SHAPE DESCRIPTION USING PHASE-PRESERVING FOURIER DESCRIPTOR

Emir Sokic, Samim Konjicija

Faculty of Electrical Engineering, University of Sarajevo
Bosnia and Herzegovina
esokic@etf.unsa.ba

ABSTRACT

Contour-based Fourier descriptors are established as a simple and effective shape description method for content-based image retrieval. In order to achieve invariance under rotation and starting point change, most Fourier descriptor implementations disregard the phase of the Fourier coefficients. We introduce a novel method for extracting Fourier descriptors, which preserve the phase of Fourier coefficients and have the desired invariance. We propose specific points, called *pseudomirror* points, to be used as shape orientation reference. Experimental results indicate that the proposed method significantly outperforms other Fourier descriptor based techniques.

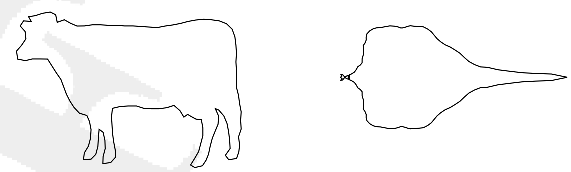
***Index Terms*—** Content based image retrieval, Fourier descriptors, phase, nominal orientation, pseudomirror points

1. INTRODUCTION

The most important property of a content-based image retrieval (CBIR) system is the ability to effectively and efficiently describe shape. Today, many shape description methods exist, but very few of them are not limited by specific application, performance or computational complexity. Thus, the design of an efficient, effective and versatile shape descriptor is still an open challenge.

Fourier descriptors (FD) are global contour-based shape descriptors that have good retrieval accuracy, compactness, are insensitive to noise, and have a hierarchical representation in the spectral domain. They are obtained by applying the discrete Fourier transform (DFT) over a shape signature [1–11]. By disregarding the phase and using only the magnitude of Fourier coefficients, these descriptors become invariant under rotation, translation, scale, and change of the starting point of the contour. Many authors adopt this simplistic approach [1–3], which renders valuable information contained in phase inevitable lost.

Since magnitude-based FD are not information preserving, the original shape cannot be reconstructed from FD. Moreover, they cannot be used in shape-retrieval tasks where rotation invariance is not desirable (e.g. traffic signs recognition [4]).



(a) (b)

Fig. 1. Reconstruction of shapes with equal FD magnitudes: a) shape reconstructed with the original phase content, b) shape reconstructed with different phase content.

When the phase information is discarded, the descriptor ability to discriminate between shapes is affected. As illustrated in Figure 1, two completely different shapes may have equal magnitudes of Fourier coefficients.

In order to use the phase of the Fourier coefficients for shape-based image retrieval, phase should be normalized under rotation and starting point change. These are the most common approaches of FD phase normalization found in literature:

- normalization using geometrical properties such as points with maximal radius [5], or central moments [6],
- normalization using the phase of the first harmonic [5, 7, 8],
- normalization using the phase of higher order harmonics [5, 9, 10],
- phase is not normalized, but shapes are implicitly aligned in the spatial domain using cross-correlation [4], or Procrustes distance as proposed in [11, 12].

Normalization of the phase using single contour point as shape reference is not suitable when noise, indentation or spurious peaks are present in the shape.

Normalization using the phase of the harmonics is an ill-conditioned task. Small variation of the phase of the first harmonic causes large variations of the phase of higher order harmonics. In the cases of noise, circular harmonic locus or

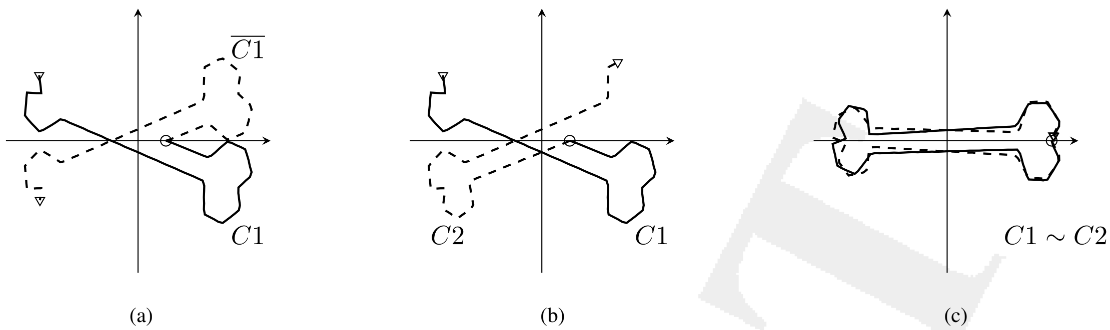


Fig. 2. Parts of the contours are omitted for clarity. Symbol o marks the starting points, while the triangle symbol $\triangle$ marks the end points. a) The original contour $C1$ and the "mirrored" contour $\overline{C1}$ of the shape are depicted for an arbitrary starting point. b) Original contour $C1$ and "mirrored" contour $C2$ of the shape are depicted for an arbitrary starting point. Contours $\overline{C1}$ and $C2$ differ only by directionality. c) One nominal orientation (pseudomirror point) for which $C1$ and $C2$ are very similar.

harmonics with negligible magnitude, these methods usually fail. Moreover, phase is computed as 2π-modulus function, thus the uniqueness of the solution is usually questionable.

It is interesting to note that some authors admit that the normalized phase is not recommended to use for shape description (e.g., in [5]). In [6], authors propose a weighting scheme on the variance of magnitude and phase ratios between descriptors, to minimize phase normalization errors. Few authors try to compensate for non-perfect normalization using more complex distance measures (such as cross-correlation [4] or time-warping distance [8]). Methods that avoid the normalization of the phase, but instead align the shapes during matching [6, 10, 11], are more robust and also more computationally demanding.

To this end, we propose a method to determine nominal orientation(s) of the shape during descriptor extraction. Invariance under rotation and starting point change is achieved by using so called *pseudomirror* points for nominal orientation. They are determined in spatial domain, yet Fourier descriptors are used to determine their semantic meaning. Experimental results indicate that phase-preserving FD extracted using pseudomirror points as orientation reference, outperform other phase-based and magnitude-based Fourier descriptor techniques.

The paper is organized as follows. Introduction of pseudomirror points and Phase-including Fourier Descriptor (PIFD) is given in Section 2. Experimental results are discussed in Section 3. Finally, a conclusion is given at the end of the paper.

2. PHASE-INCLUDING FOURIER DESCRIPTOR (PIFD)

In preprocessing stage, the coordinates of the shape boundary need to be extracted from the image. Then, the contour is re-sampled by the fixed number of points N, using equal arc-length sampling. For subsequent analysis, it will be assumed that a shape contour is given by N boundary points $P_n = (x_n, y_n)$, where $n = 0, 1, ..., N-1$. Points of the contour $P_n = (x_n, y_n)$ can be represented in the form of complex numbers:

$$Z_n = x_n + jy_n, \tag{1}$$

for which the Discrete Fourier Transform may be computed as:

$$a_k = \frac{1}{N}\sum_{n=0}^{N-1} Z_n e^{-j2\pi nk/N}, \tag{2}$$

where $k = 0, 1, ..., N-1$. Fourier coefficients a_k are used to derive Fourier descriptors. Our proposed translation and scaling invariant descriptor is given by:

$$\mathbf{F_{TS}} = \left\{\frac{a_{-M/2}}{Sc}, ..., \frac{a_{-1}}{Sc}, \frac{a_1}{Sc}, \frac{a_2}{Sc}, ..., \frac{a_{M/2}}{Sc}\right\}, \tag{3}$$

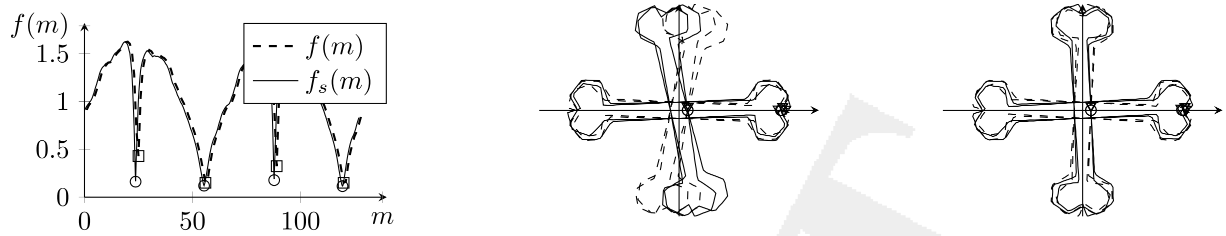
where $Sc = \sum_{k=1}^{N-1} |a_k|$ is the scaling factor, and a_k are computed using (2). The number of Fourier coefficients needed for shape representation is denoted by M. The number M is usually small (≤ 30) and invariant with the respect to the number of points N. Note that the DFT is a periodic sequence with the period N ($a_{-m} = a_{N-m}$).

The descriptor given with (3) is invariant under translation and scaling, which is very easy to demonstrate (see e.g., a procedure presented in [3]).

2.1. Pseudomirror points

In order to use information contained in the phase, one must obtain the invariance of a_k under starting point and orientation change. Let the $a_k^{(old)}$ be the Fourier coefficients of the initial shape, with starting point P_0. The Fourier coefficients of the shape with the starting point P_m, rotated for an angle ϕ, are given by:

$$a_k^{(new)} = e^{j\phi} e^{j2\pi km/N} a_k^{(old)}. \tag{4}$$



(a) (b) (c)

Fig. 3. a) Cost function $f(m)$ and its sub-sampled function $f_s(m)$, with their corresponding minima, b) shape orientations corresponding to the local minima of $f(m)$, c) shape orientations corresponding to the local minima of $f_s(m)$.

If the starting point P_m (or m) is known in advance, then the shape could be de-rotated so that the starting point always has a specific predetermined argument, therefore ϕ could be eliminated from equation (4).

We propose points with specific geometrical and shape discriminative meaning to be chosen as starting points P_m. They are determined using the following procedure. Each contour point P_m is chosen as a starting point, and the contour is de-rotated with respect to the centroid of the shape, so that the starting point of the contour lies on the positive real axis. This contour will be noted $C1$. Then another, "mirrored" contour is created. The "mirrored" contour $C2$ is derived by conjugating all points of the contour $C1$, and changing the direction of the contour. The whole process is illustrated in Figure 2. For the original contour $C1$, and for the "mirrored" contour $C2$, FD $a_k^{(1)}$ and $a_k^{(2)}$ are computed, respectively. We propose a following cost function for determining the starting points:

$$\begin{aligned} f(m) &= \sum_{k=1}^{N-1} |a_{k(m)}^{(1)} - a_{k(m)}^{(2)}|, \\ &= 2\sum_{k=1}^{N-1} |Im\{a_{k(m)}^{(1)}\}|, \end{aligned} \tag{5}$$

where $Im\{\}$ denotes imaginary part. Points P_m which correspond to the local minima of the cost function $f(m)$ are proposed to be used as starting points. Such points P_m will be called *pseudomirror* points. Depending on the shape, several pseudomirror points could be found. The rationale behind this approach is that the pseudomirror points capture the local (pseudo) similarity of the shape to its mirrored version, in terms of Fourier descriptors. The pseudomirror points are inherent to the shape, and hold some shape semantics. In case of symmetrical objects, pseudomirror points determine a line which coincides with the axis of symmetry. In case of shapes with occlusions or indentations (non-perfect symmetry), pseudomirror points determine the axis of "pseudosymmetry", hence these artifacts have less influence in determining nominal orientation. Moreover, pseudomirror points can be computed in the pre-processing stage, instead of the matching stage, which allows for a simpler matching technique.

There are two implementation issues that need to be addressed. When using a lower number of contour points, the gradient of $f(m)$ computed in points close to the centroid tends to be large, so the exact minimum of $f(m)$ could fall somewhere between m and $m+1$, leading to smaller errors. A simple workaround is to interpolate and sub-sample the original function $f(m)$ with $f_s(m)$ around the local minima. This is illustrated in Figure 3.

Experiments have shown that another small modification of the algorithm improves retrieval accuracy. The harmonics with the larger magnitude contribute to the coarse description of the shape, while those with the smaller magnitude usually contain subtle intra-class variations, details and noise. Let R_{ind} be the set that contains the indices of R largest harmonics, in terms of magnitude. The value R and the set R_{ind} are determined for the specific shape by solving the inequality $\sum_{i\in R_{ind}} |a_i| \leq 0.95\sum_{j=-M/2,i\neq 0}^{M/2} |a_j|$. Using $f(m) = 2\sum_{k\in R_{ind}} |Im\{a_{k(m)}\}|$ instead of (5) will improve retrieval performance and reduce computational time.

2.2. PIFD extraction procedure

These are the steps to follow in order to extract PIFD from shape contour:

1. In order to avoid constant normalization under translation and scale, the shape is translated so that the centroid of the shape (given with $Z_C = \sum_{n=0}^{N-1} Z_n$) is located at the origin. The shape is then scaled in the spatial domain by a factor $Sc = \sum_{k=1}^{N-1} |a_k|$.

2. For each $m \in \{0, 1, 2, ..., N-1\}$, a point P_m is chosen as the starting point, and the contour is de-rotated so that the starting point of the contour has the argument of zero degrees. For each starting point $a_{k(m)}$ may be computed using $a_{k(m)} = e^{-jarg(Z_m)}e^{j2\pi mk/N}$ $e^{jarg(Z_0)}a_{k(0)}$.

3. The cost function $f(m) = 2\sum_{k\in R_{ind}} |Im\{a_{k(m)}\}|$ is computed.

4. Let Q be a number of wanted pseudomirror points, and LM be the number of the cost function's local minima.

Table 1. Similarity measures

Mirror non-invariant	$d(P_1, P_2) = min_{(i_1,i_2)} \sum \lvert P_1(i_1, :) - P_2(i_2, :)\rvert$
Mirror invariant	$d(P_1, P_2) = min_{(i_1,i_2)} \left[min\{\sum \lvert P_1(i_1, :) - P_2(i_2, :)\rvert, \sum \lvert P_1(i_1, :) - \overline{P_2(i_2, :)}\rvert\} \right]$

(*Step 4a*) If $Q \leq LM$, then Q best local minima of $f(m)$ and corresponding points P_q are found.
(*Step 4b*) If $Q > LM$ then $Q - LM$ points with the lowest value of $f(m)$ are added to set $\{P_q\}$ of local minima of $f(m)$. This is a useful feature which improves retrieval accuracy, especially for a coarser sampling of the contour. For example, if the minimum is located between indices m and $m + 1$, it is wise to use both of these points as potential pseudomirror points in order to reduce the discretization errors.
The resulting points in $\{P_q\}$ are potential starting (pseudomirror) points.

5. For each point $P_{ms} \in \{P_q\}$ corresponding to the index m_s, the function $arg(Z_{m_S})$ is interpolated with $g(m_s)$ on the interval $[m_s - S, m_s + S]$, where S is a positive integer (usually $S = 1$ or $S = 2$), and $g(m_s)$ is a linear or cubic interpolation function.

6. For each pseudomirror candidate point index m_s, a finer pseudomirror candidate point index $m_s^{fine} \in [m_s - S, m_s + S]$ is found by minimizing $f(m)$ using $a_{k(m_S)} = e^{-jg(m_s)} e^{j2\pi m_s k/N} e^{jarg(Z_0)} a_{k(0)}$. Instead of a set of points $\{P_q\}$, the corresponding $\{P_q^{fine}\}$ is determined. For each point P_q^{fine} descriptors $a_{k(q^{fine})}$ are derived.

7. The descriptor is formed as:

$$PIFD_{Q\times M} = \begin{bmatrix} \{a_{k(q_0^{fine})}\} \\ \{a_{k(q_1^{fine})}\} \\ \vdots \\ \{a_{k(q_{(Q-1)}^{fine})}\} \end{bmatrix}_{q_i \in q, k \in M_{ind}} . \quad (6)$$

PIFD is a parameter based descriptor, which means that tuning parameters for different datasets may provide better results. Nevertheless, parameters $N = 512$, $M = 20$, and $Q = 8$ imply promising results for almost all applications, while keeping the descriptor compact and allowing fast shape matching.

2.3. Similarity measures

Clearly the relative position of the proposed pseudomirror points is mirror-invariant, which means that the points will be positioned in the same place on the shape even if the shape contour is mirrored with respect to arbitrary axis. One of the advantages of PIFD is that it is very simple to introduce mirror invariance in shape description. It is easy to show (using (5)) that mirrored shapes have conjugated PIFDs.

Suppose that two different shapes are described by PIFDs P_1 and P_2 respectively. We propose two similarity measures, given in Table 1.

3. EXPERIMENTAL RESULTS

3.1. Methodology

The proposed shape descriptor is tested on the popular MPEG7 CE-1 Set B [13]. Representative elements of MPEG-7 dataset are depicted in Figure 4. MPEG-7 CE-1 Set B consists of 1400 shapes representing real life objects, classified into 70 classes with 20 similar shapes for each class. This database is convenient for shape-based image retrieval testing since it includes rotation, scaling, skew, stretching, defection, indentation and articulation of shapes.

Two commonly adopted measures of retrieval performance are computed: Precision and recall (PR) diagrams (used in [1–3]), and Bulls-Eye score (used in [8, 12, 14, 15]).

Precision is defined as the ratio of the number of the relevant shapes to the total number of retrieved shapes, while recall is defined as the ratio of the number of retrieved relevant shapes to the total number of relevant shapes in the entire database. Average precision for all recall values for all query shapes in the database is computed and presented.

Bulls-Eye score (BE) is defined as the percentage of relevant results in the first 40 retrieved results of a query. Average Bulls-Eye score is computed after all elements in the dataset are used as a query. As opposed to PR diagrams which demonstrate precision across all recall values, Bulls-Eye scores favor algorithms that provide a higher precision for top retrieval results.

3.2. Results and discussion

We selected and implemented several most representative and best performing Fourier-based methods for comparison with PIFD: magnitude-based methods (Normalized complex coordinate FD (NCC) [3] and Combined Perimeter Area Function (CPAF) [2]), phase-based methods (Affine-Fourier descriptor (AFD) [9], First harmonic aligned FD (FHAFD) [5, 7], Variance based modified FD (VBMFD) [6]), phase-based methods with alignment in matching stage (The Correlation based

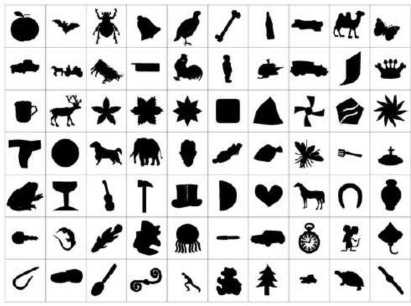


Fig. 4. MPEG7 CE-1 Set B dataset representative shapes (70 classes with 20 variations per class)

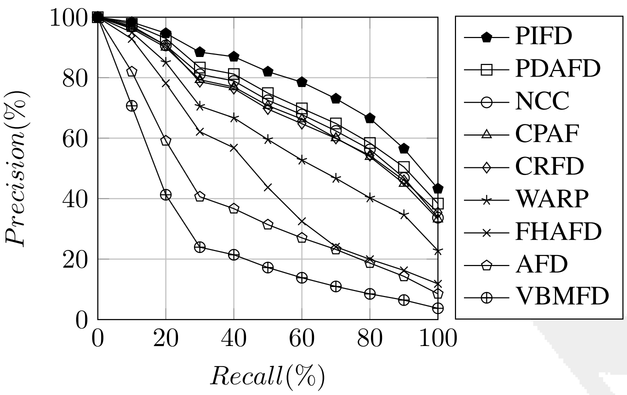


Fig. 5. PR diagrams obtained on MPEG-7 dataset

FD (CRFD) [4], WARP FD (WARP) [8] and Procrustes distance aligned FD (PDAFD) as proposed by [11]). To the best of the authors' knowledge, it is the first time some of these descriptors (CRFD, WARP, PDAFD, VBMFD) were tested on the whole MPEG-7 dataset.

The obtained PR diagrams are given in Figure 5, and Bulls-Eye scores are presented in Table 2. As it may be seen from the results, PIFD substantially outperforms the competitive methods.

The phase-based methods (AFD, FHAFD and VBMFD) cannot obtain good performance because of the inadequate phase normalization. Since they struggle to accurately determine the nominal orientation, the matching of FD is prone to errors.

Methods that align shapes during matching achieve better retrieval results, at a cost of a more computationally demanding matching process. Moreover, aligning shapes pairwise negatively affects retrieval performance. Shapes may be aligned using orientations that are not their nominal, thus "false" similarity may be computed.

The magnitude-based methods have promising performance, compactness and are simple for computation and comparison. This explains why magnitude-based FD were used more frequently in practical applications when compared to phase-based FD. However, since they discard phase information, they can hardly increase their retrieval performance above a certain level.

Computational complexity of PIFD is $O(nlogn)$ during descriptor extraction, and $O(n)$ during matching. The complexity is the same as for the most of the magnitude-based FD (NCC, CPAF and many others). The complexity of PIFD is also lower than of WARP and PDAFD (which is $O(n^2)$ in the matching stage), and CRFD ($O(nlogn)$ in the matching stage).

It is worth pointing out that PIFD and all other techniques addressed in the paper are global techniques, so they exhibit low performance when articulations of the shape are present. Thus, PIFD is slightly outperformed by structural, rich and hierarchical shape descriptors such as: Shape tree ($BE = 87.70$) [14], Hierarchical Procrustes Matching (HPM) ($BE = 86.35$) [12], Inner-Distance Shape Context ($BE = 85.40$) [15], Height functions ($BE = 90.35$) [16] etc. They achieve exceptionally good Bulls-eye scores on the MPEG-7 dataset, but because they are "rich" descriptors, consequently suffer from higher computational complexity (ranging from $O(n^2)$ to $O(n^4)$). Moreover, PIFD outperforms many descriptors that have higher computational complexity ($O(n^2)$), such as Generative Models ($BE = 80.03$) [17], Shape Context ($BE = 76.51$) [18], Curvature Scale Space (CSS) ($BE = 75.44$) [19] etc. In addition, PIFD achieves a solid 95.47 classification score on the Leaf dataset (almost as Shape tree [14]), and 91.64 on ETH-80 dataset (better than Height functions [16]), while slightly underperforms on the Kimia99 dataset which contains many shape articulations. A comprehensive evaluation of PIFD's retrieval performance on different datasets is not given in this paper due to the space limit.

4. CONCLUSION AND FUTURE WORK

Experiments have demonstrated that PIFD is a versatile global contour-based shape descriptor, characterized by solid performance, simple extraction and matching. Moreover, pseudomirror points proved to be promising shape orientation reference. They may be used wherever starting point and rotation invariance are needed. Combined with the scale and translation invariance of the Fourier descriptors, they can be used to normalize the contour in order to implement other more complex shape description techniques.

However, experimental results pointed out the main drawbacks of PIFD. It exhibits lower performance in scenarios with significant articulations, large artifacts or missing parts. Moreover, as many other contour based descriptors, it fails in region-based shape-retrieval tasks. As a part of future research, PIFD should be improved in two directions: it should exploit certain hierarchical structure in spatial domain in order to perform partial (local) matching of the shape, and it should be extended to a region-based descriptor.

Table 2. Performance of different Fourier-based methods on MPEG-7 dataset. Notation "method" refer to: "magnitude" (only magnitude is used), "phase" (phase is preserved), "phase+matching" (phase is preserved but phase normalization is done implicitly during matching).

Method	Bulls Eye score	method
PIFD	**82.03**	**phase**
PDAFD [11]	76.38	phase+matching
NCC [3]	75.75	magnitude
CPAF [2]	74.47	magnitude
CRFD [4]	72.57	phase+matching
WARP [8]	58.50	phase+matching
FHAFD [5]	51.34	phase
AFD [9]	41.08	phase
VBMFD [6]	28.89	phase

5. REFERENCES

[1] Dengsheng Zhang and Guojun Lu, "Study and evaluation of different Fourier methods for image retrieval," *Image and Vision Computing*, vol. 23, no. 1, pp. 33–49, Jan. 2005.

[2] Bin Wang, "Shape retrieval using combined Fourier features," *Optics Communications*, vol. 284, no. 14, pp. 3504–3508, July 2011.

[3] Emir Sokic and Samim Konjicija, "Novel fourier descriptor based on complex coordinates shape signature," in *Content-Based Multimedia Indexing (CBMI), 2014 12th International Workshop on*. IEEE, 2014, pp. 1–4.

[4] Fredrik Larsson, Michael Felsberg, and P-E Forssen, "Correlating fourier descriptors of local patches for road sign recognition," *Computer Vision, IET*, vol. 5, no. 4, pp. 244–254, 2011.

[5] FP Kuhl, "Elliptic Fourier features of a closed contour," *Computer graphics and image processing*, 1982.

[6] Yong Rui, Alfred C She, and Thomas S Huang, "Modified fourier descriptors for shape representation-a practical approach," in *Proc of First International Workshop on Image Databases and Multi Media Search*, 1996, pp. 22–23.

[7] E Persoon and K S Fu, "Shape discrimination using fourier descriptors.," *IEEE transactions on pattern analysis and machine intelligence*, vol. 8, no. 3, pp. 388–97, Mar. 1986.

[8] Ilaria Bartolini, Paolo Ciaccia, and Marco Patella, "WARP: accurate retrieval of shapes using phase of fourier descriptors and time warping distance.," *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, no. 1, pp. 142–7, Jan. 2005.

[9] K. Arbter, W.E. Snyder, Hans Burkhardt, and G. Hirzinger, "Application of affine-invariant Fourier descriptors to recognition of 3-D objects," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 12, no. 7, pp. 640–647, 1990.

[10] Paul Fieguth, Paul Bloore, and Adrian Domsa, "Phase-based methods for fourier shape matching," in *Acoustics, Speech, and Signal Processing, 2004. Proceedings.(ICASSP'04). IEEE International Conference on*. IEEE, 2004, vol. 3, pp. iii–425.

[11] Karl Sjöstrand, Anders Ericsson, and Rasmus Larsen, "On the alignment of shapes represented by fourier descriptors," in *Medical Imaging*. International Society for Optics and Photonics, 2006, pp. 61440Z–61440Z.

[12] Graham McNeill and Sethu Vijayakumar, "Hierarchical procrustes matching for shape retrieval," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*. IEEE, 2006, vol. 1, pp. 885–894.

[13] L.J. Latecki, R. Lakamper, and T. Eckhardt, "Shape descriptors for non-rigid shapes with a single closed contour," in *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*. 2000, vol. 1, pp. 424–429, IEEE.

[14] Pedro F Felzenszwalb and Joshua D Schwartz, "Hierarchical matching of deformable shapes," in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*. IEEE, 2007, pp. 1–8.

[15] Haibin Ling and David W Jacobs, "Shape classification using the inner-distance," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 2, pp. 286–299, 2007.

[16] Junwei Wang, Xiang Bai, Xinge You, Wenyu Liu, and Longin Jan Latecki, "Shape matching and classification using height functions," *Pattern Recognition Letters*, vol. 33, no. 2, pp. 134–143, 2012.

[17] Zhuowen Tu and Alan L Yuille, "Shape matching and recognition–using generative models and informative features," in *Computer Vision-ECCV 2004*, pp. 195–209. Springer, 2004.

[18] Serge Belongie, Jitendra Malik, and Jan Puzicha, "Shape matching and object recognition using shape contexts," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 4, pp. 509–522, 2002.

[19] Farzin Mokhtarian, Sadegh Abbasi, and Josef Kittler, "Efficient and robust retrieval by shape content through curvature scale space," *Series on Software Engineering and Knowledge Engineering*, vol. 8, pp. 51–58, 1997.