FUSION IN PHASE SPACE FOR SHAPE RETRIEVAL

Foteini Fotopoulou+, Ilias Theodorakopoulos, George Economou**

+ Department of Computer Engineering and Informatics, University of Patras, 26500, Patras, Greece
phone: (+30)2610 996058, fax: +(0030) 2610997456, email: fotopoulou@ceid.upatras.gr

* Department of Physics, University of Patras, 26500, Patras, Greece
phone: (+30)2610 996059, fax: +(0030) 2610997456, email: economou@physics.upatras.gr

ABSTRACT

2-D shape recognition and content-based image retrieval using the boundary information as a 1-D time series is a common way of addressing the shape-matching problem. A significant factor for successful object identification is insensitivity to basic shape deformations, such as scaling, shifting, rotation, partial distortions etc. In this paper, a method for handling the time series boundary information as a set of feature vectors in the reconstructed phase space is proposed. Besides enhanced discrimination, projection in phase space facilitates the fusion of several boundary descriptors as well. Optimal information extraction from the multivariate time series is achieved by proper embedding dimension and time-lag parameter selection. Centroid contour distance (CCD) and Angle Sequence (AS), capturing both local and global shape information in a circular manner, are the descriptors employed in the fusion stage. Invariance to rotations and insensitivity to partial deformations is achieved while retrieval performance is enhanced. The method is applied and evaluated in MPEG-7 part B shape database using early or late fusion.

1. INTRODUCTION

Content based 2-D shape retrieval is an important issue in computer vision. It requires efficient descriptors that take into account the possible shapes' deformations, as well as an effective comparison method. Shape descriptors are divided into three main categories: contour based, image based and skeleton based [1], [2]. A plethora of shape descriptors can be found in literature. Descriptors that originate from the contour [3], [4] have been enriched with the presence of geodetic distance [5], aiming to include information from the shape interior. In the Shape Context [3] method a discrete set of points sampled from the contours serve for shape description, while the method is improved [6] by the adoption of inner distances, in place of Euclidian (IDSC). Our method is also contour based, implementing a discrete time series representation of the pixels that belong to the shape outline. The above shape contour – time series modeling is quite common in shape research as it makes distance computation and similarity assessment between shapes easier. In [7] the scaling invariant problem in boundary image matching is addressed in the time series domain instead of image domain. In [8] the rotation insensitivity problem is addressed, using time series conversion of shapes while in the present work deduction of a time series enables the shape description in phase space.

In this paper multivariate time series are extracted from 1-D contour point sequences, in order to use reconstructed phase space [9]. From the serial 1-D data, a set of embedding vectors is formed. In this representation the time lag parameter and the embedding dimension have to be found. Their selection, which is mostly influential to algorithm performance, is studied in this work. Proper combination of these parameters can capture both global and local information of the original sequence. The phase space method is applied in two well known shape descriptors, the Angle Sequence (AS) [10] and the Centroid Contour Distance (CCD) curve [11].

The idea of phase space reconstruction from a 1-D series, as well as a multivariate statistical method for comparing multidimensional points, has already been tested on the specific shapes of leaves database (subset of Smithsonian database) in a forthcoming publication by the present authors [13]. The phase space approach to the 1-D series representation, analysis and processing has also been used in several other cases [12],[14],[15]. In this work the phase space concept is farther utilized in combined use for improving shape retrieval.

Phase space representation is suitable for the fusion of various descriptors originating from differently parameterized time series or descriptors that encode diverse information. Fusion of descriptors can be implemented either in the first stage, by concatenating characteristics that provide different shape information (early fusion), or in the final stage utilizing the dissimilarity between two shapes (late fusion). It is the scope of this work to demonstrate that the richness of information available to each separate descriptor, combined with the advantages of phase space representation, result in improving shape retrieval performance.

The rest of the paper is organized as follows. In section 2 the core of the multidimensional phase space method is described. The two phase space reconstruction parameters, i.e. time lag parameter and embedding dimension are studied and the matching stage is explained. The main part which focuses on the fusion process is described in section 3 and experimental results for evaluation of the method are given in section 4. Conclusions and future work are given in section 5.

2. PHASE SPACE REPRESENTATION OF TIME SERIES

The transformation of a 1-D sequence into a set of multidimensional points is a well known signal analysis technique encountered in the field of non-linear dynamics. There, using appropriate delays and dimension on the time series data, the reconstructed phase space can provide valuable information about the nature and dynamics of the underlying generating system. For the present work, borrowing the same basic ‘phase space’ methodology and terminology we examine the 1-D series signal that comes from the shape’s contour which can be considered containing significant information about the hidden variables behind its formation.

2.1 Feature extraction

The feature extraction stems from the mapping of a time sequence to phase space using a certain time lag and an embedding dimension. The resulting phase space reconstruction provides the multivariate trajectory and the associated set of vectors that form the shape feature.

Let x(n) be the time sequence consisting of p points. Then the transformation into phase space is achieved using the following equation (1), where **X(n)** is an m-dimensional vector:

$$\mathbf{X(n)} = \{x(n)\ x(n+1\cdot T)\ \ldots\ x(n+(m-1)\cdot T)\}\ ,\ n=1:p\ . \qquad (1)$$

T is the time-lag parameter, and m is the embedding dimension. Therefore from each x(n) sequence, a set of p, m-dimensional vectors is produced: **[X(1) X(2) ... X(p)]**

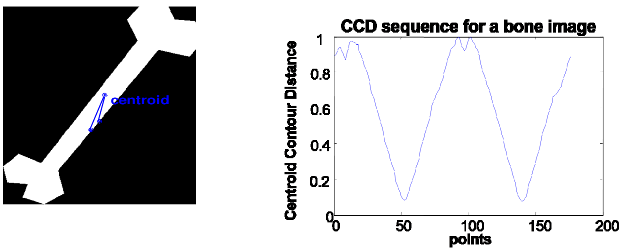

Figure 1 – CCD calculation & its Centroid Contour Distance

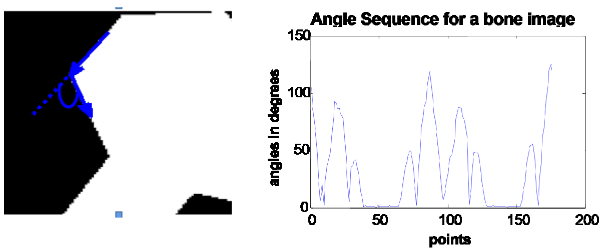

Figure 2 –Angle calculation & a bone's angle sequence.

Several point sequences are eligible for representing the shape boundary. No constraint is imposed, and the selection is dictated by the specific application. In this work the Centroid Contour Distance (CCD), and Angle Sequence (AS), two well known descriptors are selected [10],[11]. CCD is a descriptor formed by consecutive boundary -to- centroid Euclidean distances (Figure 1). It has been extensively used in shape description [4],[2],[16], with the main drawback being its sensitivity to shape deformations. Adopting proper normalizations it can be made scale invariant, while translation invariance is innate. The CCD sequence does not possess rotation invariance, resulting in the need of defining a stable starting point. However this property is acquired in the phase space mapping. The second descriptor is the Angle Sequence (AS), which is a sequence of angles formed by the contour points (Figure 2). The AS, properly encoded into 9 values has been widely used in the form of Angle Code Histogram (ACH) [16]. ACH ignores structural information but achieves rotation invariance, which otherwise is difficult to attain. The angle computed in the contour coordinate system is assigned to each boundary point and concatenation of these angles form the AS, which takes values in the range 0 to 180 (degrees). The AS is tolerant to shape deformations but it is sensitive to noise. The descriptive capability of these two descriptors is complementary providing diverse information and their combination enhances the retrieval procedure.

Having decided for the time series selection the phase space representation follows, where a time sequence x(n) is transformed to a set of m-dimensional vectors, forming points in phase space R^m. The number p of these vectors is equal to the number of boundary points. It should be noticed that these time series of shape boundary points are calculated in a periodical way and problems related to start and end effects are dismissed.

In the phase space representation new properties emerge. Rotation invariance, which is not inherited from the time series is acquired if the time index of trajectory is dropped. In the absence of time index the whole set of multidimensional feature vectors describes the shape and probabilistic measures are employed to quantify the shape's - distance. Although global information (like histogram) is utilized in the matching stage, the structure is still taken indirectly into account by each vector **X(n)**, which conveys information of **m** consecutive boundary points. This could be also considered as partial matching. Partial matching is important when shapes have distortions in some of their parts. The acquisition of global or local information depends on the parameters (time-lag T, embedding dimension m) selection. Therefore different scales of shapes are received.

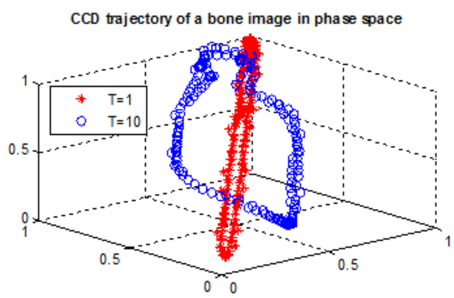

Figure 3 – Phase space reconstruction of CCD for the image “bone”, for time-lag T=1 and T=10.

In conclusion this shape descriptor (1) is rotation invariant, since the rotation does not alternate the set of p, m - dimensional vectors that represent a given sequence. Translation invariance is inherited from the original sequence (CCD, AS). Proper down sampling and normalization is essential for scaling invariance. All CCD, AS values are normalized and their range is [0,1].

An example of a phase space representation in 3-D (m=3) for the CCD time series of image “bone” is given in Figure 3, where trajectories for time-lag T=1 and T=10 are shown. The corresponding representation in phase space for the AS time series is given in Figure 4.

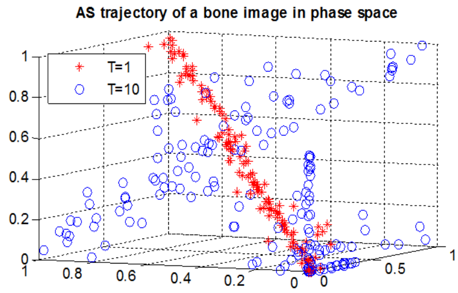

Figure 4 – Phase space reconstruction of AS for the image “bone”, and for time-lag T=1 and T=10.

2.2 Parameter selection

When transforming a time series into phase space, a proper estimation of time-lag parameter T and embedding dimension m should be adopted. There exist several methods for selecting time lag T, such as mutual information criterion or by utilizing autocorrelation information [20]. In this paper autocorrelation functions were used to obtain an approximation of time-lag parameter. The autocorrelation between x(n) and x(n-T) is calculated using the following equation (2), where E{ } stands for the mean value:

$$C\{x(n),x(n-T)\}=\frac{E\{x(n)x(n-T)\}-E\{x(n)\}^2)}{E\{[x(n)-x(n-T)]^2\}} \quad (2)$$

The value of T, where the first zero crossing occurs, is the necessary value. The above equation (2) is repeated for every time-series in the database, including both AS and CCD. Results showed that a mean value for optimum T when dealing with CCD is T=20, while for the AS case T=10.

Selection of the embedding dimension m is accomplished by using the False Nearest Neighbour (FNN) method [17]. This method is based on the fact that m-dimensional phase space must preserve the topological properties of the original phase space (Taken's theorem). Otherwise a higher dimension should be set for correct phase space reconstruction. The algorithm of finding the FNN is described as follows:

For each data point $X^m(i)$ in R^m space find the nearest neighbour, as the point that shares the smaller Euclidean distance:

$$X^m(j)_{NN}=d^{min}(X^m(i),X^m(j))$$

using a predefined embedding dimension m. Augment m by b, $b \in Z^+$ and calculate the distance (3) between the two points :

$$r_i=\sqrt{\frac{d[X^{m+b}(i)-X^{m+b}(j)_{NN}]^2-d[X^m(i)-X^m(j)_{NN}]^2}{d[X^m(i)-X^m(j)_{NN}]^2}} \quad (3)$$

When r_i exceeds a given threshold, then $X^m(i)$ is marked as having a false nearest neighbour. Application of FNN in our database indicated an optimum embedding dimension approximately at m= 5.

The above methods provide a rough estimation for these parameters and were used as initial guidelines. However, experimental results for retrieval tasks reveal different optimum values which are adopted in this work.

2.3 Matching process

A multivariate comparison method is adopted [19] in order to quantify similarity between vector sets in phase space. Given two sets of vectors - trajectories in phase space - $X\{x_i\}$, $i=1{:}p_1$, $Y\{y_j\}$, $j=1{:}p_2$ where p_1, p_2 stand for the number of points in phase space in each set, Mutual Nearest Point Distance is computed using the following equation (4):

$$\mathrm{MNPD}(X,Y)=\frac{\sum_{i=1}^{p_1} d^{\min}(x_i-y_{j=1:p_2})+\sum_{j=1}^{p_2} d^{\min}(y_j-x_{i=1:p_1})}{p_1+p_2} \quad (4)$$

From each point x_i in the m - dimensional data set X the nearest distance to the m - dimensional data set Y is defined. The MNPD is calculated in a bidirectional manner in order to avoid extreme cases where trajectories in phase space belonging to different data sets share small distances when calculating one-directional distances. Summation of these distances leads to X-to-Y dissimilarity. The denominator is used for normalization purposes. The WW statistical test [18] could have been also used instead, but he MNPD is a lower complexity algorithm [19] and for that reason it was preferred.

3. FUSION METHODS IN PHASE SPACE

Combination of the two selected features can be implemented in two different stages of the retrieval process. The objective is to advantageously combine information contained in the two different descriptors either in the early stage producing a new feature, or in a later stage producing another distance measurement. In both cases the present method is very well suited for fusion purposes. An indication of the improved discriminating capability achieved by the combination of the above two descriptors is given in Figures 5, 6. For visualization purposes, the phase-space parameters have been set to m=2 and T=1.

3.1 Early Fusion

Early fusion yields a truly feature representation, since the features are integrated from the start. In that way computational cost is comparable to that of individual descriptors. This approach generally encounters the problem of combining features into a common representation. Our method deals with vectors in phase space which are easily combined in a concatenation scheme producing a new longer feature vector. Given two set of vectors X{x_i}, i=1:p, Y{y_j}, j=1:p a new vector Z is produced,

Z{z_k}, k=1:p, z_k = concatenate (x_i, y_j).

In order to incorporate information that corresponds to the same shape part, feature vectors of the same embedding dimension and time-lag parameter should be used. Let x_i values represent boundary-to-centroid distances points in phase space and y_i represent boundary angles points in phase space, z_k includes both descriptors:

$z_1 = [x_1\ y_1]$, $z_2 = [x_2\ y_2]$, ..., $z_p = [x_p\ y_p]$

The new features Z{z_k} constitute the input for the matching algorithm. Retrieval results of the new feature are given in section 4 and prove the optimized performance which is about 7% above the performance of individual descriptors (CCD or AS).

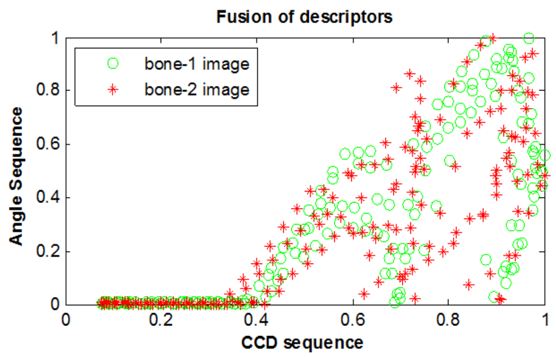

Figure 5 – Fusion of AS and CCD descriptors in 2-D for alike shapes. The strong similarity of phase space trajectories is noticeable.

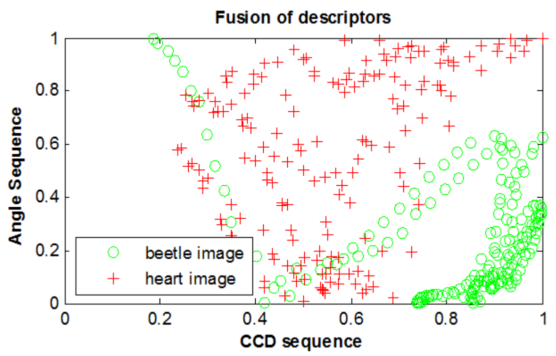

Figure 6 – Fusion of AS and CCD descriptors in 2-D. The dissimilarity of different shapes' trajectories is clear.

3.2 Late fusion.

Late fusion refers to scores' combination. It is based on the individual strength of methods that are combined. A common disadvantage of late fusion which is the potential loss of correlation in mixed feature space, does not apply in our case. Let MNPD(X_i, X_j) and MNPD(Y_i, Y_j), measuring the dissimilarity between set of vectors in phase space originating from CCD and AS sequence respectively. Fusion of these descriptors is implemented by simply averaging the dissimilarity measures MNPD(X_i, X_j) and MNPD(Y_i, Y_j). Retrieval proceeds using the new average distance. It should be noticed that an advantage of score's merging is that combination of phase spaces with different embedding dimension or/and time-lag parameter is feasible. Results of late fusion are given in next section 4. Late or early fusion both optimize the retrieval performance a great deal and to almost the same level.

4. EVALUATION

The proposed method was tested in MPEG-7 Part B shape database, containing 1400 shape silhouettes (70 classes of various shapes, each class containing 20 images), including both scaled and rotated versions of shapes. A small sample of this database is given in Figure 7. These are the four images used in the previous phase trajectory examples.

Performance evaluation was implemented using the Bullseye test. Each image served as a query in the retrieval process. The number of images belonging to the same class was counted at each trial, over the 40 most similar retrieved images. Table 1 shows the improvement of individual descriptors when fusion is used.

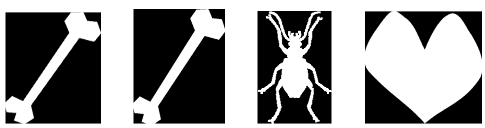

Figure 7 – Sample of MPEG-7 database

Table 1 - Improvement of individual descriptors using fusion

Method	Time-lag	Embedding dimension	Bullseye score
AS	5	15	64.9 %
CCD	12	5	68.8 %
AS+CCD (early fusion)	5	15	75.1 %
AS+CCD (late fusion)	5, 12 (respectively)	15,5 (respectively)	75.1%

Our method reached a Bullseye score 75%, while the state-of-art method of [1] yielded a score of 86.56%. The algorithm proposed by [1] uses the IDSC [6] as a feature, while introduction of the EMD-L1 as the matching method - which is an improvement of the EMD algorithm- leads to such a high retrieval rate. The difference in performance is due to the choice of complex algorithms, accompanied by a relevant computational cost, in the feature extraction stage.

5. CONCLUSIONS

The proposed embedding of 1-D shape descriptor time sequences in phase space creates an efficient representation for shape indexing and discrimination. Invariance to rotation as well as insensitivity to partial distortions of shapes is easily achieved. Proper selection of parameters (time-lag and embedding dimension) that are used for phase space reconstruction results in the efficient integration of diverse information of the shape contour. The ease of implementing fusion in phase space is an additional advantage of the suggested method, as combination of different descriptors improves the retrieval process. Furthermore, the method has a wide applicability and can be easily modified to extract phase space portraits from any 1-D time series descriptor. Application of the proposed method to MPEG 7 shape database proved very promising, while application to other sequences such as signatures, biometrics etc is almost straight forward. Moreover, in the proposed method a predefined set of phase space parameters was adopted in order to represent the whole database's shapes. Implementation of methods that select dynamically the time lag parameter or/and the embedding dimension for each shape individually, is left for further research. Furthermore, improvement in the matching stage could be achieved by considering only representative boundary points, such as curvature zero crossings, in order to enhance the discriminative capability of the method. Finally, a family of global descriptors, e.g. solidity, eccentricity etc can be incorporated in order to enhance the overall retrieval results.

REFERENCES

[1] H. Ling, and K. Okada, "An efficient Earth Mover's Distance Algorithm for Robust Histogram Comparison," *IEEE Trans. PAMI*, vol. 29, pp. 840–853, 2007.

[2] D.Zhang, G. Lu,' Review of shape representation and description *techniques' Pattern Recognition,* vol.37 , pp. 1 – 19, 2004

[3] S. Belongie, J. Malik, and J. Puzicha , "Shape Context: a new descriptor for shape matching and object recognition," *IEEE Trans. Patt. Anal. and Mach. Intell*, vol. 24 (4), pp. 509–522, Apr. 2002.

[4] X Shu, X-J. Wu "A novel contour descriptor for 2D shape matching and its application to image retrieval" ,*Image and Vision Computing,* vol. 29, pp. 286–294, 2011

[5] K. Danillidis, P. Maragos, and N. Paragios (Eds.): "Geodesic Shape Retrieval via Optimal Mass Transport" *ECCV 2010,* Part V, LNCS 6315, pp. 771–784, 2010.

[6] H. Ling , D. W. Jacobs "Shape classification using the inner distance" *IEEE Trans. Patt. Anal. and Mach. Intell.*, vol. 29(2), pp. 286–299, Feb. 2007.

[7] Y-S. Moon, B-S. Kim, M-S Kim, and K-Y Whang, "Scaling invariant boundary image matching using time-series matching techniques, "*Data and Knowledge Engineering*, vol. 69, pp. 1022–1042, 2010

[8] D. Yankov, E. Keogh, L-W Xiaopeng, and W Hodges, "Fast Best-Match Shape Searching in Rotation-Invariant Metric Spaces," *IEEE Transactions on multimedia*," vol. 10, pp.230-239, Feb. 2008

[9] F. Takens, "Detecting strange attractors in turbulence" In *Proc. DynamicalSystems and Turbulence*, Lecture Notes in Mathematics 898, pages 366–381.Springer, 1980.

[10] H-L. Peng, and S-Y Chen, "Trademark shape recognition using closed contours," *Pattern Recognition Letters*, vol. 18, pp.791-803, Aug. 1997

[11] C.Chang, S. Hwang, and D. Buehrer, "A shape recognition scheme based on relative distances of feature points from the centroid," *Pattern Recognition*," vol. 11, pp. 1053-1063, 1991

[12] H.-L. Chan, S.-C. Fang, P.-K. Chao, C.-L. Wang and J-D. Wei "Phase-space Reconstruction of Electrocardiogram for heartbeat Classification," *World Congress on Medical Physics and Biomedical Engineering*, Munich, Germany, September 7 - 12, . 2009, IFMBE Proceedings, vol. 25/4, pp.1234-1237,2010

[13] F. Fotopoulou, N. Laskaris, G. Economou and S. Fotopoulos, "Improvement in Leaf Image Retrieval using the Multidimensional Embedding Sequence Similarity (MESS) Method ," *Pattern Analysis with Applications*, under review

[14] R.J. Povinelli, M. T. Johnson, A. C. Lindgren, F. M. Roberts, J. Ye, "Statistical models of reconstructed phase spaces for signal classification," *IEEE Transactions on Signal Processing*, 54(6-1): 2178-2186 (2006)

[15] J. Richiardi, K. Kryszczuk and A. Drygajlo "Static Models of Derivative-Coordinates Phase Spaces for Multivariate Time Series Classification: An Application to Signature Verification," *Advances in Biometrics Lecture Notes in Computer Science*, 2009, Volume 5558/2009, 1200-1208

[16] Wang, Z., Chi, W., Feng, D.: Shape based leaf image retrieval. IEE Proc. Vision, *Image and Signal Processing* vol.150(1), 34–43, 2003

[17] H.D.I Abarbanel, "The Analysis of Observed Chaotic Data in physical systems," *Rev. Mod.Phys.* , vol. 65, pp. 1331-1392, 1993

[18] J.H.Frieman, L.C Rafsky, "Multivariate generalizations of the Wald-Wolfowitz and Smirnov two-sample tests, "*Ann.Stat* 7, vol. 11, pp. 697-717, 1979

[19] Shih-Chin Fang, Hsiao-Lung Chan, "Human Identification by quantifying similarity and dissimilarity in electrocardiogram phase space," *Pattern Recognition*, vol. 42, pp. 1824–1831, 2009.

[20] H.Kantz and T.Schreiber "Nonlinear time series analysis," *Cambridge:Cambridge university press* , pp. 388, 1997

ACKNOWLEDGMENTS

This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: Heracleitus II. Investing in knowledge society through the European Social Fund.