameters: $\theta = \{0^\circ, 60^\circ, 120^\circ, 180^\circ\}$ and $\alpha = \{2, 3\}$. Similarly, the boundaries are extracted and, for each test boundary, the corresponding invariant function is derived.

The proposed scheme is compared with a popular wavelet-based affine invariant function [6], denoted as I_2 . For the computation of I_2 , we first parameterize the object boundary using the signed enclosed area parameterization and a scheme based on the methods proposed in [3] and [4]. We do not claim that our implementation is optimum, however, for each model, we have used the parameters (see [3]) that appear to yield the best possible performance (On the other hand, in our method, all parameters remain fixed and independent of the model under examination). The signed enclosed area parameter is employed, since it is found to be less sensitive to noise [13]. The invariant function I_2 is computed at scales (5, 6) and (6, 7), which provide the best trade off between discrimination capability and robustness to noise. The quadratic B-spline wavelet has been used, which is reported to give the best results [8].

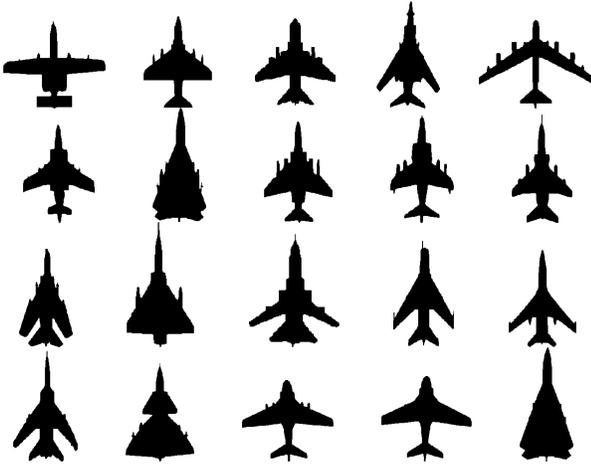


Fig. 3. The aircraft models used in the experiment.

The performance of the two invariant functions is evaluated in a noisy environment. Uniformly distributed noise is artificially added to the x and y coordinates of each contour point, after the extraction of the boundary from each test image. The amount of noise added is controlled using the Signal to Noise Ratio (SNR) defined in [4]. We have considered a large noise level of $\text{SNR} = 20$ dB [6]. For each test image, the experiment is repeated 100 times. The classification results are given in Table I. As it can be observed, the proposed algorithm features robust performance and appears to be much more stable than the standard wavelet-based method.

Model	I_1	$I_2(5, 6)$	$I_2(6, 7)$
1	0.9788	0.9550	0.9650
2	0.9350	0.7700	0.8250
3	0.9387	0.7475	0.7740
4	0.9575	0.7912	0.6588
5	0.9613	0.9500	0.9038
6	0.9950	0.8288	0.9912
7	0.9475	0.7488	0.8050
8	0.9375	0.7462	0.7537
9	0.9688	0.7188	0.9838
10	0.8938	0.7700	0.7363
11	0.9800	0.8988	0.9725
12	0.9925	0.9813	0.9900
13	0.9537	0.8075	0.6700
14	0.9575	0.9312	0.9163
15	0.9988	0.7650	0.9963
16	0.9375	0.6875	0.8762
17	0.9988	0.9975	0.9975
18	0.8887	0.7713	0.7100
19	0.9475	0.7400	0.7138
20	0.9538	0.8125	0.8150
Overall	0.9561	0.8209	0.8527

TABLE I

RECOGNITION RATE USING THE PROPOSED FUNCTION I_1 AND THE WAVELET-BASED FUNCTION I_2 .

IV. CONCLUSIONS

We have presented a PCA-based framework with the goal of robust shape-based object recognition under affine transformations. The role of Principal Component Analysis, in

the proposed scheme, is twofold. First, it is used to map the affine transform shape to its canonical form in order to tackle the problem of sampling the object contour appropriately. Then, it is used to train a set of basis functions with desired properties on the signals extracted from the object boundaries. The derived bases functions are then used for the construction of a novel invariant function. The proposed framework is applied to the problem of aircraft silhouette identification and compared with a popular wavelet-based method. Simulation results show that our method outperforms the wavelet-based method. In addition, the extra problem of identifying the most appropriate wavelet and decomposition levels does not exist in the proposed scheme.

V. ACKNOWLEDGEMENTS

The authors would like to thank Dr. I.E. Rube for his advice and help to produce some of the ideas presented in this work.

REFERENCES

- [1] K. Arbter, W.E. Snyder, H. Burkhardt, and G. Hirzinger, "Application of affine-invariant fourier descriptors to recognition of 3-d objects," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 7, pp. 640–647, 1990.
- [2] A. Zhao and J. Chen, "Affine curve moment invariants for shape recognition," *Pattern Recognition*, vol. 30, no. 6, pp. 896–901, 1997.
- [3] D. Cyganski and R.F. Vaz, "A linear signal decomposition approach to affine invariant contour identification," *Pattern Recognition*, vol. 28, pp. 1845–1853, 1995.
- [4] Q.M. Tieng and W.W. Boles, "Wavelet-based affine invariant representation: a tool for recognizing objects in 3d space," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, no. 8, pp. 846–857, 1997.
- [5] R. Alferéz and Y.F. Wang, "Geometric and illumination invariants for object recognition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 21, no. 6, pp. 505–536, 1999.
- [6] M.I. Khalil and M.M. Bayoumi, "A dyadic wavelet affine invariant function for 2-d shape recognition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 23, no. 10, pp. 1152–1164, 2001.
- [7] M.I. Khalil and M.M. Bayoumi, "Affine invariants for object recognition using the wavelet transform," *Pattern Recognition Letters*, vol. 23, pp. 57–72, 2002.
- [8] I.E. Rube, M. Ahmed, and M. Kamel, "Wavelet approximation-based affine invariant shape representation functions," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 28, no. 2, pp. 323–327, 2006.
- [9] J. Heikkilä, "Pattern matching with affine moment descriptors," *Pattern Recognition*, vol. 37, 2004.
- [10] J. Karhunen and J. Joutsensalo, "Generalization of principal component analysis, optimization problems and neural networks," *Neural Networks*, vol. 8, no. 4, pp. 549–562, 1995.
- [11] I.T. Jolliffe, *Principal Component Analysis*, Springer-Verlag, 1986.
- [12] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis and Machine Vision*, PWS, second edition, 1999.
- [13] K. Arbter, *Affinvariante Fourierdeskriptoren ebener Kurven*, Ph.D. thesis, Technische Universität Hamburg-Harburg, Germany, 1990.