# A Method of Shape Recognition Using the Smoothed Group Delay Function

Sreyasee Das Bhattacharjee
Research Scholar
Dept of Computer Science & Engineering
Indian Institute Of Technology Madras

Dr. Sukhendu Das
Associate professor
Dept of Computer Science & Engineering
Indian Institute of Technology Madras

Dr.Amitabha Datta
Associate professor
Dept of Computer Science & Engineering
University of Western Australia, Perth

## Abstract

*Fourier descriptors have been used in the past as a feature for describing the shape through its boundary contour. The contour based Fourier transform and the corresponding phase information have been used implicitly, to improve the results for image matching and retrieval. The phase of a signal stores the higher order moments and hence can be used to discriminate among various shapes. This paper exploits phase information for signal explicitly by taking the derivative of the phase, termed as the smoothed Group Delay, and then uses it for shape matching. We try to extract the features of a shape from the Group Delay function (of the boundary) and try to classify them. Results are shown on a standard database of object shapes.*

*Keywords: Phase, SGD, Correlation, and Similarity*

## 1. Introduction

Much work has been done on shape analysis, recognition [1,2,3,4,5,6,8,13] and verifications using $2^{nd}$ order statistical parameters derived from shapes, such as moments, Euler number, eccentricity, distance transforms, Gabor/Wavelet and Hough transforms etc. On the other hand, it has been observed from simple experiments that phase (spectral domain) of a signal has important information required for signal reconstruction. Not much work has been reported on the use of phase for discriminating shapes (say, even in 2-D). Potential applications of this work will find scope in Retrieval of object/shapes from a large database, object recognition, visual attention in artificial vision etc.

Neuro-biologists have reported that the brain typically may not use phase of a signal to represent or recognize shapes. But phase is considered to contain the higher order statistics of a signal/shape, which is vital for discrimination, recognition and retrieval. Phase of a signal has been used as a feature for speech analysis [10, 11, 12], signal coding etc. We plan to explore the use of phase of a 1-D signal to discriminate between shapes. A 2-D shape (we shall initially assume, 2-D shapes with no holes) will be converted to a 1-D function by obtaining the coordinates of points on the perimeter of the shape with respect to the centroid of the shape (or even origin of image plane). For affine-transformation invariant representation of a shape, as well as to eliminate the problem of wrap-around in phase, we shall use the derivative of phase (unwrapped), called the group delay function for shape analysis. There has been no well-defined work reported in this context and it should be quite novel for us to investigate this concept for shape matching and retrieval.

The Fourier Descriptor uses a boundary representation of an object as a discrete-time complex periodic signal and then takes the frequency component of the representation as done in [1, 2, 3, 5]. The phase was made invariant to any affine transformation in the approach of WARP [1, 3]. But in our experiments, we found group delay of a signal to be much more tolerant and automatically invariant to any affine transformation, rather than other features, such as moments, Fourier spectrum (magnitude) and even wavelet transforms. This is the main motivation for the use of group delay as a feature to represent shapes. Extracting phase becomes troublesome due to its wraparound problem, which is because of the value of the phase lying within the range [-π, π]. So it is important to unwrap the phase while extracting the features, by adding multiples of (±2π) when absolute jumps between consecutive elements of phase are greater than the default jump tolerance of [π] radians.

Our objective is to extract the shape information from the spectral component of the structure using the Fourier Transform and hence using the Smoothed Group Delay (SGD) [10] of the signal as an important cue towards finding the shape features. Traditionally the Group Delay (GD) has been used for feature extraction in speech signals as in [12]. We present here a novel idea for shape retrieval using the smoothed group delay as a feature.

## 2. Theoretical background

### 2.1 Computation of Signature Function

The signature function defined on the boundary of a shape can be traced and sampled in order to obtain uniformly distributed N points as illustrated in the images in the left column of Figure 1.Coordinates of the boundary points of the shape are expressed as: $(x_0, y_0)$ $(x_1, y_1)\ldots(x_N, y_N)$. The signature function can be expressed as a sequence of complex numbers of these coordinates (with respect to the centroid of the shape - this incorporates the translation invariance) as:

$$z_i = x_i + jy_i \quad i=1\ldots.N \tag{1}$$

The coefficient of the Fourier Transform of $z_i$ is given as

$$Z(e^{j\omega}) = \sum_n z_k e^{(-j\omega n)} \tag{2}$$

When a shape is rotated by an angle θ, and the starting point by $l_0$, then the corresponding DFT coefficients are as given in [1].

$$Z_m^{'} = Z_m e^{j\theta} e^{(-jm2\pi l_0)} \tag{3}$$

where

$$Z_m = R_m e^{j\theta_m}, \qquad m=1\ldots N$$

(4)

Thus the shift in phase, due to rotation of the shape and change in starting point is

$$\theta_m^{'} = \theta_m + \theta - 2\pi l_0 m / N \tag{5}$$

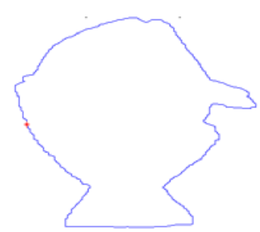
The scale factor is normalized by [1]as: $R_m^{'} = R_m / S$, where S is the scaling factor and S=$R_1$ as given in [1].

This leads to the derivation of the modified DFT coefficients [1] for normalization against the scaling, rotation and shift in the starting point, which is given in the table 1 below:

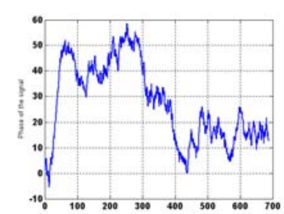
Table 1: Modified DFT Coefficients

| Invariance | Modified coefficients |
|---|---|
| Translation | **$Z_0$ ‘ =0** |
| Scale | **$R_m$ ‘=$R_m$ /S** |
| Rotation | **$\Theta_m$’= $\Theta_m$+ θ -($\Theta_{-1}$ + $\Theta_{+1}$ )/2** |
| Starting point | **$\Theta_m$’= $\Theta_m$ +m($\Theta_{-1}$ - $\Theta_{+1}$ )/2** |

The work by Bartolini et. al [1] comes up with the formulation as in the table 1 above to take care of the normalization. In case of group delay, where the derivative of the phase is taken, the added terms in the phase are constant terms and hence they should not alter the group delay coefficients. In addition, translation and scale do not alter the phase of the signature. This is the basic justification for the use of the derivative of phase for shape discrimination invariant to affine-transformations. In the following, we discuss the concept of group delay function and its implementation. Consider the example of four rotated version of the same shape in Figure 1, where this results in four different shifts in the phase domain vertical axis is in radians and hence rotation angle in degrees appear negligible. The overall nature of phase variation is unaltered. This can also be verified from the row-3 of table 1.

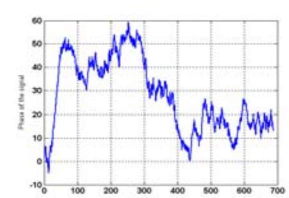


Shape(a1)



(b1)

Shape(a2)



(b2)

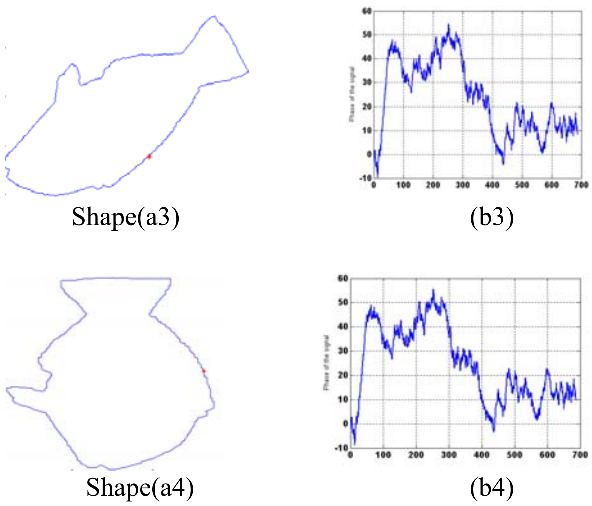



Figure 1. Four different orientations of the same fish shape (structure) shown in a1-a4; b1-b4 show the corresponding phase of the signature extracted from the boundary contour.

### 2.2 Group delay functions

If the phase spectrum θ(ω) of a signal is defined as a continuous function of ω, the group delay function [11,12] is defined as

$$\tau(\omega) = -d(\theta(\omega))/d(\omega) \quad (6)$$

In our experiment, we take the phase and smooth it using a one-dimensional operator [16]. The resultant signals are the smoothed version of the derivative of the original phase obtained from the shape `signal, which we refer as the smoothed group delay (SGD). This has been used in our proposed approach as a feature for discrimination between the various shapes.

Analytically, this (SGD) can also be expressed as

$$\tau_s(\omega) = G_\sigma(\omega) * d(\theta(\omega)/d(\omega)$$

or equivalently,

$$\tau_s(\omega) = d(G_\sigma(\omega) * (\theta(\omega))/d(\omega)$$

where, $G_\sigma(\omega)$ is a Gaussian function with suitable standard deviation $\sigma$.

## 3. Proposed Methodology

### 3.1 The shape descriptor (SGD)

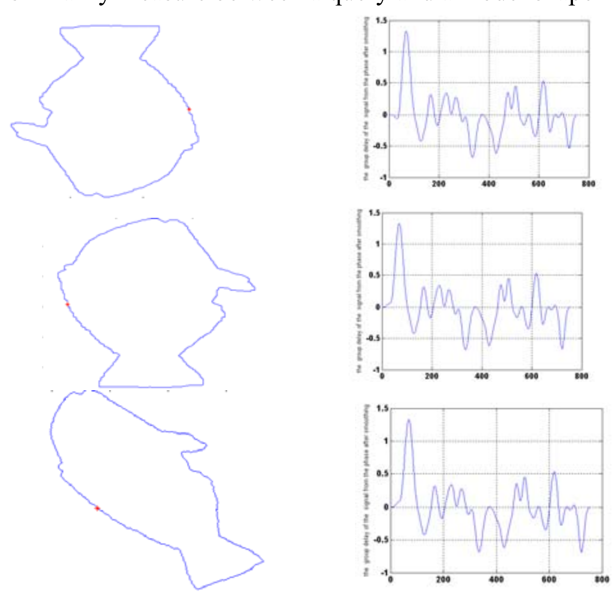
Phase being sensitive to the translation, scaling and rotation; alter the phase spectrum of an object shape. However, the affect becomes evident in terms of only a scaling factor, the overall trend remains same, i.e. for different rotations of a single shape the phase spectrum only gets a translation in the frequency space, as we can see in the column (b) of Figure 1. This illustrates the constant shift in phase for rotation of the object. The shape is made translation invariant by altering boundary contour points with respect to its centroid. The scale doesn't alter the phase as shown in row 2 of Table 1.

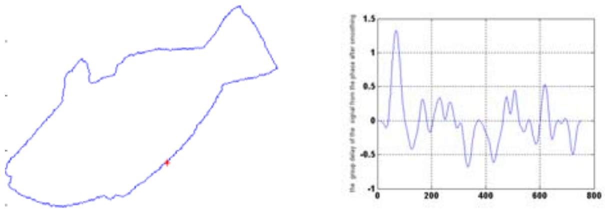
In our approach we use the group delay as the Fourier descriptor for shapes. As shown in Figure 2**,** group delay function is rotation, translation and scale invariant and remains unaltered for the same shape for any rotation. Different shapes will have different group delay functions (SGD), as illustrated in Figure 3 and therefore can act as a unique feature for shapes.

We extract the phase from the DFT coefficient, smooth it using a one-dimensional Canny operator [16] with standard deviation σ = M/70, where M is the length of the signal. This specific choice of σ makes the extent of the Gaussian proportional the length of the signal. The choice of this σ is crucial, as the performance that was observed with the value at σ = M/10 was much unsatisfactory. Gradually increasing the value of σ does improve the accuracy and an optimum choice was found at M/70. The resultant group delay, obtained after convolving the phase with the operator, is referred as a smoothed group delay (SGD). This is stored in a database of object templates.

### 3.2. Shape similarity measure using SGD

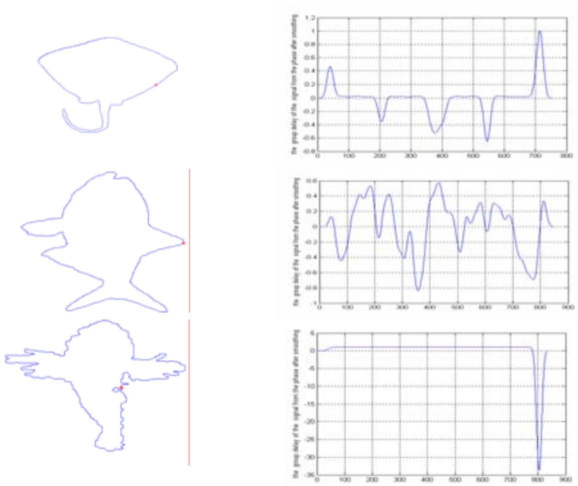
Then for any query image, we extract the smoothed group delay and take a correlation measure with all models stored in the database of object templates. The similarity measure between a query and a model shape in





(a) (b)

Figure 2. Examples of a fish shape at different rotation angles and different starting points. (a) Fish shape and (b) corresponding smoothed group delay (SGD) functions.



(a) (b)

Figure 3. Examples of four different (a) fish shapes, (b) corresponding smoothed group delay function.

the database is computed as the ratio of the maximum and mean (average) of the absolute correlation values obtained from the SGDs of the two shapes. This is because we expect a peak in the correlation output of two similar SGDs extracted from two similar shapes as shown in Figure 4.

If, $C = g_0 \otimes g_1$ (7)

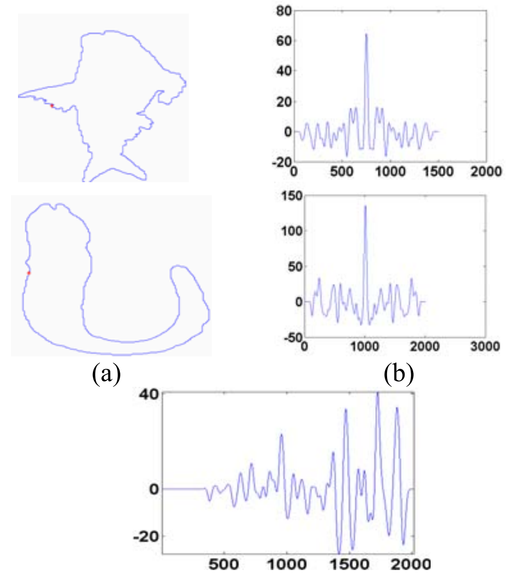
where, $g_0$ and $g_1$ are the smoothed group delays of the query shape and a model in the database respectively, and $\otimes$ indicates the correlation function.

If, b = mean(abs(C)) and m = mean(abs(C)

Then the following measure is used to compute the similarity between the two shapes

$S_m = b/m$ (8)

The result of query is reported for the first few rank-ordered samples in descending order of the value '$S_m$'. Figure 4 shows how the proposed similarity measure works well. A sharp peak is observed when the SGD of a shape is correlated with itself (auto-correlation) or a very similar shape, whereas there is no clear sharp peak when two dissimilar shapes are being correlated. Hence only nearby (similar) shapes will be retrieved. Results of our proposed method are given in the following section.




(c)

Figure 4. (a) Two dissimilar shapes; (b) Auto-correlation of the respective group delay functions of the shapes; (c) Correlation of the SGDs of the shapes in (a).

## 4. Experimental Results

We performed our experiment on the FISHES dataset [15]. Some sample images of the dataset used are shown in Figure 5.

Table 2
Some sample images from the Semantic categories

| | | | |
|---|---|---|---|
| Shark |  | U-Eels |  |
| Soles |  | Rays |  |

Figure 5. Some samples of Fish images from the FISHES dataset [15].

We first manually classified each image based on 5 semantic categories: shark (28 images), soles (34 images), Rays (20 images), U-Eels (19 images), Pipefishes (25 images). Others, which are not in any category, were chosen to be in a default category. Table 2 shows some images along with their respective categories. In the testing stage, we used 30 query images from each of the semantic categories. Some of the observed results are shown in the Figure 6.

(a) (b)

Figure 6. (a) Query shapes, (b) Retrieved shapes in the decreasing order of similarity measure, i.e. ‘$S_m$’ using equation (8).

## 5. Comparison with the prior work

The performance of our method is compared with that of the results obtained in [1] in figure 7. The method used in [1] computes the DFT coefficients, normalizes it to give it a translation, rotation and starting point invariance. The resultant normalized phase as shape descriptor, is used to reconstruct the corresponding time domain representation of the shape signal by IDFT. The reconstructed and the original signals are then compared using DTW distance. This is a computationally cumbersome process, which does not exploit the phase information explicitly.

In [1] while comparing with the samples already stored in database, a signal of the query shape undergoes a pair of DFT and IDFT transformations along with a procedure of normalization. Whereas, our proposed method consists of a DFT and a convolution with a one-dimensional Canny operator [16] to obtain the SGD representation for shapes. Similarity measure involves a correlation and peak detection procedure, which is simpler than DTW. Hence the method proposed in [1] is more expensive than the method we used. Comparative study of our results with that obtained from the method in [1] is given in Figure 7**.**

(a) (b)

(c) (d)

Figure 7. (a) The query shape (b) The shapes retrieved using the method recommended in [1], (c) the same query shape (d) the shapes retrieved using our method.

## 6. Conclusion

In this paper, we have introduced a new way to analyze the shape using a new Fourier based descriptor, which is the derivative of the smoothed phase. It is extracted to represent the complex boundary of the shape, and is termed the smoothed group delay (SGD). Correlation of the SGDs of two similar shapes is used as a similarity measure. We have tried our method on a small set of shapes where this method has performed satisfactorily well. This may be considered as a first step towards the use of phase information explicitly, for the purpose of shape matching and retrieval invariant to transformation, rotation and scaling. Future scope of work involves testing the SGD on a large and complex dataset, exploiting the occurrence of zero crossings (ZC) for improving the robustness of SGD in presence of noise and use the proposed method of shape information extraction on more complicated shapes (objects with holes).

## 7. References

[1] Ilaria Bartolini, Paolo Ciaccia, M.I., Patella, M.: “Warp: Accurate retrieval of shapes using phase of Fourier descriptors and time warping distance”. *IEEE*

*Tran. On Pattern Analysis and Machine Intelligence*. Volume 27, (2005), pp.142-147.

[2] Chen, Q., Petriu E., Yang, X.: "A comaparative study of Fourier descriptors and Hu's seven moment invariants for image recognition". *IEEE Canadian Conference on Electrical and Computer Engineering*. Volume 1, (2004), pp.103–106.

[3] Reeves A.P., Prokop R.J., Andrews S.E., Kuhl F.P.: "Three-dimensional shape analysis using moments and Fourier descriptors*". IEEE Tran. on Pattern Analysis and Machine Intelligence*. Volume 10, (1988), pp. 937–943.

[4] H.K. Seppanen, T.Pietikainen, M.: "An experimental comparison of autoregressive and Fourier-based descriptors in 2d shape classification". *IEEE Tran. on Pattern Analysis and Machine Intelligence*. Volume 17, (1995), pp.201–207.

[5] Arbter K., Snyder W.E., Burkhardt H., Gred Hirzinger: "Application of affine-invariant Fourier descriptors to recognition of 3-d objects". *IEEE Tran. on Pattern Analysis and Machine Intelligence*. Volume 12, (1990), pp. 640–647.

[6] Cai J., Liu Z.Q.: "Hidden Markov models with spectral features for 2d shape recognition". *IEEE Tran. on Pattern Analysis and Machine Intelligence*. Volume 23, (2001), pp.1454–1458.

[7] Sheng Chen, Q., Defrise, M., , Deconinck, F.: "Symmetric phase-only matched filtering of Fourier-mellin transforms for image registration and recognition". *IEEE Tran. on Pattern Analysis and Machine Intelligence*.Volume16, (1994), pp.1156–1168.

[8] D.W.Paglieroni, E.ford, G., M.Tsujimoto, E.: "The position-orientation masking approach to parametric search for template matching". *IEEE Tran. on Pattern Analysis and Machine Intelligence, Volume 16,* (1994), pp.740–747.

[9] Duda R.O., Hart P.E., D. G. Stork: *Pattern Classification*. John Wiley and Sons, Singapore (2001).

[10] Murthy, H.A., Gadde, V.R.R.: "The modified group delay function and its application to phenome recognition". *IEEE International Conference on Accoustics,Speech, and Signal Processing*. Volume 1, (2003), pp. 68–71.

[11] Murthy H.A., Madhu Murthy, K.V. Yegnanarayana, B.: "Formant extraction from phase using weighted group delay functions". *Electron. Lett*. Volume 25, (1989), pp.1609–1611.

[12] Yegnanarayana, B. Murthy, H.A.: "Significance of group delay functions in spectrum estimation". *IEEE Transactions on signal processing,* Volume 40, (1992), pp.2281–2289.

[13] Bartolini, I., Ciaccia, P., Patella, M.: Using the time warping distance for Fourier based shape retrieval. (2002), www-db.deis.unibo.it/research/**pap**ers/TRs/IEIIT-03-02.**pdf**

[14] Oppenhiem, A.V., Schafer, R.W., Buck, J.R.: *Discrete-Time Signal Processing*. Pearson Education, New Delhi (1999).

[15] Abbas, S., Mokhtarain, F., Kittler, J.: Squid demo dataset1,500. http://www.ee.surrey.ac.uk/Research/VSSP/imagedb/demo.html (1997).

[16] P. Kiran Kumar, S. Das, B. Yegnanarayana.:"*One Dimensional processing of images"*. International Conference on Multimedia Processing and Systems. (2000), pp.181-183.