Shape Based CBIR Model using Segmentation Techniques

Ayush Sharma
15BCE1335
3rd Year, B. Tech, SCSE
VIT Chennai
ayush.sharma2015@vit.ac.in

Shashank Pathak
15BCE1287
3rd Year, B. Tech, SCSE
VIT Chennai
shashank.pathak2015@vit.ac.in

Shashank Shekhar Shukla
15BCE1012
3rd Year, B. Tech, SCSE
VIT Chennai
shashankshekhar.shukla2015@vit.ac.in

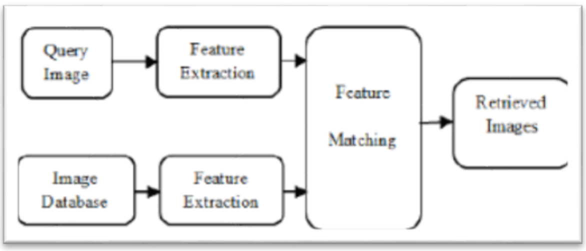
***Abstract*— CBIR (Content Based Image Retrieval) is a technique to perform retrieval of images from a large database based on the query image features. Traditional Image Search Technique use metadata such as tags, keywords to perform a retrieval whereas CBIR model uses the content in the image to retrieve the images. The word content with respect to images means various features of the image like color, texture, histogram. One such feature set is Shape Based feature which is used to find the shapes of objects in the image and used as a descriptor for shapes of objects. We will be using Image Segmentation Methods to segment the images and discriminate various objects present in the image. We will implement a CBIR based Model using Shape Based Features as descriptors on the segmented images. We will be using Cell Dataset provided by the Murphy Labs for implementation of this model.**

Keywords—CBIR, shape features, segmentation, thresholding,

Watershed, edge detection, contour

I. Introduction

CBIR model finds the similarity between query image and images in the database using the content of the image. The content of the image is decided by the different set of features used to describe the image. The architecture of CBIR model is shown below.



The CBIR model comprises of finding the features which describe the contents of the image and these features are extracted for the whole set of images in the database and these features serve as a measure for similarity and they serve as descriptors for the images [1][2][3]. When a query image is received by the system the same set of features (as used for feature extraction of images in the database) are extracted for the query image.

Then the feature of the query image is matched with the features of the images in the image database. The similarity between the images is calculated and the images are sorted according to the similarity value. The images with higher similarity are retrieved by the CBIR system. There are different similarity metrics used for feature matching like distance based metrics. Distance based metric include various distances like Euclidean, Manhattan and others. This is the basic working of a CBIR model.

There are various types of features used for calculating the content of the image. Color Based Features are also used as descriptors. Color space is used to represent color images. The RGB space is where the gray level intensity is represented as the sum of red, green and blue gray level intensities. There are so many methods used to retrieve the color feature. They include color histogram, color moment. Color Histogram is the commonly used method for color feature extraction in digital images [4][5]. In CBIR systems Color histograms are widely used in the image retrieval area. It is one of the most common methods for predicting the features of an image. The great advantages of the Color are speed and low memory space. Color histogram method is invariant to rotation but it is not invariant to scaling. It also varies with the angle of view. The color moments widely used are mean, standard deviation, and kurtosis. Color moments are mainly used for color indexing. The other advantages of color moments are: - they are good under lighting conditions, the requirement for their storage are very less.

Texture is another important property of images. Texture is basically a regular repetition of an element or pattern on a surface. The six texture properties are coarseness, contrast, directionality, line likeness, regularity and roughness. Texture Based Features are also used in CBIR. Texture Based Features are mainly of two types of methods Structural and Statistical. Structural methods, including morphological operator and adjacency graph, describe texture by identifying structural primitives and their placement rules. They tend to be most effective when applied to textures that are very regular. Statistical methods, including Fourier power spectra, co-occurrence matrices, shift-invariant principal component analysis (SPCA), Tamura feature, Word decomposition, Markov random field, fractal model, and multi-resolution filtering techniques such as Gabor and wavelet transform, characterize texture by the statistical distribution of the image intensity.

Shape based features are used for representing and describing an image and its objects or regions of interest [7][8][9]. We will be using segmentation to find region of interest in our dataset and use Shape based feature on the segmented images to retrieve the image based on the query image. A binary object can be defined as the following mathematical formula:

$$O_i(x, y) = \begin{cases} 1 & \text{if } f(x, y) \in O_i \\ 0 & \text{otherwise} \end{cases}$$

There are various shaped based features that can be used for the binary objects in the image like

a. Area

The area of the object can be calculated by

$$A_i = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} O_i(x, y)$$

b. Centroid

The centroid of the image can be calculated by

$$\bar{x}_i = \frac{1}{A_i} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x O_i(x, y)$$

$$\bar{y}_i = \frac{1}{A_i} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} y O_i(x, y)$$

c. Euler Number

The Euler number of an image can be found by subtracting the number of connected components from the number of holes.

E = C – H

d. Perimeter

The binary object has value of pixels equal to 1 so we can find the perimeter of the object by counting the number of pixels which have value 1 and have neighbouring pixels with value as 0. This will give the number of pixels comprising the boundary of the image and can be successfully used to find the perimeter of the object.

e. Thinness Ratio

$$T_i = \frac{4\pi A_i}{P_i^2}$$

Where A_i is the area of the object and P_i is the perimeter of the image.

f. Image Moment

image moment is a certain particular weighted average of the image pixels' intensities, or a function of such moments. Image moments have some particular property and can be used for interpretation. The 2D moment of order (p + q) of a digital image f (x, y) is defined as

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q f(x, y)$$

where M and N are the image height and width, respectively, and p and q are positive nonzero integers.

Central moments are the translation-invariant equivalent of moments. They are defined as

$$\mu_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x - \bar{x})^p (y - \bar{y})^q f(x, y)$$

Image moments can be used to define features of the image like area, centroid and information about the orientation of the image.

g. Convex Hull

A set is said convex if the straight line connecting any two points of the set lies entirely within A. Convex Hull of set S is the smallest convex set A that contains S. The set difference A-S is called the convex deficiency of S.

Shape Based feature discussed above will be used on the region of interest.

II. Related Works

Many systems use segmentation to calculate image features in order to build a Content-based image retrieval system. In fact, the segmentation step is necessary to extract information from images. Indeed, how is it possible to find an object in an image without first extracting its own regions. Several kinds of segmented region may be extracted according to the homogeneity predicates used: color, texture or else semantic criteria. Many approaches exist and many of them propose an adapted and optimized version of some well-known classical algorithms or other. Even if a segmentation step is required in a feature extraction process, the expected results from this low-level process still have to be discussed. Of course, if the segmentation was able to extract semantic regions, the discussion would be close. However, this is not the case and we cannot expect from the future such an algorithm because of the underlying diversity. If we consider the semantic extraction impossible, we have to first estimate the robustness of this segmentation step in order to really select the suitable parameters during the feature extraction. Nevertheless, at this point of research in image retrieval by content, the question of using or not using segmentation may be asked. And, in this case, which method or approach may be used. Whatever these interrogations are often inquired, few objective studies permit to answer it. Our goal in this paper is then to propose an objective evaluation of the stability of classical image segmentation process. Our goal is not to set up a new evaluation protocol but this paper must be considered as a punctual step forward to overcome the use of segmentation algorithms in image retrieval by content. In the context of natural scene images, we will first merely measure the stability of some algorithms made for feature extraction, without judge the quality of segmentation itself.

Many traditional techniques have been proposed for the segmentation of nuclei in histopathology images, ranging from simple background subtraction and color threshold techniques to much more sophisticated approaches, such as marked point processes [4][5]. Many of these methods have been recently reviewed. More recently, advances in deep convolutional neural networks (CNN), and in particular in their optimization have made them become the state-of-the-art model for object recognition. Indeed, deep neural networks have been used for the localization and classification of nuclei in histopathology data from breast and colon cancer, but these previously published approaches do not allow for segmentation.

On the other hand, deep CNN have also been used for semantic segmentation in other fields, where use" deconvolution layers" and up-sampling in order to directly output binary images. Learning based methods rely on annotated data sets. But for the problem of nuclei segmentation in histopathology images, there are only few manually segmented images freely available. There are many partially annotated datasets available. However, the problem with those datasets are that they do not cover the morphological variability typically encountered in real histopathology data.

Thus, to overcome these problems, various types of content-based image retrieval (CBIR) have been proposed over the last few decades. Unlike text-based retrieval, CBIR indexes images using color, texture, shape, and sound, which are then used for retrieval instead of keywords. Ideally, the goal is to create an interactive system for retrieving images that is semantically related to the user's query from the database. More recent research has also focused on region-based retrieval that allows the user to specify a particular region or object in an image and request the system to retrieve similar images containing similar regions.

Most existing region-based image retrieval systems rely on image segmentation and require extraction of the region of interest (ROI), which occupies a large portion of the entire image. Therefore, semantic ROI segmentation is essential for efficient region-based image retrieval.

In the medical field, region-based image retrieval is also helpful for diagnostic purposes. For example, diagnosis systems based on cytology and histophysiology are used to analyze tissue specimens to detect lesions as an early signal of latent cancer. Plus, measuring the cell cycle using a diagnosis system can enhance the effectiveness of drug discovery and development. However, existing diagnosis systems are restricted when dealing with cells and due to the subjective variance of an observer. Therefore, an integrated diagnosis system with an automatic aid method was recently developed to assist with the detection of cancer.

For a semantic analysis, an automatic aid method requires an ROI-based approach rather than a pixel-based approach to detect an abnormal nucleus or lesion. In particular, the ROI segmentation is a crucial preprocess to enable successful cell classification or diagnosis.

Comaniciu and Meer developed the Image Guided Decision Support system to analyze tissue structures and organ states to support diagnosis and identify factors in clinical pathology. The system extracts an ROI within the attention scope using the mean-shift segmentation method. However, to extract the ROI before segmentation, the attention window must be defined by hand.

Chen et al.6 developed a method of image analysis to resolve the problems of touching cells and ambiguous correspondence, resulting in a computational bio image system that facilitates the automated segmentation, tracking, and classification of cancer cell nuclei in time-lapse microscopy images.

Tscherepanow et al.10 proposed a method for classifying segmented regions in bright field microscope images. However, since an active contour is used for the cell segmentation, the performance can deteriorate when an image contains cells with a complex structure.

III. USED METHODS

We have analyzed and implemented various segmentation methods. The basic ones are Thresholding, Adaptive Thresholding and Otsu Binarization. We have also implemented edge detection techniques based segmentation. We have also implemented some advanced techniques such as Kmeans clustering based segmentation and Watershed method.

1. Thresholding

   It is the simplest method of image segmentation. In this method, we choose a threshold value to create binary image. The intensities above the threshold are maximised and the ones below are minimised. For example, anything that is greater than 127 in the grayscale, can be set to 1 in the binary image and anything that is less than or equal to 127 in the grayscale image can be set to 0 in the binary image.

2. Adaptive Thresholding

   In adaptive threshold unlike fixed threshold, the threshold value at each pixel location depends on the neighbouring pixel intensities. To calculate the threshold T(x, y) i.e. the threshold value at pixel location (x, y) in the image, we perform the following steps –

   a. A bxb region around the pixel location is selected. b is selected by the user.
   b. The next step is to calculate the weighted average of the bxb region. We can either use the average (mean) of all the pixel location that lie in the bxb box or we can use a Gaussian weighted average of the pixel values that lie in the box. In the latter case, the pixel values that are near to the center of the box, will have higher weight. We will represent this value by WA(x, y).
   c. The next step is to find the Threshold value T(x, y) by subtracting a constant parameter, let's name it param1 from the weighted average value WA(x, y) calculated for each pixel in the previous step. The threshold value T(x, y) at pixel location (x, y) is then calculated using the formula given below –

   T(x, y) = WA(x, y) - param1

3. Otsu Binarization

   This is a really advanced Thresholding technique. It follows the mechanism of Gaussian Mixture Models (GMM). In Otsu's method we exhaustively search for the threshold that minimizes the intra-class variance (the variance within the class), defined as a weighted sum of variances of the two classes:

$$\sigma_w^2(t) = \omega_0(t)\sigma_0^2(t) + \omega_1(t)\sigma_1^2(t)$$

   Where, w1 and w2 are the weights of the classes, t is the threshold value and sigma(s) are the variance of the classes.

4. Edge Detection Based segmentation

   There are a lot of methods for edge detection. Faster ones being Sobel, Canny and Laplacian. We have implemented these methods to detect the edge of the object of interest and then applied Thresholding for Binarization.

   Most of these algorithms use first or second derivative gradient measure and apply convolution. The kernels and formula for these methods are as follows:

   a. Sobel

$$\mathbf{G}_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * \mathbf{A} \quad \text{and} \quad \mathbf{G}_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * \mathbf{A}$$

   b. Canny

$$H_{ij} = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{(i-(k+1))^2 + (j-(k+1))^2}{2\sigma^2}\right); 1 \le i, j \le (2k+1)$$

   c. Laplacian

0	-1	0
-1	4	-1
0	-1	0

The laplacian operator

-1	-1	-1
-1	8	-1
-1	-1	-1

The laplacian operator (include diagonals)

5. Kmeans clustering based segmentation

   k-means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster. This results in a partitioning of the data space into Voronoi cells. Kmeans can be used to binarize the image by taking k=2 and clustering the pixels into 2 clusters according to their intensities.

   Given a set of observations (x1, x2, …, xn), where each observation is a d-dimensional real vector, k-means clustering aims to partition the n observations into k (≤ n) sets S = {S1, S2, …, Sk} so as to minimize the within-cluster sum of squares (WCSS) (i.e. variance). Formally, the objective is to find:

$$\underset{\mathbf{S}}{\arg\min} \sum_{i=1}^{k} \sum_{\mathbf{x} \in S_i} \|\mathbf{x} - \boldsymbol{\mu}_i\|^2 = \underset{\mathbf{S}}{\arg\min} \sum_{i=1}^{k} |S_i| \operatorname{Var} S_i$$

   Then, Thresholding is done to create binary images using the clusters as separated pixels.

6. Watershed
   Watershed is a transformation defined on a grayscale image. The name refers metaphorically to a geological watershed, or drainage divide, which separates adjacent drainage basins. The watershed transformation treats the image it operates upon like a topographic map, with the brightness of each point representing its height, and finds the lines that run along the tops of ridges.
   The basic idea consisted of placing a water source in each regional minimum in the relief, to flood the entire relief from sources, and build barriers when different water sources meet. The resulting set of barriers constitutes a watershed by flooding.

   The Priority-Flooding algorithm works on a gray scale image. During the successive flooding of the grey value relief, watersheds with adjacent catchment basins are constructed. This flooding process is performed on the gradient image, i.e. the basins should emerge along the edges. Normally this will lead to an over-segmentation of the image, especially for noisy image material, e.g. medical CT data. Either the image must be pre-processed or the regions must be merged on the basis of a similarity criterion afterwards.
   a. A set of markers, pixels where the flooding shall start, are chosen. Each is given a different label.
   b. The neighbouring pixels of each marked area are inserted into a priority queue with a priority level corresponding to the gradient magnitude of the pixel.
   c. The pixel with the lowest priority level is extracted from the priority queue. If the neighbours of the extracted pixel that have already been labelled all have the same label, then the pixel is labelled with their label. All non-marked neighbours that are not yet in the priority queue are put into the priority queue.
   d. Redo step 3 until the priority queue is empty.

   The non-labelled pixels are the watershed lines.

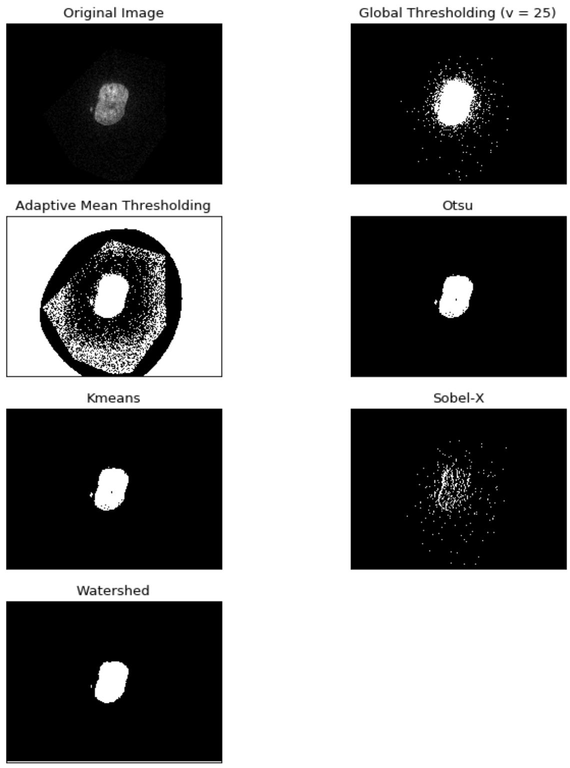
After segmentation, we have done some morphological enhancements to prepare the image for contour extraction. We have used Dilation and Opening methods for enhancement.

1. Dilation
   In binary morphology, dilation is a shift-invariant (translation invariant) operator, equivalent to Murkowski addition. The dilation operation usually uses a structuring element for probing and expanding the shapes contained in the input image. The structuring element or the kernel is used to expand the focused pixel or duplicate it into its neighbours so that the object seems dilated or expanded.
2. Closing
   It is the combination of dilation and erosion. Actually, it is the erosion of the dilation of an image. Dilation operation removes the holes that are smaller than structuring element and erosion operation restores the shape of remaining objects. The object seems to be expanded and filled up.

Finally, we have implemented Contour extraction and shape based features which are explained in detail in experimentation.

IV. Experimental Results

We have used the cell dataset from murphy labs for creating a shape based CBIR model using segmentation techniques. For this we started with analysis and implementation of various segmentation techniques of some sample images. The following shows the segmentation on an image of Nucleus.

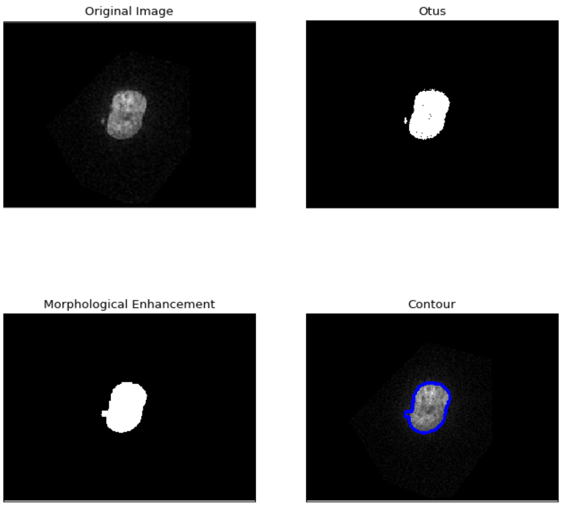


As we can see in this figure, the Otsu Binarization, Kmeans and Watershed are performing best for our experiment.
But since we want to create a content based image retrieval system, we want it to be efficient and fast. Out of these 3

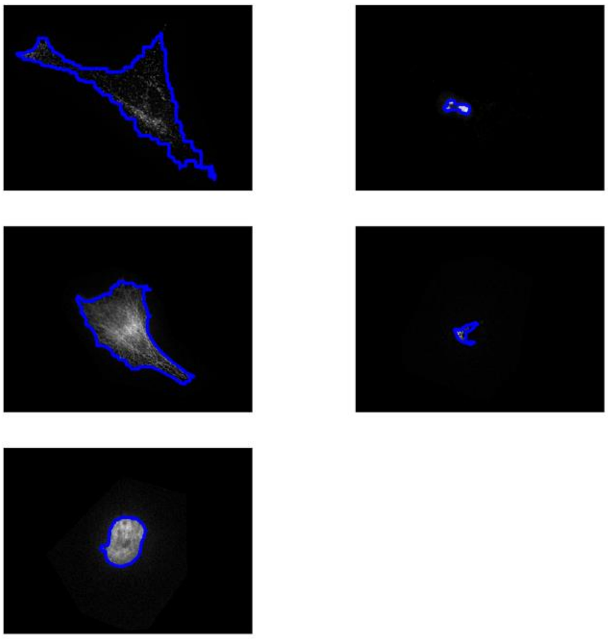
algorithms, Otsu Binarization was the fastest thus we have used it in our further experimentation.

Now after we have acquired the best algorithm for segmentation, we need to import the dataset and the query images, perform the segmentation and extract features for our CBIR model.
We have first applied some morphological enhancements on these images. We have used Closing method to fill the gaps in the images and to dilate them a little for easy contour extraction.

For contour extraction, we have used OpenCV package. The contour of the object is used to extract the moments of the image.
An image moment is a certain particular weighted average of the image pixels' intensities, or a function of such moments. Image moments are useful to describe objects after segmentation. Simple properties of the image which are found via image moments include area (or total intensity), its centroid, and information about its orientation.
We have extracted a collection of 24 moments from each segmented image. These moments are used to extract shape based features as explained above. The contour of the sample image is as follows:



We have about 500 images in our dataset belonging to 5 classes: Microtubules, Endosome, Nucleus, Golgi gia and Nucleolus. We have taken one image from each class as query image and performed the CBIR model. The contours for the query images are as follows:



Clearly we can see that the contour extraction works fine for this dataset. The moments were calculated for all these query images and were used in CBIR. The precision was calculated for these images and the average of these precision values was taken for the evaluation of our CBIR Model.

V. Conclusion

We have successfully constructed a shape based CBIR model using segmentation techniques. We have implemented the CBIR model for various query images and have achieved an average precision of 66%.

This Model can definitely be improved using a combination of shape and texture based features because the density of tissues may play a great role in the classification of the cell based images in the given classes.

Our Model is really fast because of the use of best segmentation method for this dataset and we have achieved these results by deeply analysing the mechanism of various segmentation techniques.

To improve this CBIR model we can also use data parallelism. The segmentation of the images can be done parallel using threads. The image can be separated into quadrants for the implementation of 4 threads and then each thread would perform segmentation on its quadrant. Finally, these quadrants can be merged into a single image and used for further computation.

REFERENCES

[1] Chen X, Zhou X, Wong STC. Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy. IEEE Trans Biomed Eng. 2006;53:762–766. doi: 10.1109/TBME.2006.870201.

[2] Tscherepanow M, Zollner F, Kummert F. Classification of segmented regions in brightfield microscope images. Proceedings of the IEEE International Conference on Pattern Recognition. 2006;3:972–975.

[3] Comaniciu D, Meer P. Mean-shift: a robust approach toward feature space analysis. IEEE Transactions on Pattern Recognition and Machine Intelligence. 2002;24:603–619. doi: 10.1109/34.1000236.

[4] Comaniciu D, Meer P, Foran D. Shape-based image indexing and retrieval for diagnostic pathology. Proceedings of the International Conference on Pattern Recognition.

[5] Ropers S-O, Bell AA, Wurflinger T, Bocking A, Meyer-Ebrecht D. Automatic scene comparison and matching in multimodal cytopathological microscopic images. Proceedings of the IEEE International Conference on Image Processing. 2005;1:1145–1148.

[6] L.Dorini et al, “White blood cell segmentation using morphological operators and scale-space analysis”, IEEE conference on Computer Graphics and Image Segmentation,2007.

[7] Lei Zheng et al, “Design and Analysis of a ContentBased Pathology Image Retrieval System”IEEE Transactions on Information Technology in Biomedicine, vol.7, no.4, December 2003

[8] J. Kong, L. Cooper, T. Kurc, D. Brat, J. Saltz, "Towards building computerized image analysis framework for nucleus discrimination in microscopy images of diffuse glioma", Proc. IEEE 33rd Annu. Int. Conf. Eng. Med. Biol. Soc., pp. 6605-6608, Boston, MA, USA, Aug. 30–Sep. 3, 2011.

[9] T. Mouroutis, S. J. Roberts, A. A. Bharath, "Robust cell nuclei segmentation using statistical modeling", Bioimaging, vol. 6, pp. 79-91, 1998.

[10] S. Brandt, Use of shape features in content-based image retrieval, 1999.

[11] Distasi R, Nappi M, Tucci M, Vitulano S: Context: A technique for image retrieval integrating CONtour and TEXTure Information. Proceedings of the IEEE International Conference on Image Analysis and Processing 224–229, 2001