AN AFFINE INVARIANT FUNCTION USING PCA BASES WITH AN APPLICATION TO WITHIN-CLASS OBJECT RECOGNITION

Georgios Tzimiropoulos, Nikolaos Mitianoudis and Tania Stathaki

Imperial College, Exhibition Road, SW7 2AZ London, UK

ABSTRACT

The problem of shape-based recognition of objects under affine transformations is considered. We focus on the construction of a robust and highly discriminative affine invariant function that can be used for within-class object recognition applications. Using the boundaries of the objects of interest, a training scheme, based on Principal Component Analysis (PCA), is proposed to derive a set of basis functions with desired properties. The derived bases are then used for the construction of a novel affine invariant function. The proposed invariant function is evaluated for the problem of aircraft silhouette identification and appears to achieve comparable performance to a popular wavelet-based affine invariant function. At the same time, the proposed framework is much simpler than that based on wavelet analysis.

Index Terms— affine transformation, shape-based object recognition, PCA bases

1. INTRODUCTION

Shape information, extracted from objects' boundaries, as depicted in $2D$ images, has played a crucial role in pattern recognition applications in computer vision. An important problem, related to all methods involving shape analysis, arises from the fact that the shape of an object varies due to arbitrary camera viewpoint positions. In all cases, the shape variation can be modeled by the perspective transformation. However, when the object is far from the camera, the shape distortion can be approximately described by the affine transformation which is linear. The affine transformation models possible scaling, rotation, translation and shearing.

A number of methods have been proposed, which compute descriptors from the boundary representation that remain invariant under affine transformation. Recently, methods based on wavelet analysis have become popular in shape representation and matching, since they appear to outperform tradition techniques such as Fourier analysis and moments. The main idea is to analyze the object boundary using the dyadic wavelet transform. The decomposition yields to the approximation and detail signals which uniquely characterize the boundary. These signals are then used to construct affine invariant functions. The choice of the signals, the decomposition levels and the wavelet functions used, have all resulted in a number of different approaches [1], [2], [3], [4], [5].

In this work, we propose a linear generative model and *Principal Component Analysis* (PCA) to derive kernels (bases) that can be used for shape analysis in within-class object recognition applications. The bases are obtained by a simple training process using the boundaries of the objects of interest. They can be used to analyze the boundary locally similarly to the wavelet analysis. In contrary to the wavelet-based methods, our approach does not require a thorough investigation of the parameters used (for example the choice of the mother wavelet and the dyadic levels) to optimize performance for a specific application. The significance of the obtained analysis kernels is naturally derived from the training process. Based on the theory of invariants and PCA bases analysis, an affine invariant function is constructed, that can be used for the recognition of the objects of interest under affine transformation.

2. BACKROUND

Let $\mathbf{c}(t) = [x(t), y(t)]^T$ denote a parametric closed curve in $2D$ space, where t is an arbitrary parameter used for curve parameterization (e.g. the arc length). If the curve has undergone an affine transformation, then it can be described by the following parametric equation:

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

where $\mathbf{c}'(t') = [x'(t'), y'(t')]^T$ 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:

$$A = \left[\begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \end{array} \right] = s \left[\begin{array}{cc} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{array} \right] \left[\begin{array}{cc} 1 & \alpha \\ 0 & 1 \end{array} \right] \quad (2)$$

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' is, in general, a function of t.

To preserve the linearity of the affine transformation, the object boundary must be parameterized using a parameter, which transforms linearly under affine transformations [1]. In

this work, we will assume $t' \approx t + t_0$, where t is the arc length parameter and t_0 represents a shift between the starting points of the two contours, and we will consider the parameter transformation problem as an additional factor which deteriorates the performance of the matching process [5]. The effect of translation is also eliminated by setting the origin of the coordinate system to the centroid of the object contour (i.e. $\mathbf{b} = \underline{0}$). Taking the above into consideration and ignoring, for the time being, the effect of t_0, Eq. (1) reduces to:

$$\mathbf{c}'(t) = A\mathbf{c}(t) \tag{3}$$

Let I be a quantity computed from $\mathbf{c}$ and I' be the same quantity computed from $\mathbf{c}'$. If I is an affine invariant, then it is related to I' as follows:

$$I' = \mu I \tag{4}$$

where $\mu \neq 0$ is a constant. If $\mu = 1$, then I is called an absolute invariant, otherwise I is called a relative invariant. In the following section, a relative invariant function will be derived using the proposed methodology.

3. AN AFFINE INVARIANT FUNCTION USING PCA BASES

In literature, a number of techniques have been proposed, which derive invariants from the boundary representation of objects, operating in the transform domain. Fourier analysis and wavelet decomposition are two well-studied examples [6], [1], [2], [3], [4], [5]. In these cases, the bases, used for boundary analysis, are mathematically well-defined to serve some specific analysis tasks. However, in many applications, it is useful to derive the transform from the data itself. A number of arbitrary bases can be obtained by a training process using the set of the objects of interest.

Assume a collection of objects stored in our database and let $\mathbf{c}(n) = [x(n), y(n)]^T$ be the sampled parametric boundary representation of any of these objects. All objects are normalized to be of equal length N_c. We also denote by $\mathbf{c}_s(n) = [x_s(n), y_s(n)]^T$ any segment along the curve $\mathbf{c}$. Then, the whole curve can be seen as a concatenation of curve segments which, in this work, are chosen to be of equal length N_s. Let $r_s(n)$ represent either $x_s(n)$ or $y_s(n)$. If we assume that r_s follows a linear generative model, then it can be written as a linear combination of L basis functions as follows:

$$\mathbf{r}_s = F\mathbf{u}_s \tag{5}$$

where $\mathbf{r}_s$ is a $N_s \times 1$ vector, $F = [\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_L]$ is a $N_s \times L$ matrix whose i-th column corresponds to the basis function $\mathbf{f}_i$ and $\mathbf{u}_s = [u_{s1}, u_{s2}, \ldots, u_{sL}]^T$ is a $L \times 1$ vector whose components are scalar constants. The matrix F represents the synthesis kernels (bases), while the matrix $G = F^{-1} = [\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_L]^T$ gathers the analysis kernels. Our main target is to compute a set of appropriate basis functions. Since $\mathbf{r}_s$ is of length N_s, we essentially need N_s basis functions.

In this work, a set of uncorrelated bases is obtained using Principal Component Analysis (PCA) [7]. PCA provides an optimal solution to several signal representation problems, making it useful for feature extraction and dimensionality reduction in a wide range of applications [8], [7]. Its main drawback is that it assumes a Gaussian distribution of the data. However, it is not unreasonable to assume that the low-pass nature signals, obtained from the objects' boundaries, can fit the Gaussian profile imposed by PCA.

Let us define by $R = [\mathbf{r}_{s1}, \mathbf{r}_{s2}, \ldots]$ the matrix obtained by gathering all curve segments from all objects. The PCA bases are derived from the eigenvalue decomposition of the covariance matrix of R, C_R. Let H be the matrix containing the eigenvectors of C_R and D the diagonal matrix containing the eigenvalues of C_R, such that the i-th diagonal element corresponds to the i-th column of H. Then, the PCA bases are defined as the rows of the following matrix V:

$$V = D^{-1/2}H^T \tag{6}$$

An important benefit gained from PCA analysis is that a measure of the significance of each basis function is given by the corresponding eigenvalue. One can discard the functions corresponding to the smallest eigenvalues to obtain a reduced set of L bases $\tilde{V}$. This reduced set can be used for a compact data representation. The function corresponding to the largest eigenvalue is discarded as well, since it represents a DC component, i.e. a change in signal level.

In this work, we are interested in boundary analysis rather than compact representations. Likewise, the most significant bases will be used to capture most of the signals structure and, at the same time, to perform significant noise reduction. Similarly to the wavelet transform, the object contour is analyzed locally, using a series of bandpass filters. The eight most important PCA bases, for the application of aircraft contour identification (considered in the following section), along with their frequency content can be seen in Fig. 1. However, in our case, there is no need for an exhaustive investigation to identify the wavelet or the decomposition levels which should be used for optimum performance. Such a procedure is described in detail in [1]. Instead, the analysis functions are directly provided by the data itself, together with a measure which indicates their significance. The boundary is analyzed using the analysis kernels and a set of coefficients is generated, which are used for the construction of an affine invariant function.

Let $P_i s$ be the coefficients obtained by projecting the signal s to the i-th analysis kernel. Since the transform lacks the shift-invariance property, a sliding window approach should be followed. By applying the transform to both parts of Eq. (3), we get:

$$\begin{aligned} P_i x' &= a_{11} P_i y + a_{12} P_i y \\ P_i y' &= a_{21} P_i y + a_{22} P_i y \end{aligned} \tag{7}$$

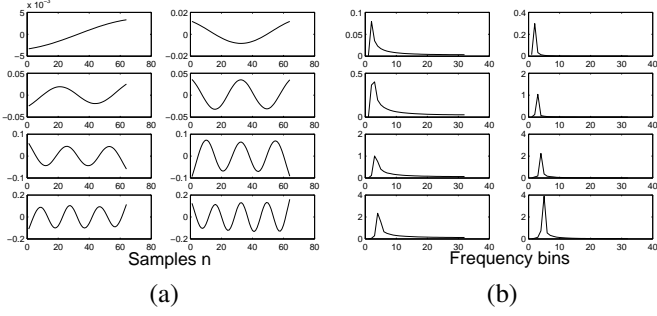Likewise, using the coefficients obtained by applying the j-th

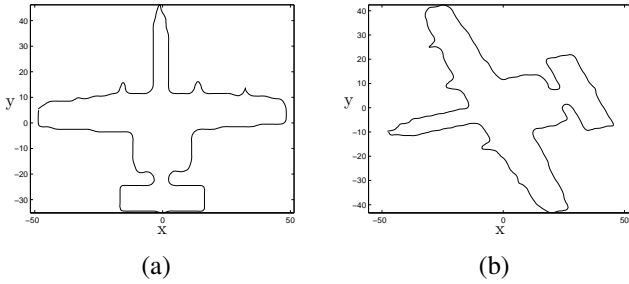**Fig. 1**. (a) The derived PCA bases ($N_s = 64$) and (b) their frequency content.

Fig. 2. (a) Original boundary and (b) boundary under affine transformation

analysis kernel, one can form:

$$\left[\begin{array}{cc} P_i x' & P_j x' \\ P_i y' & P_j y' \end{array}\right] = A \left[\begin{array}{cc} P_i x & P_j x \\ P_i y & P_j y \end{array}\right] \qquad (8)$$

Taking the determinants in both parts yields:

$$\begin{aligned} I_1'(i,j) &= P_i x' P_j y' - P_j x' P_i y' \\ &= det(A)(P_i x P_j y - P_j x P_i y) = det(A) I_1(i,j) \end{aligned} \qquad (9)$$

Therefore, the function I_1 is a relative invariant with $\mu = det(A)$. It can be made an absolute invariant by dividing it with another relative invariant function or its maximum value [5]. Fig. 2a and 2b show the boundary of an aircraft model and the boundary of its affine transformed version respectively. The corresponding derived invariant functions I_1 and I_1' can be seen in Fig. 3, which demonstrates their invariance to affine transformations.

4. RESULTS

To evaluate the algorithm performance, the method is applied to a database of $K = 20$ aircraft models which can be seen in Fig. 4. These models have also been used in [3] to test the discriminative power of the invariant function and its ability to capture small variations. Since scale transformations
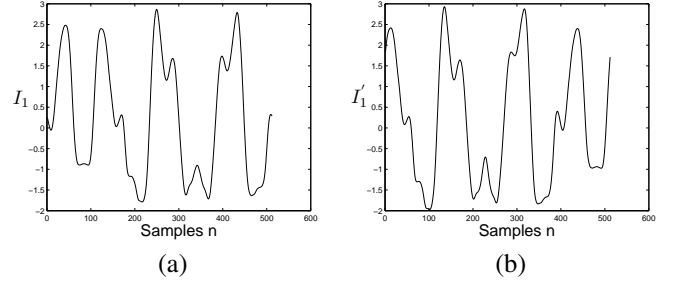

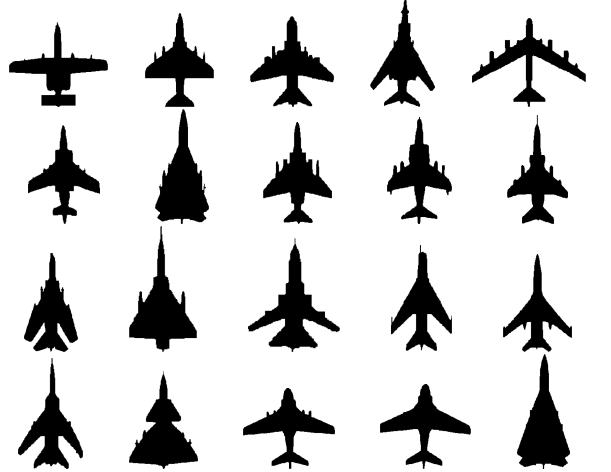**Fig. 3**. The derived invariant functions (a) I_1 and (b) I_1'.

Fig. 4. The aircraft models used in the experiment.

depend on the resolution of the original image, in our experiments, the effect of scaling is not investigated. Instead, to model a 128×64 image resolution, the resolution of all model images in Fig. 4 is such that each aircraft approximately fits in a $d_1 \times d_2$ rectangular grid, where $d_1 d_2 = 128^2/2$. From these images, the aircraft contours are extracted using a simple $8-$point connectivity algorithm [9] and normalized to have equal length $N_c = 512$ points. The length of the curve segment N_s is chosen to be equal to 64 points. We also note that, when constructing the matrix R, we allow a small overlap of the curve segments to ensure signal's stationarity. 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 model image, a set of test images is generated using the following affine parameters: $\theta = \{0°, 60°, 120°, 180°\}$ and $\alpha = \{0, 1/3, 2/3, 1\}$. Similarly, the boundaries are extracted using the same $8-$point connectivity algorithm and, for each test boundary, the corresponding invariant function is derived.

For two invariant functions, $I_{1,\kappa}$ and $I_{1,\lambda}$, derived from objects κ and λ respectively, the degree of similarity is measured as the maximum value of the normalized circular cross-

correlation [3]:

$$S_{1,\kappa,\lambda}(m) = \frac{\sum_m \sum_n I_{1,\kappa}(n) I_{1,\lambda}(n-m)}{\sqrt{\sum_n I_{1,\kappa}(n)^2 \sum_n I_{1,\lambda}(n)^2}} \quad (10)$$

The use of the circular cross-correlation also eliminates the effect of the unknown shift between the starting points of the two contours.

The performance of the proposed invariant function 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 [1]. We have considered a large noise level of SNR $= 20$ dB. For each test image, the experiment is repeated 100 times. The classification results are given in Table 1. In the same table, we also present the recognition rate obtained by applying a popular wavelet-based affine invariant function [3], denoted as I_2. The invariant function I_2 is computed at scales $(5,6)$ and $(6,7)$ using the quadratic B-spline wavelet which is reported to yield the best performance [5]. As it can be observed, the proposed algorithm features comparable performance to the standard wavelet-based method. We further note that any linear combination of invariant functions (derived using more than two PCA bases) can be used to increase the recognition rate with the cost of additional computational complexity.

5. CONCLUSIONS

We have presented a methodology which aims at the construction of a robust affine invariant function that can be used for within-class object recognition applications. The proposed framework is based on PCA basis functions trained from the boundaries of the objects of interest. Simulation results show that the derived function appears to have similar performance to a popular wavelet-based affine invariant function. In addition, the developed framework is much simpler than that based on wavelet analysis.

6. ACKNOWLEDGEMENTS

7. REFERENCES

[1] 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.

[2] R. Alferez 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.

Model	I_1	$I_2(5,6)$	$I_2(6,7)$
1	1.0000	1.0000	0.9487
2	1.0000	1.0000	1.0000
3	0.9981	1.0000	0.9387
4	0.7988	0.7863	0.8125
5	0.6600	0.6856	1.0000
6	0.9531	0.7837	0.8163
7	0.9838	0.9525	0.8425
8	1.0000	1.0000	1.0000
9	0.9144	0.8500	0.7500
10	0.8781	0.9081	1.0000
11	1.0000	1.0000	1.0000
12	1.0000	1.0000	1.0000
13	0.9931	0.9775	0.7500
14	0.5000	0.4944	0.3538
15	0.6806	0.6656	0.6300
16	0.6806	0.6900	0.5200
17	0.8962	1.0000	1.0000
18	1.0000	1.0000	1.0000
19	0.9862	1.0000	0.8200
20	0.9481	0.8775	1.0000
Overall	0.8936	0.8836	0.8595

Table 1. Recognition rate using the proposed function I_1 and the wavelet-based function I_2.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] I.T. Joliffe, *Principal Component Analysis*, Springer-Verlag, 1986.

[8] J. Karhunen and J. Joutsensalo, "Generalization of principal component analysis, optimization problems and neural networks," *Neural Netwoks*, vol. 8, no. 4, pp. 549–562, 1995.

[9] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis and Machine Vision*, PWS, second edition, 1999.