A NEW MULTISCALE, CURVATURE-BASED SHAPE REPRESENTATION TECHNIQUE FOR CONTENT-BASED IMAGE RETRIEVAL

JanKees van der Poel
Instituto de Educação Superior da Paraíba – IESP
João Pessoa – Paraíba – Brazil

Leonardo Vidal Batista, Carlos Wilson Dantas de Almeida
Departamento de Informática – Universidade Federal da Paraíba
João Pessoa – Paraíba – Brazil

Keywords: Curvature Scale Space, CBIR, correlation coefficient, resampling, resizing, shape classification, full curvature values.

Abstract: This work presents a new multiscale, curvature-based shape representation technique for planar curves. One limitation of the well-known Curvature Scale Space (CSS) method is that it uses only curvature zero-crossings to characterize shapes and thus there is no CSS descriptor for convex shapes. The proposed method, on the other hand, uses bidimentional→unidimentional→bidimentional transformations together with resampling techniques to retain the full curvature information for shape characterization. It also employs the correlation coefficient as a measure of similarity. In the evaluation tests, the proposed method achieved a high correct classification rate (CCR), even when the shapes were severely corrupted by noise. Results clearly showed that the proposed method is more robust to noise than CSS.

1 INTRODUCTION

Since the beginning of the 90s there has been an increased research activity in the area of Content Based Image Retrieval (CBIR). Not only large research teams, such as IBM with the QBIC project, but also small project groups in the academic and industry worlds have devoted themselves to this task.

Image retrieval systems are supposed to retrieve images in an effective manner based on a user's input query. This work aims to investigate ways of providing fast, easy and *reliable* access to images in electronic databases. It also intends to create a classification and retrieval system that will help to browse, search, classify and retrieve images in large digitized databases.

Image similarity is very subjective and relative to each person actual needs: the more a person encounters an object, the more detailed and differently he or she can describe it.

There are several simple image retrieval solutions to the query problem. One of them is to annotate images with text and then use a traditional textbased search and retrieval. While fast, this is not effective when dealing with large collections of complex images: variability of interpretation is enormous and also is the human effort required for database annotation. Effective image retrieval systems should exploit image attributes such as color distribution, motion, shape (Niblack et al., 1993), structure and texture (Batista and Meira, 2004).

Image retrieval is based on an ordering of match scores obtained by searching through a database. The key challenges in building a retrieval system are the choice of attributes, their representations, query specification methods, match metrics and indexing strategies.

Shape matching performs an important issue in image retrieval by recognizing and classifying the images shapes. Some contour shape representations include eccentricity, circularity (Niblack et al., 1993), chain code (Freeman and Saaghri, 1978), centroid distance, cumulative angles (Davies, 1997), Fourier descriptors, Wavelet descriptors (Zahn and Roskies, 1972; Persoon and Fu, 1977; Kauppinen et al., 1995; Tieng and Boles, 1997; Yang et al., 1998; Loncaric, 1998) and Curvature Scale Space (CSS) descriptors (Mokhtarian and Mackworth, 1992; Mokhtarian, 1995; Mokhtarian et al., 1996b; Abbasi et al., 2000).

It is generally accepted that CSS descriptors achieve a good compromise between representation power and computational efficiency (Dudek and Tsot-

van der Poel J., Vidal Batista L. and Wilson Dantas de Almeida C. (2006).
A NEW MULTISCALE, CURVATURE-BASED SHAPE REPRESENTATION TECHNIQUE FOR CONTENT-BASED IMAGE RETRIEVAL.
In *Proceedings of the First International Conference on Computer Vision Theory and Applications*, pages 401-406
DOI: 10.5220/0001372304010406
Copyright © SciTePress

sos, 1997; Daoudi and Matusiak, 2000; Mokhtarian et al., 1996a). The curvature of a planar curve has perceptual characteristics that have proven to be useful for shape recognition (Pomerantz et al., 1977), being one of the most powerful tools for representation, interpretation and recognition of objects in an image (Mokhtarian and Mackworth, 1992; Pavlidis, 1980; Mokhtarian and Mackworth, 1986; Dudek and Tsotsos, 1997)). Such characteristics made CSS to be selected as one of the MPEG-7 contour shape descriptors (Mokhtarian and Bober, 1998).

This paper presents a new method for shape characterization suitable to be used as a retrieval tool in large image collections. This method is based on the classical CSS method but but *all* the available curvature information is used to perform shape matching, instead of using *only* the curvature zero-crossing points. It also uses the correlation coefficient as a measure of similarity between shapes, allowing a high correct classification rate, even when the shapes are severely corrupted by uniform random noise.

The rest of the paper is organized as follows: Section 2 describes and discusses the classical CSS descriptor; Section 3 describes the proposed method and compares it with the classical CSS method; Section 4 shows qualitative results of retrieval tests obtained from direct comparison between the classical CSS method and the proposed method; and Section 5 presents a discussion of the results and the concluding remarks.

2 CURVATURE SCALE SPACE METHOD

The classical Curvature Scale Space method for contour representation captures, describes and compares characteristic shape features of objects based on their closed contours. It is a multiscale representation of the inflexion points of a closed contour (Mokhtarian and Mackworth, 1992), is considered a reliable and very fast method to perform shape analysis in large databases and has a number of important properties, such as:

1. It captures the main features of a shape, enabling similarity-based retrieval;
2. It reflects properties of the human visual system perception and offers good generalization;
3. It is robust to non-rigid motion, partial occlusion of the shape, noise and changes in scale and orientation;
4. It is robust to perspective transformations that result from common changes of camera parameters in images and video; and
5. It is compact, reliable and fast.

The idea behind the classical CSS is that a contour can be represented by its curvature values. It is possible to compute the curvature at each contour point based on its neighboring points (Haralik and Shapiro, 1992). The classical CSS takes into account only those points where the curvature goes from a positive to a negative value, or vice-versa (curvature zero-crossings), as illustrated in Figure 1.

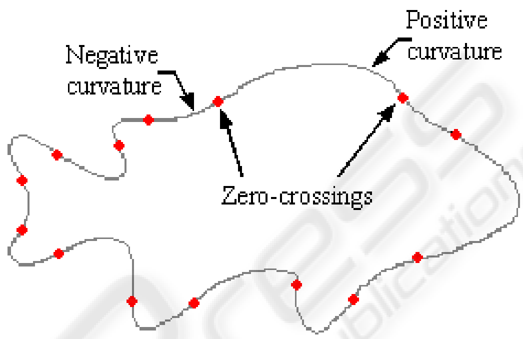

Figure 1: Curvature zero-crossings.

To compute the classical CSS representation of a given object, its contour Γ is initially obtained and parametrized as in Equation 1.

$$\Gamma(u) = \{(x(u), y(u))\,|u \in [0,1]\} \qquad (1)$$

where u is the normalized arc length parameter, varying between 0 and 1, and $(x(u), y(u))$ are parametric coordinates sampled from the contour at equidistant values of u, starting at an arbitrary contour point and following in counterclockwise direction.

Convolving the parametric coordinates of Γ with a progressively higher standard deviation σ 1-D Gaussian kernel generates the *evolved version* Γ_σ (Mokhtarian and Mackworth, 1992) of Γ, defined as in Equation 2.

$$\Gamma_\sigma(u,\sigma) = (X(u,\sigma), Y(u,\sigma)) \qquad (2)$$

where $X(u,\sigma) = x(u) * g(u,\sigma)$, $Y(u,\sigma) = y(u) * g(u,\sigma)$, "$*$" is the convolution operator and $g(u,\sigma)$ is a Gaussian with standard deviation σ.

The curvature $k(u,\sigma)$ of Γ_σ is given by Equation 3 (Mokhtarian and Mackworth, 1992):

$$\kappa(u,\sigma) = \frac{X_u(u,\sigma)\cdot Y_{uu}(u,\sigma) - X_{uu}(u,\sigma)\cdot Y_u(u,\sigma)}{(X_u(u,\sigma)^2 + Y_u(u,\sigma)^2)^{3/2}} \qquad (3)$$

where $X_u(u,\sigma) = x(u) * g_u(u,\sigma)$, $Y_u(u,\sigma) = y(u) * g_u(u,\sigma)$, $X_{uu}(u,\sigma) = x(u) * g_{uu}(u,\sigma)$ and $Y_{uu}(u,\sigma) = y(u) * g_{uu}(u,\sigma)$.

After each convolution step, the curvature zero-crossings of $k(u,\sigma)$ are located by computing the curvature for all contour points with Equation 3 using progressively higher values of σ and determining where the curvature goes from a positive to a negative value, and vice-versa. This is done until the curve Γ becomes completely convex.

Finally, the CSS Image of Γ is a binary image defined by the zero-crossing points of $k(u, \sigma)$, with u values in the horizontal axis and σ values in the vertical axis.

As each convolution step smooths the shape contour, the zero-crossings will group two by two, approach each other, merge and finally disappear, forming what is called a *CSS peak*. Each zero-crossing does not necessarily group with an adjacent zero-crossing: at the end smaller peaks can exist inside larger ones, due to contour sections delimited by two zero-crossings that are close together.

To perform a CSS shape-based retrieval, the similarity between two shapes is measured by the sum of the peak differences between all the matched peaks and the peak values of all the unmatched peaks (Mokhtarian, 1995; Mokhtarian et al., 1996a; Mokhtarian et al., 1996b; Abbasi et al., 2000).

Below there is a summary of the steps performed when using the classical CSS method:

1. Image segmentation;
2. Contour extraction;
3. Shape scale normalization, done by sampling the shape boundary into a fixed number of points to allow matching shapes with different boundary sizes;
4. Shape curvature computation;
5. CSS Image computation; and
6. Shape matching.

Matching CSS Images is difficult because they usually have a different number of CSS peaks, these peaks are usually not matching and also can be ordered in a quite different way. Mirrored and flipped shapes need to be considered separately too.

In addition, the classical CSS method only captures local shape features, missing the global ones (which are important to shape representation too). To overcome this, global features such as eccentricity, circularity and number of CSS peaks should be combined in order to correctly describe the shapes (Zhang and Lu, 2001).

Due to the dependence on curvature zero crossings, convex objects may not be well represented with the classical CSS method. This means that shapes like circles, ellipsis or convex polygons may not be recognized using this method.

As a final drawback of the classical CSS method, the boundary sampling and the thresholding processes done when extracting the CSS peaks causes the CSS Image to not reflect the true number of convex (or concave) segments on the shape boundary.

3 PROPOSED METHOD – FULL CURVATURE SCALE SPACE METHOD

Considering all the classical CSS method problems, a new approach to retrieve an image based on its contour will now be described.

The proposed method basically performs the same computations of the classical CSS method. The main differences are:

1. Its unnecessary to perform the CSS Image computation step;
2. Full curvature information usage (and not only the zero-crossing information); and
3. The shape matching step, which uses a different approach.

The proposed method matches shapes using a *Full CSS matrix*. This matrix has all the curvature values computed from Equation 3 with progressively higher values of σ.

To perform shape matching, the CSS matrices of the shapes under analysis are computed and compared using the 2-D correlation coefficient r, shown in Equation 4.

$$r = \frac{\sum_{\sigma}\sum_{u}\left(A_{\sigma u}-\bar{A}\right)\left(B_{\sigma u}-\bar{B}\right)}{\sqrt{\left(\sum_{\sigma}\sum_{u}\left(A_{\sigma u}-\bar{A}\right)^2\right)\left(\sum_{\sigma}\sum_{u}\left(B_{\sigma u}-\bar{B}\right)^2\right)}} \tag{4}$$

where A, B are the CSS matrices of shapes A and B and $\bar{A}$, $\bar{B}$ their means, respectively.

As the classical CSS method, the proposed method suffers from the problem related to the the contour following arbitrary starting point. To solve this, a one column rotation is performed to the matrix belonging to the original shape (the one that is being compared to the other shapes in the database) each time the correlation coefficient between them is computed.

If one shape has deeper concavities than other, the σ value used to smooth them completely can be different, leading to curvature matrices with non-matching heights. In this case, directly computing r with Equation 4 will not be possible. To overcome this, the curvature matrix with the smaller number of lines is resampled to the same number of lines of the higher curvature matrix.

For noise contaminated shapes, initially the proposed method has results that are worse than those achieved by the classical CSS method: as the proposed method takes into account *all* the curvature information, and as the first lines of the Full CSS matrix

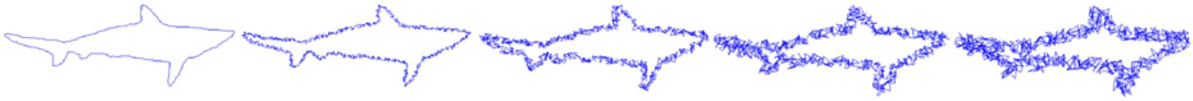

Figure 2: Fish *kk4* and its noise contaminated versions (noise ranges of $[-3,3]$, $[-6,6]$, $[-9,9]$ and $[-12,12]$, respectively, from column two to five).

are those corresponding to the contour without being smoothed enough, this leads to wrong classifications results. To solve this problem, the shape smoothing process of the proposed method can be started with a higher σ value, generating a new Full CSS matrix, which then will be used to perform the comparison tests.

Having computed all the correlation coefficients between the two curvature matrices, one for each rotation, the highest coefficient among all is chosen to be the similarity measure between the two shapes. The higher this coefficient is, the more similar are the shapes.

This procedure has the advantage of keeping the information related to the curvatures of the shapes, allowing for a more precise matching between them.

4 EXPERIMENTAL RESULTS

In this section, a comparison between the classical CSS method and the proposed Full CSS method in terms of retrieval results is done.

For sake of comparison, tests were made with two initial σ values ($\sigma = 1$ and $\sigma = 10$) when computing the curvature matrices at the classification stage.

To test the retrieval performance of the classical CSS and the proposed FCSS method, a Matlab®-based indexing and retrieval framework was implemented on Microsoft® Windows® XP Professional running on a PC Athlon64 platform.

For the tests, a small database set, comprising the first one hundred fish contours (from *kk1* to *kk100*) from the fish contours database available at (Abbasi et al., 2005), was selected. The classical CSS method was implemented using the same matching algorithm described in (Mokhtarian and Abbasi, 2002).

Classification accuracy was measured by the Correct Classification Rate (CCR), as in Equation 5:

$$CF = \frac{c}{t} \times 100\% \tag{5}$$

where c is the number of *correctly* classified contours, and t is the number of classified contours.

Each contour was compared to itself and to all others belonging to the test database. To test the robustness of the proposed methods under noisy conditions, uniform random noise in the ranges $[-3,3]$, $[-6,6]$, $[-9,9]$ and $[-12,12]$ was added to each contour. Again, each contour was compared to all noise contaminated contours of the test database. Figure 2 shows one of the contours (and its noise contaminated versions) used to perform the comparison tests between the methods discussed in this work.

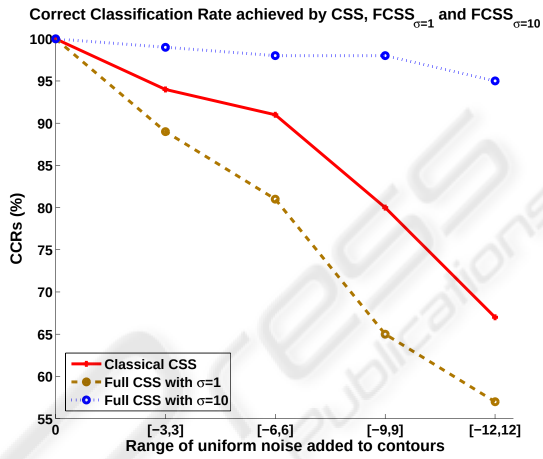

Figure 3: CCRs achieved by CSS, FCSS$_{\sigma=1}$ and FCSS$_{\sigma=10}$.

Table 1 and Figure 3 both summarize the test results for the classical CSS method along with the results for FCSS$_{\sigma=1}$ and FCSS$_{\sigma=10}$, presenting the CCRs achieved by each method *versus* the range of uniform noise added to the shapes.

Figure 4 presents the retrieval results achieved by FCSS with $\sigma = 10$ in a shape similarity query using the fish contour *kk2*. This test was done to visually assess the accuracy of the proposed method: the sys-

Table 1: CCRs achieved by the classical CSS method, FCSS$_{\sigma=1}$ and FCSS$_{\sigma=10}$ for the test shapes.

Amplitude of the random uniform noise	CCR (%)		
	Classical CSS method	FCSS $\sigma=1$	FCSS $\sigma=10$
0 (no noise)	100	100	100
$[-3,3]$	94	89	99
$[-6,6]$	91	81	98
$[-9,9]$	80	65	98
$[-12,12]$	67	57	95

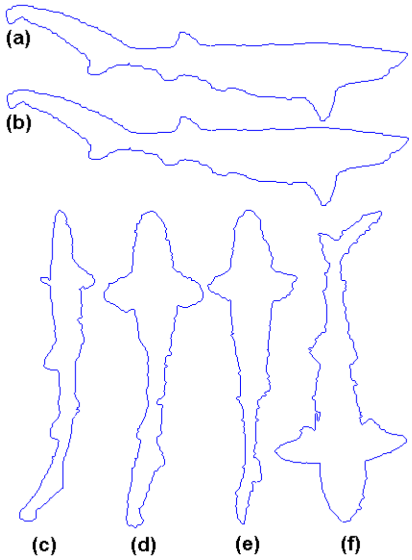

Figure 4: Similarity retrieval test results, ordered from the most similar to the less similar according to the system answer: (a) shape used as query (*kk2*); (b) *kk2*; (c) *kk13*; (d) *kk31*; (e) *kk22*; and (f) *kk20*.

tem was set to return the five more similar shapes to the one chosen. The five fishes returned were *kk2* itself, *kk13*, *kk31*, *kk22* and *kk20*, ordered by similarity according to the system answer.

The time spent by the proposed method to retrieve one shape from the database was about 19 seconds. This means that, given one shape, the system spent 19 seconds to retrieve 100 other shapes and return the correct (or incorrect) shape matching the given one. The classical CSS method spent 15 seconds to perform the same retrieval task.

The memory needed for each Full CSS matrix is the same as for an image of size about 300×200. The Matlab® implementation used in this work stored the FCSS data as an array of 16 bit-unsigned integers, meaning that one FCSS matrix has about 480 K bytes. This memory amount can be reduced by using an array of 8 bit-unsigned integers.

5 CONCLUSIONS

This paper proposed a new, simple and highly accurate shape classification scheme based on the well known CSS method. The main innovations of the new representation scheme over the classical CSS method are the use of all curvature information available, of different initial σ values to improve the retrieval performance and of the correlation factor as a similarity measure.

It is worth to say that in (Junior and da Costa, 1998), a similar multiscale, curvature-based shape representation method was shown, but the author only describes his technique and does not show any results regarding its shape retrieval performance. Besides that, his method uses Fourier Transforms and is not directly related to the CSS method.

To assess the performance of the classical CSS and the FCSS method, the test shapes were contaminated with random uniform noise ranging from low to high amplitudes and a database shape retrieval was done, as described in Section 4.

Evaluating the effects of adding noise to the test shapes over classification accuracy, as in Figure 3, showed that the performance of the classical CSS method quickly degrades. This is due to the fact that when noise range goes beyond $[-6, 6]$, important regions of the contour will overlap (as can be seen in Figure 2). In contrast, when using $\sigma = 10$ the proposed classifier maintains its performance even in presence of severe noise, as can be seen in Figure 3.

Table 1 and Figure 3 together show that the FCSS method with $\sigma = 10$ achieved CCR = 95% with noise amplitudes up to $[-12, 12]$, while the classical CSS method achieved only CCR = 67% under the same condition.

The robustness of the proposed method under noisy conditions should be pointed: a retrieval for similar shapes showed very good results, as can be seen from Figure 4.

Table 1 shows that simply increasing the σ initial value also increased the classifier performance: for a noise with amplitude range $[-12, 12]$, the classical CSS method achieved CCR = 67% while the FCSS method with $\sigma = 10$ achieved CCR = 95% (a much better result than the FCSS method with $\sigma = 1$, which only achieved CCR = 57%).

The time spent in a simple query by the proposed method is still high. This is due to the fact that all the test platform was implemented using interpreted code in Matlab® without any speed optimization. The authors are now implementing the same test framework in Java, seeking for a more efficient retrieval.

It is clear that the proposed method consistently outperforms the classical CSS method. This superiority is still more remarkable when shapes are severely contaminated by noise and shows that the proposed method is suitable to be used in such conditions.

Other research directions include more tests to assess the classification and retrieval results under different types of noise; investigate the performance of the FCSS method in other domains, such as tumor classification and recognition (de Almeida et al., 2005), and in shapes with large boundary indentations and protrusions; and investigate the use of dictionary created by lossless data compression algorithms, such as LZW or PPM, instead of the correlation coefficient as the similarity measure (Batista and Meira, 2004).

ACKNOWLEDGMENTS

The authors wish to thank Professor Fernando Manuel Bernardo Pereira, from Instituto Superior Técnico, Lisboa, for his valuable contribution to our Image Processing Group, and Professor Farzin Mokhtarian for his fish contour database.

REFERENCES

Abbasi, S., Mokhtarian, F., and Kittler, J. (2005). Search for similar shapes in the squid system: Shape queries using image databases. World Wide Web, `http://www.ee.surrey.ac.uk/Research/VSSP/imagedb/demo.html`.

Abbasi, S., Mokhtarian, F., and Kittler, J. V. (2000). Enhancing CSS-based shape retrieval for objects with shallow concavities. *Image and Vision Computing*, 18(3):199–211.

Batista, L. V. and Meira, M. M. (2004). Texture classification using the Lempel-Ziv-Welch algorithm. In *Proceedings of the 17th Brazilian Symposium on Artificial Intelligence*, pages 444–453, São Luís, Maranhão, Brazil.

Daoudi, M. and Matusiak, S. (2000). Visual image retrieval by multiscale description of user sketches. *Journal of Visual Languages and Computing*, 11(3):287–301.

Davies, E. R. (1997). *Machine Vision: Theory, Algorithms, Practicalities*. Academic Press, 1st edition.

de Almeida, C. W. D., v. d. Poel, J., Batista, L. V., and de Amorim, H. L. E. (2005). Análise de Formas Baseada no Método da Curvature Scale Space para Tumores de Câncer de Mama. In *Proceedings of the Brazilian Computer Society – SBC – WIM 2005*, pages 1–4. Brazilian Computer Society – SBC.

Dudek, G. and Tsotsos, J. K. (1997). Shape representation and recognition from multiscale curvature. *Computer Vision and Image Understanding*, 68(2):170–189.

Freeman, H. and Saaghri, A. (1978). Generalized chain codes for planar curves. In *Proceedings of the 4th International Joint Conference on Pattern Recognition*, pages 701–703.

Haralik, R. M. and Shapiro, L. G. (1992). *Computer and Robot Vision. Vol. I.* Addison-Wesley, Reading, MA.

Junior, R. M. C. and da Costa, L. F. (1998). Towards effective planar shape representation with multiscale digital curvature analysis based on signal processing techniques. *Pattern Recognition*, 29(9):1559–1569.

Kauppinen, H., Seppanen, T., and Pietikainen, M. (1995). An experimental comparison of autoregressive and Fourier-based descriptors in 2d shape classification. *IEEE Trans. Pattern Anal. Mach. Intell.*, 17(2):201–207.

Loncaric, S. (1998). A survey of shape analysis techniques. *Pattern Recognition*, 31(8):983–1001.

Mokhtarian, F. (1995). Silhouette-based isolated object recognition through Curvature Scale Space. *IEEE Trans. Pattern. Anal. Mach. Intell.*, 17(5):539–544.

Mokhtarian, F. and Abbasi, S. (2002). Shape similarity retrieval under affine transforms. *Pattern Recognition*, 35(1):31–41.

Mokhtarian, F., Abbasi, S., and Kittler, J. (1996a). Robust and efficient shape indexing through Curvature Scale Space. In *Proceedings of British Machine Vision Conference*, pages 53–62.

Mokhtarian, F., Abbasi, S., and Kittler, J. V. (1996b). Efficient and robust retrieval by shape content through Curvature Scale Space.

Mokhtarian, F. and Bober, M. (1998). *Curvature Scale Space Representation: Theory, Applications and MPEG-7 Standardisation*. Springer, 1st edition.

Mokhtarian, F. and Mackworth, A. K. (1986). Scale based description and recognition of planar curves and two-dimensional shapes. *IEEE Trans. Pattern. Anal. Mach. Intell.*, 8(1):34–43.

Mokhtarian, F. and Mackworth, A. K. (1992). A theory of multiscale, curvature-based shape representation for planar curves. *IEEE Trans. Pattern. Anal. Mach. Intell.*, 14(8):789–805.

Niblack, W., Barber, R., Equitz, W., Flickner, M., Glasman, E., Petkovic, D., Yanker, P., Faloutsos, C., and Taubin, G. (1993). The QBIC project: querying images by content using color, texture and shape. In *Proceedings of the SPIE - The International Society for Optical Engineering*, volume 1908, pages 173–187.

Pavlidis, T. (1980). Algorithms for shape analysis of contours and waveforms. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2(4):301–312.

Persoon, E. and Fu, K. S. (1977). Shape discrimination using Fourier descriptors. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-7(3):170–179.

Pomerantz, J. R., Sager, L. C., and Stoever, R. J. (1977). Perception of wholes and of their component parts: some configural superiority effects. *Journal of experimental psychology*, 3(3):422–435.

Tieng, Q. and Boles, W. W. (1997). Recognition of 2d object contours using the wavelet transform zero-crossing representation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(8):910–916.

Yang, H. S., Lee, S. U., and Lee, K. M. (1998). Recognition of 2d object contours using starting-point-independent wavelet coefficient matching. *Journal of Visual Communication and Image Representation*, 9:171–181.

Zahn, C. and Roskies, R. (1972). Fourier descriptors for plane closed curves. *IEEE Transactions on Computers*, C–21:269–281.

Zhang, D. and Lu, G. (2001). Content-based shape retrieval using different shape descriptors: A comparative study. In *Proc. of IEEE International Conference on Multimedia and Expo (ICME2001)*, pages 317–320, Tokyo, Japan.