An Efficient Contour based Shape Descriptor
For Representation and Classification of Medical Images

Saad M. Saad¹, Shawkat K. Guirguis², and Ashraf M. Emam³
Department of Information Technology
Institute of Graduate Studies and Research
University of Alexandria
P.O. Box 832, El Shatby 21526, Alexandria, Egypt.

Abstract

Foundation of a robust image classifier depends on having an accurate image’s description based on its visual content. Recently, contour-based shape descriptors are employed for object-based representation due to its applicability to generic shapes and tolerance to noise and distortion. In this paper, an approach for extracting a hierarchical contour-based shape descriptor is described. It benefits compactness, flexibility, and suitability for image classification. This approach mingles the global description of moments with the multiresolution decomposition property of wavelet transform for robust shape representation and extracts approximate information about object’s shape. Furthermore, by integrating a suitable segmentation algorithm into the descriptor generation schema, the proposed approach turns into object-based. Experimental results show that the resultant descriptor provides promising results in terms of accuracy for medical image classification.

Keywords: Shape Descriptors – Shape Recognition – Wavelet based Shape Representation

1 Introduction

The irregular format of digital image tends to defy standard categorization and classification techniques. Classification problems generally, aim to identify the characteristics that indicate the group to which each instance belongs. Classification can be used both to understand the existing data and to predict how new instance will behave [1]. The benefits of image classification include better image storage and management, and optimized image-indexing scheme for fast and efficient image retrieval [2]. Traditional systems used to store and process multimedia images provide little to no means of automatic classification. Existing image retrieving systems [3, 4] limit classification mechanism to describing an image based on metadata such as color, texture, or shape features. Other systems use neural networks together with image features (descriptor) to provide automatic classification of images [5, 6]. Antonie et al [5] exploited the use of neural networks in classification of breast cancer image using back-propagation which proved to be less sensitive to database imbalance at a cost of high training time.

In medical application, shape is one of the most essential low-level images feature attributes that have received attention for classification [6]. Shape is a very significant feature in image perception. Human beings tend to perceive scenes as being composed of individual objects, which can be best identified by their shapes [7]. One main advantage of using the shape feature of a target is that it contains fewer parameters since it can be represented by one dimensional curve. Another advantage is that the shape feature can be normalized so that it has invariant properties. These properties greatly reduce the size of the target database and the complexity of target search algorithms [8]. In general, analysis of shapes requires the identification of proper descriptors that differentiate the shapes in the image for further classification and recognition.

¹ saad_darwish@yahoo.com
² shawkat_g@yahoo.com
³ Ashraf_emam@gmail.com
Many researches on shape indexing and description have been proposed [9, 10, 11, 12]. These techniques can be classified into two distinct categories: contour based and region based shape description. The first one describes the considered region by using its external characteristics while the second one represents the considered region by using its internal characteristics. Adamrk et al [13] compared the various shape representation schemes. The comparison shows that contour-based shape properties are more significant for retrieval of semantically similar objects than region-based shape properties. The contour-based descriptor is very efficient in applications where high variability in the shape is expected, caused by, e.g., deformation in the object, and tolerant to noise presented in the contour [14].

Two major shape description methods are considered: Wavelet theory-based method and the Moment theory-based method. As far as the former approach is concerned, the wavelet transform are generally used to describe the shape from its boundary. Wavelet transform produces a sequence of coefficients called Wavelet descriptors (WD). These coefficients represent the shape of an object in the frequency domain where the lower frequencies symbolize its general contour, and the higher frequencies represent its detailed contour. Only a few coefficients are enough to describe even complex shapes. Wavelet descriptors have the advantage over Fourier descriptors in that, they achieve localization of shape features in joint-space, i.e., in both spatial and frequency domains. The wavelet transform’s localization property concentrates on local differences between shapes within a limited set of coefficients [14]. The latter method uses region-based moments to characterize the contour of an object. It also considers global features such as center, variance, and symmetry. In general, this moment description is suited for a simple similarity search, but does not consider local contour variations, because it gives lower performance in the frequency domain. Thus, there have been efforts on incorporating another frequency-based technique to it for effective and compact contour representation. Babu et al [15] compared the performance of boundary-based representations, region-based representations and combined representations. Their experiments showed that combined representations outperformed any single representation.

Although system efficiency in terms of speed and computational complexity has been also subject of many researches, many related problems are still unsolved. A major shortcoming of the available shape description techniques is the lack of a sensible integration between segmentation and indexing when object-based description is envisaged. The research leading to this paper focuses on this problem, i.e. efficiency. The objective is to develop a scalable and efficient descriptor approach. This approach integrates a global moment’s description and the wavelet hierarchical decomposition property, to form a contour-oriented representation tool of shapes. The approach is independent to object’s translation, rotation, and scaling. This invariance is achieved by application of invariance transformation for the wavelet descriptor and the global nature of moment values for the moment descriptor. The generated descriptor is used in conjunction with neural networks to perform classification at various levels of detail.

This paper is organized as follows: Section 2 details the steps taken to extract contour descriptor. Section 3 introduces image modeling that depicts the proposed descriptor approach. Section 4 presents a system for medical image classification outlines the steps carried out by the neural network as it process an image. Selected results from the experimental evaluation are reported in section 5. Finally, the paper concludes with a summary in section 6.

### 2 Contour-based Shape Descriptor Extraction

The following subsections outline the steps taken to successfully extract the contour features vector. Figure 1 shows the block diagram of extracting the wavelet-based contour descriptor.

![Fig.1 Steps of extracting the wavelet-based contour.](image)

#### 2.1 Pre-Processing Stage

Medical images are difficult to interpret, and a pre-processing stage is necessary to improve the quality of the images and make the feature extraction stage more reliable [5]. Three techniques
are utilized in this work: a cropping operation, an image enhancement, and finally morphological noise removal. The first one is employed in order to remove the unwanted parts of the image usually peripheral to the area of interest. Image enhancement helps in qualitative improvement of the image with respect to a specific application.

In order to diminish the effect of over brightness or over darkness in the images and emphasize the image features, a widely used technique in the image processing to improve visual appearance known as histogram equalization is applied. Histogram equalization increases the contrast range in an image by increasing the dynamic range of the colors.

Regarding medical images, feature based on image object’s shape can differentiate images better. Alas, the image quantization to black and white produces noise around the shape boundary, therefore, the use of de-noising process is vital. The de-noising process eliminates isolated pixels and isolated small regions or segments. In the developed approach, a morphological closing filter is used in order to remove noise, which can be defined as [10]:

\[
B = B \ast K = (B \circledast K) \odot K ,
\]

where \( \circledast \) and \( \odot \) are the morphological dilation operator and morphological erosion operator defined, respectively, by

\[
B \circledast K = \bigcup_{k \in K} (k + B) \quad B \odot K = \bigcap_{k \in K} (k + B),
\]

where K is the following 3 x 3 binary mask

\[
K = \begin{bmatrix}
0 & 1 & 0 \\
1 & 1 & 1 \\
0 & 1 & 0
\end{bmatrix} \quad (1)
\]

an example of the three applied techniques is given in Figure2.

![Fig2. Pre-processing stage on an example image.](image1)

2.2 Segmentation Stage

For the feature extraction of contour descriptor, segmented image material has to be provided. Medical images are rich in both color and texture, and a wide range of them can be considered as a mosaic of regions with different colors and textures [16]. In this work, segmentation is achieved by Expectation Maximization (EM) algorithm applied on a model based on Gaussian mixture distribution of the color and texture feature space [17]. After running EM algorithm, each image pixel is labeled with the cluster values for which it attains the highest likelihood. A 3x3 max-vote filter is then applied to smooth the image and run a connected component algorithm to produce a set of homogeneous image regions (classes). Figure.3 gives an example of EM segmentation result with two classes.

![Fig3. Example of segmented image.](image2)

2.3 Contour Representation and Sampling

Edge detection and edge tracing are very important tasks in contour features extraction. Technically, edge detection is the process of locating edge pixels, and edge tracing is the process of following the edges [17]. In a first step an ordered sequence of contour points approximating the input shape is extracted. For this purpose, a contour detection algorithm similar to that of Canny [18] is applied to the segmented image. For some shapes, the shape boundary is not always connected; therefore a clockwise edge-linking algorithm is used to fill the gaps between boundary points and to ensure that a closed contour is extracted from the image [19].

![Class 1](image3)
![Class 2](image4)

In the next step, a vector of equally distributed points along the curve is extracted from the ordered sequence of contour pixels (sampling process). The sampling process not only normalizes the sizes of shapes but also has the effect of smoothing the shape [19]. The smoothing eliminates the noise in the shape boundary and the small details along the shape boundary as well. By varying the number of sampled points, the accuracy of the shape representation can be adjusted. The technique used begin with sampling the edge coordinates using a uniform number of points so that no matter how long a curve is, the number of sampling points is the same, this property is crucial for scale invariance. If there are m original
edge points \((x_0(i),y_0(i))\), \(i = 0, 1, \ldots, m-1\). The incremental curve length can be calculated as \([20]\):

\[
L = \sum_{i=1}^{m-1} \sqrt{(y_0(i) - y_0(i-1))^2 + (x_0(i) - x_0(i-1))^2}.
\]

The total length of the curve is denoted by \(L = l(m)\). The curve is then uniformly sampled at \(n\) points, i.e.

\[
x_{n}(j) = x_{0}(s(j-1)) + x_{0}(s(j+1))f, \quad y_{n}(j) = y_{0}(s(j-1)) + y_{0}(s(j+1))f, \quad j = 1, \ldots, m
\]

with

\[
s = \left[ \frac{L \times J}{n} \right], \quad f = \frac{L \times j / n - l(s)}{l(s+1) - l(s)}
\]

From this point, the contour of the shape is presented by the ordered sequence \((x(s), y(s))\) of contour points, where \(s\) denotes the position along the contour. In order to facilitate the use of the fast wavelet transform. The number of sampled points is chosen to be power of two.

### 2.4 Contour Normalization

Shape must be normalized before wavelet transform is performed. Edge coordinates are in general dependent on the displacement, scaling, and rotation of the target. A normalization process is performed so that the normalized edge coordinates have 2D shift, scale, and rotation invariant properties. To normalize the contour representation with respect to displacement, the curve is shifted so that the center of the curve is located at the origin. This is achieved by calculating the new coordinates \((x_n(j), y_n(j))\), using

\[
x_n(j) = x(s(j)-\bar{x}), \quad y_n(j) = y(s(j)-\bar{y})
\]

where \((\bar{x}, \bar{y})\) is the mean of the edge coordinates. To normalize the representation with respect to scaling and 2D rotation, edge points are transformed from Cartesian coordinates to Polar coordinates via

\[
R(j) = \sqrt{x_n(j)^2 + y_n(j)^2}, \quad \Phi(j) = \tan^{-1}\left(\frac{y_n(j)}{x_n(j)}\right)
\]

\(R(j)\) is then normalized so that it has the unit radius. Since all targets with different sizes have the same radius, they are scale-invariant. Finally, for rotational invariance, the coordinates are rotated so that the center of mass \(\sum_i R^2_k(i)\) for radius \(R\) is minimized \([19]\).

\[
\bar{R}_k(i) = \begin{cases} R(i+k), & 0 \leq i < N - k \\ R(i - N + k), & N - k \leq i < N \\
\end{cases}
\]

where

\[
k = \arg \min_k \sum_{i=1}^{N} \bar{R}_k(i).
\]

### 2.5 The Wavelet Descriptor

The wavelet transform is a tool that divides up data, functions, or operators into different frequency components and then studies each component with a resolution matched to its scale \([20]\). Therefore, the wavelet transform is anticipated to provide informative mathematical representation of many objects of interest. Wavelets have many favorable properties, such as vanishing moments, hierarchical and multiresolution decomposition structure, dilating relation, compactness of the support, decelerated coefficients, and a wide variety of basis functions \([1, 21]\). These properties provide considerably more efficient and effective solutions to many shape description problems.

Let \(\{Z^2_t\}\) be a discrete parameterised closed planar curve that represents the contour of an object of interest consisting of \(N\) points (obtained in previous section) of the form \([20]\)

\[
\{Z^2_t\} = \begin{bmatrix} R_t \\ \phi_t \end{bmatrix},
\]

or also, in matrix notation

\[
Z^2 = \begin{bmatrix} R \\ \phi \end{bmatrix}
\]

where the parameter \(t\) corresponds to the normalized arc length. If a level \(J_0\) partial one-dimensional “Discrete Wavelet Transform” (DWT) is applied independently to each component, the planar curve can be described in terms of the multiresolution analysis of \(\{R_t\}\) and \(\{\phi_t\}\):

\[
Z^2 = \begin{bmatrix} R \\ \phi \end{bmatrix} = \begin{bmatrix} w^T W_k \\ w^T \phi \end{bmatrix}
\]

\[
= \begin{bmatrix} v^T J_{0H} R_{J_{0}} \\ v^T J_{0H} \phi_{J_{0}} \end{bmatrix} + \sum_{j=1}^{J_{0}} w^T W_{J_{j}}
\]

\[
= \begin{bmatrix} v^T J_{0H} R_{J_{0}} \\ v^T J_{0H} \phi_{J_{0}} \end{bmatrix} + \sum_{j=1}^{J_{0}} w^T W_{J_{j}}
\]

4 In this paper, \(\{Z^2_t\}\) indicates that \(Z\) is bi-dimensional, not power operations. The same holds for matrix notation.
where \( w \) is an \( N \times N \) real-valued matrix defining the partial DWT, and \( W_R, W_\Phi \) are column vectors of length \( N \) in which the first \( (1-2^{-J_0})N \) elements are the wavelet coefficients, and the last \( 2^{-J_0}N \) elements are the scaling coefficients. The \( \Phi_j \) are detail vectors whose elements are related to changes in the \( R \) component at scale \( 2^{-j} \); and the smooth vectors \( v_{R,j_0}^T v_{R,j_0} \) represent averages at scale \( 2^{-j_0} \). The wavelet descriptor \( (WD) \) for the planar curve \( Z^2 \) is defined as the coefficients of the partial DWT for \( R \) and \( \Phi \), that is

\[
WD(Z^2) = \begin{bmatrix} W_R \\ W_\Phi \end{bmatrix}
\]

(9)

With this definition, a multiresolution representation of the closed curve has been obtained by using only a subset of wavelet coefficients consisting of primarily coarser scale components. In general, the two properties of wavelets that are exploited for shape characterization are the capability for detecting and representing local features, and the energy compaction in the transformed coefficients that make the description process more effect and accurate [1].

According to [22], the \( R \)-component of the contour points is most important, while the influence of \( \Phi \)-component could be neglected. Therefore, a contour is represented by a single one-dimensional wavelet transform \( W_R \). Because, the most important wavelet coefficients are the ones describing the lowest frequencies, only the first 10 coefficients are used as wavelet-based contour descriptor.

2.6 Global Descriptor of Moment

The distribution of different regions is a global feature used to characterize the image content. Thus, one way to apply this global feature is storing information for each region separately, such as its positions and sizes. In this work, the features to be considered as global descriptor are: Dispersions (see Eq.10) and Modification ratio (see Eq.11)[23]. Assume that the result of segmenting image \( I \) is the image \( S \) and that the pair \((x, y)\) represents a spatial location. The mass of the contour labeled with the class \( c \) after the segmentation is given by:

\[
m_c = \sum_s h_c(x, y) ,
\]

where

\[
h_c = \begin{cases} 1 & \text{if } S_{x,y} \in c \\ 0 & \text{if } S_{x,y} \notin c \end{cases}
\]

the center of mass (centroid) of the contour labeled as \( c \) is given by:

\[
x_c = \frac{\sum_{x,y} x h_c(x, y)}{m_c} , \quad y_c = \frac{\sum_{x,y} y h_c(x, y)}{m_c} ,
\]

(10)

The dispersion feature is a summation of distances. Therefore, for each contour \( i \), the centroid is computed and then their distances to the global class center \((x_c, y_c)\) are summed as follows:

\[
d_c = \sum_i \sqrt{(x_c-x_i)^2 + (y_c-y_i)^2} 
\]

The modification ratio is defined by the ratio of inscribed diagonal \( d_i \) that has the same second-moments as the region, to maximum diagonal \( d_m \). This value describes whether a contour is thin (oval like) or circular. The parameters inscribed and maximum diagonal, represent the dimensions of surrounding rectangle, as depicted in Figure 4.

\[
\text{Modification ratio} = \frac{d_i}{d_m} 
\]

(11)

3 Image Mathematical Modeling

An image is mathematically modeled according to a so-called object model \( O \) to describe the content of image regions [24]. Every image region is specified by an image object \( O_n \) of type \( O \), with \( n=1,\ldots,N \). The image is modeled as:

\[
M=(d, O_n),
\]

where \( d \) is the dispersion, and \( O \) is defined by the following 2-tuple:

\[
O=(W, G)
\]
where \( W \) presents the wavelet feature values, and \( G \) presents the moment feature values. Following the definition of image descriptors, which consist of the wavelet-based feature vector and the modification ratio per image region, plus the dispersion for entire image, each vector is associated with an index that corresponds to a particular category.

4 Descriptor Accuracy Evaluation

A system is developed for testing purposes. In this system, the contour-based shape descriptor extractor is followed by a “Learning Vector Quantization” neural network classifier for object identification based on its content. This system circumvents the low precision classification technique of other systems [5, 6] by examining the actual objects within an image and using them to discover relationships that reveal information useful in classifying the entire image, Figure 5 shows an overview of the classification system.

For the evaluation, samples of morphological medical image are selected. Sixty five percent of which are used to train the neural network, while the rest are used to test the trained network. Five different network configurations are developed. After training of each neural network configuration, it is tested on all the images using the contour-shape descriptor. Table 1 shows the result of the testing the five configurations.

Table 1. Results from five neural network configuration of the LVQ Classifier.

<table>
<thead>
<tr>
<th>Network Configuration</th>
<th>R.M.S. (Training)</th>
<th>No. of Feature Vector (per region)</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.045</td>
<td>31</td>
<td>83.75</td>
</tr>
<tr>
<td>2</td>
<td>0.023</td>
<td>11</td>
<td>85.83</td>
</tr>
<tr>
<td>3</td>
<td>0.023</td>
<td>16</td>
<td>87.8</td>
</tr>
<tr>
<td>4</td>
<td>0.034</td>
<td>21</td>
<td>85.83</td>
</tr>
<tr>
<td>5</td>
<td>0.031</td>
<td>25</td>
<td>87.69</td>
</tr>
</tbody>
</table>

The second configuration delivers the highest classification rate, although it uses a lower numbers of feature vectors. Since the classification process is performed using non-complex contour descriptor, the classification time remains almost constant when the database grows.

To justify the results obtained from this study, the LVQ classifier has been compared with three other classifiers: K-means classifier, Multiple Layers Perceptron (MLP), and the Support Vector Machine (SVM). Using the same training image sets, and the same contour vector is fed to the classifiers. The results of testing the classifier can be founded in Figure 7.

The results obtained indicate that the LVQ compared favorably with three other classifiers and outperformed them on average by 15%.
5 Conclusions

In this paper, an integrated contour-based shape description approach that demonstrates the feasibility of using wavelet transform with global description of moments for automatic medical images classification is presented. It is shown that descriptor defined from contour features extracted during the segmentation is very suitable for object-based description. Image objects are represented with a small number of the largest magnitude wavelet coefficients that describe the local features for each contour and global features of moments allowing for an efficient contour representation.

The developed descriptor is compact and flexible abstraction that inherently fits any classification strategy. The techniques of target contour extraction from the pathological medical images and the shape normalization for 2-D shift, scale, and rotation invariance are also described.

The advantages of using the developed shape descriptor are threefold. First, the descriptor represents well each region’s overall color and texture distribution because they are based on the features generated by the segmentation algorithm. Second, the shape descriptor is compact: the contour-features distribution of each region is summarized by approximately 11 values. Finally, the descriptor is flexible abstractions for which it can define a hierarchical classification strategy. This strategy is performed by varying the number of wavelet coefficients used to describe the contour. Experimental evaluation of the developed descriptor shows its effectiveness.

References


