System for Screening Objectionable Images Using Daubechies' Wavelets and Color Histograms ⋆

James Ze Wang Gio Wiederhold Oscar Firschein

Department of Computer Science, Stanford University, Stanford, CA 94305

Abstract. This paper describes WIPE$_{TM}$ (Wavelet Image Pornography Elimination), an algorithm capable of classifying an image as objectionable or benign. The algorithm uses a combination of Daubechies' wavelets, normalized central moments, and color histograms to provide semantically-meaningful feature vector matching so that comparisons between the query image and images in a pre-marked training set can be performed efficiently and effectively. The system is practical for real-world applications, processing queries at the speed of less than 10 seconds each, including the time to compute the feature vector for the query. Besides its exceptional speed, it has demonstrated 97.5% recall over a test set of 437 images found from objectionable news groups. It wrongly classified 18.4% of a set of 10,809 benign images obtained from various sources. For different application needs, the algorithm can be adjusted to show 95.2% recall while wrongly classifying only 10.7% of the benign images.

I Introduction

Every day, large numbers of adults and children use the internet for searching and browsing through different multimedia documents and databases. Convenience in accessing a wide range of information is making the internet and the world-wide web part of the everyday life of ordinary people. To protect the freedom of speech, people are allowed to publish various types of material or conduct different types of business on the internet. However, due to this policy, there is currently a large amount of domestic and foreign objectionable images and video sequences available for free download on the world-wide web and usenet newsgroups. Accessing objectionable media by under-aged "netters" is increasingly a problem that many parents are concerned about.

I.1 Related Work in Industry

There are many attempts to solve the problem of objectionable images in the software industry. Pornography-free web sites such as the *Yahoo! Web Guides for Kids* have been set up for protecting those children too young to know how to use the web browser to get to other sites. However, there is still a possibility that children may be able to gain access to an objectionable site starting from such a site.

Programs such as *NetNanny* or *CyberSitter* have been put on market for parents to prevent their children from accessing objectionable documents. However, the algorithms used in this software are still in an elementary stage. Some software stores more than 10,000 IP addresses and blocks access to objectionable

⋆ *Correspondence to:* wangz@cs.stanford.edu

sites by matching the site addresses. Other software blocks all unsupervised image access. Apparently, there are problems with this software. The internet is so dynamic that more and more new sites and personal pages are added to it everyday. Memorizing many sites is neither efficient in speed nor feasible. Eliminating all images is not a feasible solution since the internet will not be as useful and attractive to children if we do not allow children to view images.

I.2 Related Work in Academia

Academic researchers are also actively investigating alternative algorithms to screen and block objectionable media. Many recent developments in shape detection, object representation and recognition, people recognition, face recognition, and content-based image and video database retrieval are being considered by researchers for use in this problem. A detailed summary of previous work can be found in [8].

To make such algorithms practical for our purposes, extremely high recall with reasonably high speed and precision are necessary. In this application, *recall* is defined as the ratio of the number of objectionable images identified, to the total number of objectionable images in the database; *precision* is defined for a set of images labeled as objectionable as the ratio of images that are actually objectionable to the number of images in the set. A perfect system would identify all objectionable images and not mislabel any benign images, and would therefore have a recall and precision of 1. In the present application, a high recall is desirable, i.e., the correct identification of almost every image likely to be objectionable even though this may result in some benign images being mislabeled. As one can imagine, even a few exposures of objectionable images would be harmful enough for children.

The following properties of the internet objectionable images make the problem extremely difficult:

- mostly contain non-uniform image background;
- foreground may contain textual noise such as phone numbers, URLs, etc;
- may range from grey-scale to 24-bit color;
- may be of very low quality (sharpness);
- taken by all possible camera positions.
- may be an indexing image containing many small icons;
- may contain more than one person;
- persons in the picture may be of different skin colors;
- may contain both people and animals;
- may contain only some parts of a person;
- person in the picture may be partially dressed;
- person in the picture may be fully dressed, but the facial expression is objectionable;
- and so on.

Forsyth's research group [7, 8, 9] has designed and implemented an algorithm to screen images of naked people. Their algorithms involve a skin filter and a human figure grouper. As indicated in [8], 43% recall and 57% precision have been obtained for a test set of 565 images with naked people and 4289 assorted benign images. However, it takes about 6 minutes on a workstation for the grouper in their algorithm to process a suspect image passed by the skin filter.

I.3 Overview of Our Work

Our approach is different from previous approaches. Instead of carrying out a detailed analysis of an image, we match it against a small number of feature vectors obtained from a training database of 500 objectionable images and 8,000 benign images. If the image is close in content to a threshold number of pornographic images, e.g., matching two or more of the marked objectionable images in the training database within the closest 15 matches, it is considered objectionable. To accomplish this, we attempt to effectively code images based on image content and match the query with statistical information on the feature indexes of the training database. The foundation of this approach is the content-based feature vector indexing and matching developed in our multimedia database research.

Image feature vector indexing has been developed and implemented in several multimedia database systems such as the IBM QBIC System [5, 16] developed at the IBM Almaden Research Center. Readers are referred to [10, 13, 17, 18, 21] for details on this subject.

In the WIPE project, we developed a new algorithm to efficiently and accurately compare the semantic content of images mainly consisting of objects such as the human body. Using Daubechies' wavelets, moment analysis, and histogram indexing, the algorithm produces feature vectors that provide excellent accuracy in matching images of relatively isolated objects such as the human body. We use a novel multi-step metric to compute the distance between two given images. A training database of about 500 objectionable images and about 8,000 benign images has been indexed using such an algorithm. When a query comes in, we compute the feature vector and use it to match with the training database. If it matches with objectionable images in the training database, we classify it as an objectionable image. If it does not match with objectionable images in the training database, we classify it as a benign image. Promising results have been obtained in experiments using a test set of 437 objectionable images and 10,809 benign images.

II Related Background

II.1 Content Based Image Indexing and Retrieval

There are several ways to index the images so that queries can be performed by comparing the indexes. The color histogram is one of the many ways to index color images and it preserves the color information contained in images very well. However, a global histogram does not preserve the color locational information within the images. Using this measure, two images may be considered to be very similar to each other even though they have completely unrelated semantics. Shape and texture-based detection and coding algorithms are other techniques for indexing images. They both have substantial limitations for general-purpose image databases. For example, current shape detection algorithms such as the *snake* algorithm only work effectively on images with relatively uniform backgrounds. Texture coding is not appropriate for non-textural images.

Storing color layout information is another way to describe the image content. A wavelet and statistical analysis-based layout indexing technique developed at the Stanford WBIIS project [21] performs best when the images contain a lot of high frequency information such as sharp color changes.

II.2 Daubechies' Wavelets and Fast Wavelet Transform

Wavelets, developed in mathematics, quantum physics, and statistics, are functions that decompose signals into different frequency components and analyze each component with a resolution matching its scale. Applications of wavelets to signal denoising, image compression, image smoothing, fractal analysis and turbulence characterization are active research topics.

Theoretical details on wavelet analysis, wavelet basis and Daubechies' wavelets can be found in [6, 15, 4, 14, 21]. Daubechies' wavelets give remarkable results in image analysis and synthesis due to its mathematical properties.

Daubechies' wavelets transform is more like a weighted averaging which better preserves the trend information stored in the signals if we consider only the low-pass filter part. Various experiments and studies have shown that Daubechies' wavelets are better than other wavelet forms for dealing with general-purpose images.

Original image — *Traditional edge detection* (with post-processing) — *Wavelet edge detection* (without post-processing)

Fig. 1. Comparison of the traditional edge detection and our wavelet algorithm. The amount of background noise can be controlled by using different Daubechies' wavelets.

In shape-based image indexing, we want to represent the object shape in the image as exactly as possible in the coefficients of the feature vector. When using the Haar wavelet, we get too much noise in the high-pass bands. Traditional edge detection algorithms have the same problem, as illustrated in Figure 1. Daubechies' wavelets offer (1) a multiresolution analysis, which has the potential for high speed algorithm design, and (2) a wide range of flexibility. For example, we may select the appropriate wavelet basis to obtain the exact amount of fluctuation we desire in the high-frequency bands to represent the object shape.

II.3 Moments

Moments are descriptors widely used in shape and regional coding [11]. For a 2-D continuous surface $f(x, y)$ embedded on the xy-plane, the *moment of order* $(p+q)$ is defined as

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy \tag{1}$$

for $p, q \in \mathbb{N} \bigcup \{0\}$. Theory of moments has shown that the moment sequence $\{m_{pq}\}$ is uniquely determined by $f(x, y)$ and vice versa.

The *central moment* is defined as

$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(x - \frac{m_{10}}{m_{00}}\right)^p \left(y - \frac{m_{01}}{m_{00}}\right)^q f(x, y) dx dy. \tag{2}$$

For discrete cases such as a digitized image, we define the *central moment* as

$$\mu_{pq} = \sum_x \sum_y \left(x - \frac{m_{10}}{m_{00}} \right)^p \left(y - \frac{m_{01}}{m_{00}} \right)^q f(x,y). \tag{3}$$

Then the *normalized central moments* are defined as

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}} \quad where \quad \gamma = \frac{p+q+2}{2} \tag{4}$$

for $p + q = 2, 3, 4, ...$

A set of seven *translation, rotation, and scale invariant moments* can be derived from the 2nd and 3rd moments. A detailed introduction to these moments can be found in [12, 11]. These moments can be used to match two objectionable images containing people having the same posture but taken from different camera angles.

III Screening Algorithm in WIPE

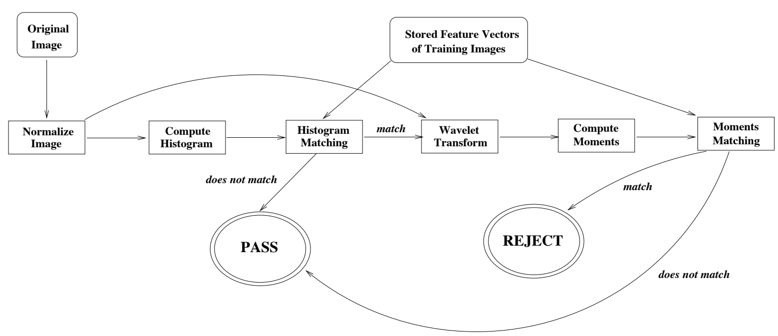


Fig. 2. Basic structure of the algorithm.

III.1 Overview

We have developed a new shape-based indexing scheme using forward and backward Daubechies' wavelet transforms, variant and invariant normalized central moments, and color histogram analysis that is able to capture the object posture. The screening algorithm uses several major steps. Figure 2 shows the basic structure.

We apply a 1-level fast wavelet transform (FWT) with Daubechies-3 wavelet to each image in the training set of objectionable and benign images. Then we perform an inverse wavelet transform on each of the three high frequency blocks obtained from the wavelet transform using the same wavelet basis. Multidirectional edge detection is then applied to the three inverse transforms to obtain fluctuation information stored in the original image. The normalized central moments of the edge image, invariant moments of the edge image and the color histogram of the rescaled (R, G, B)-space image are stored as a feature vector.

Given a query image, the search is carried out in three steps after the feature vector for the query is computed. In the first step, the color histogram of the query is used. If the amount of yellow in the query image is lower than a threshold, we classify it as a benign image since we have found that objectionable images contain large areas of yellow or colors close to yellow. In the second step, a crude selection based on the individual moments stored is carried out. Then a weighted version of the Euclidean distance between the moments of an image in the training set selected in the first step and those of the querying image is calculated, and the images with the smallest distances are selected and sorted. In the third step, a color histogram is again used to perform the final match so that images with both human body shape and skin color histogram will be selected as matching images to the query. If one or more objectionable images is found in the best matching images, we mark the query as an objectionable image. Otherwise, the query is marked as a benign image.

Our design has several immediate advantages.

1. It does not rely too much on color when detecting sharp edges. That means that naked people of different races can be detected without bias. It also has the potential for shape-based matching of benign images. Image background does not affect the querying results unless the background is not reasonably smooth. Also, the query image can be of different color quality.
2. We used Daubechies' wavelet rather than a traditional edge detector to capture the shape information in the images. This reduced the dependence on the quality or the sharpness of the images.
3. We used a combination of variant and invariant normalized central moments to make the querying independent of the camera position.
4. Our algorithm is expandable for different types of objectionable images and different needs. Queries in our algorithm are based on feature indexes of an expandable training set of images in the database. By including images of various types, it is likely to improve the recall and the precision of the algorithm. Besides, this algorithm can be easily applied to other shape-based searching such as object matching.

III.2 Normalize the Images

Many color image formats are currently in use, e.g., GIF, JPEG, PPM and TIFF are the most widely used formats. Because images in an image database can have different formats and different sizes, we must first normalize the data for histogram computation. For the wavelet computation and moment analysis parts of our algorithm, any image size is acceptable.

A rescaled thumbnail consisting of 128×128 pixels in Red-Green-Blue (i.e., RGB) color space is adequate for the purpose of computing the histogram feature vectors.

III.3 Color Histogram Indexing and Matching

It is clear to us that objectionable images contain mostly large areas of yellow or colors close to yellow. This is not the case for benign images.

The color histogram indexing and searching techniques used in the WIPE system are fairly standard when compared to systems such as the IBM QBIC System. We use a total of 512 bins to compute the histogram.

A much more efficient approach is implemented. We manually define a certain color range in the color spectrum as yellow. If we define

$$yellow(r, g, b) : [0, 255] \times [0, 255] \times [0, 255] \rightarrow [0, 1] \tag{5}$$

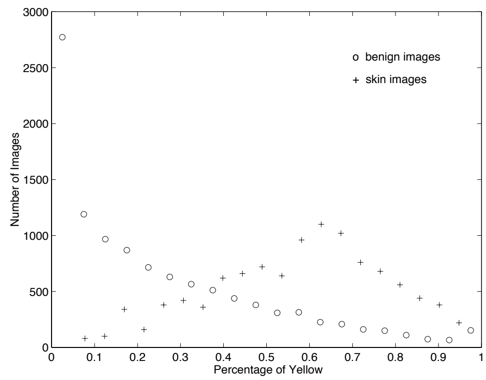


Fig. 3. Percentage of yellow in benign images and skin images.

as the weight of yellow for a certain color (r, g, b), then $yellow(r, g, b) = 1$ means the color is the most yellow color and $yellow(r, g, b) = 0$ means the color is the least yellow color. Denote

$$histogram(r, g, b) : [0, 255] \times [0, 255] \times [0, 255] \rightarrow \mathbb{R}+ \tag{6}$$

as the histogram of a given image. Then the function

$$YLW(image) = \sum_{r=0}^{255} \sum_{g=0}^{255} \sum_{b=0}^{255} \Big(yellow(r, g, b) \times histogram(r, g, b) \Big) \tag{7}$$

indicates the weighted amount of yellow that the given image contains.

For simplicity, we set

$$yellow(r, g, b) = \begin{cases} 1 & (r, g, b) \ close\ to\ yellow \\ 0 & otherwise \end{cases} \tag{8}$$

Figure 3 shows histograms of the values of $YLW(image)$ for benign images and skin images. If we set a threshold of, say, 0.15, about one half of the benign images are then classified correctly, while only a small amount of skin images are classified incorrectly.

Using this approach, the run time of the histogram matching step for skin images has been reduced from $O(n)$ to $O(1)$, where n is the number of images in the training database.

III.4 Edge and Shape Detection and Matching Using Wavelets and Moments

Clearly the color histogram approach alone is not sufficient. Sometimes two images may be considered very close to each other using this measure when in actuality they have completely unrelated semantics.

We apply the wavelet transform to perform multidirectional and multiscale edge detection. Readers are referred to [1] for the theoretical arguments on the effectiveness of a similar algorithm. Our purpose is not to obtain a high quality edge detection algorithm for this application. Rather, since the goal here is to effectively extract the main shape information for objects and textural information for areas from the image, it is not necessary to produce a perceptually pleasant edge image. Consequently, we try to keep the algorithm simple to achieve a fast computation speed.

We start the edge detection process by transforming the image using the Daubechies-3 wavelet basis. The image is decomposed into four frequency bands with corresponding names LL, HL, LH and HH. The notation is borrowed from the filtering literature [20]. The letter 'L' stands for low frequency and the letter 'H' stands for high frequency. The left upper band is called 'LL' band because it contains low frequency information in both the row and column directions. The details of the filtering terminologies are given in [20]. An even number of columns and rows in the querying image is required due to the downsampling process of the wavelet transform. However, if the dimensions of the image are odd, we simply delete one column or one row of pixels from the boundaries.

The LH frequency band is sensitive to the horizontal edges, the HL band is sensitive to the vertical edges, and the HH band is sensitive to the diagonal edges [4, 1]. This property has been used in a textual information detection and elimination system designed for secure medical image distribution [22].

We detect the three types of edges separately and combine them at the end to construct a complete edge image. To detect the horizontal edges, we perform an inverse Daubechies-3 wavelet transform on a matrix containing only the wavelet coefficients in the LH band. Then we apply a zero-crossing detector in vertical direction to find the edges in the horizontal direction. The mechanism for using zero-crossing detector to find the edges can be found in [1]. Similar operations are applied to the HL and HH band. The difference lies with the zero-crossing detector. For the HL band, we use a zero-crossing detector in the horizontal direction to find vertical edges and for the HH band, we use zero-crossing detector in the diagonal direction to find diagonal edges.

After we get the three edge maps, we combine them to get the final edge image. To numerically show the combination, let us denote[2] the three edge maps by $E_1[1:m, 1:n]$, $E_2[1:m, 1:n]$ and $E_3[1:m, 1:n]$. The image size is $m \times n$. Then the final edge image, denoted by $E[1:m, 1:n]$, can be obtained from

$$E[i,j] = \left(E_1[i,j]^2 + E_2[i,j]^2 + E_3[i,j]^2\right)^{\frac{1}{2}}. \qquad (9)$$

Once the edge image is computed, we compute the normalized central moments up to order five and the translation, rotation, and scale invariant moments based on the gray scale edge image using the definitions in Section II.3. A feature vector containing these $21 + 7 = 28$ moments is computed and stored for each image in the training database. When a query comes in that has passed the histogram matching step, a moment feature vector is computed and a weighted Euclidean distance is used to measure the distance of the query and an image in the training database. The weights are determined so that matching of the 21 normalized central moments has higher priority than the matching of the 7 invariant moments. In fact, many objectionable images are of similar orientation.

If the query matches with objectionable images in the training database, we classify it as an objectionable image. If it does not match with objectionable images in the training database, we classify it as a benign image.

---

[2] Here we use MATLAB notation. That is, $A(m_1 : n_1, m_2 : n_2)$ denotes the submatrix with opposite corners $A(m_1, m_2)$ and $A(n_1, n_2)$.

IV Experimental Results

Type of Images (total)	Eliminated (percentage)	Passed WIPE (percentage)	Eliminated (percentage)	Passed WIPE (percentage)
Pornographic (437)	426 (97.5%)	11 (2.5%)	416 (95.2%)	21 (4.8%)
Benign (10809)	1993 (18.4%)	8816 (81.6%)	1155 (10.7%)	9654 (89.3%)
	Configuration A		*Configuration B*	

Table 1. Overall Classification Performance of the WIPE System.

This algorithm has been implemented on a Sparc-20 workstation. We selected about 500 objectionable images from news groups and 8,000 benign images from various sources such as the Corel Photo CD-ROM series for our training database. When we downloaded the objectionable images, we tried to eliminate those from the same source, i.e., those of extremely similar content. To compute the feature vectors for the 8,000 color images of size 640×480 in our database requires approximately 3 hours of CPU time.

We have also selected 437 objectionable images and 10,809 benign images as our queries. The matching speed is very fast. Using a SUN Sparc-20 workstation, it takes less than 10 seconds to process a query and select the best 100 matching images from the 8,500 image database using our similarity measure. Once the matching is done, it takes almost no extra CPU time to determine the final answer, i.e., if the query is objectionable or benign.

Besides the fast speed, the algorithm has achieved remarkable accuracy. Table 1 shows the overall classification performance of two system configurations of the WIPE system. In System Configuration A, we set the yellow color threshold to 10% in the color matching and the the minimum number of objectionable images within the top 15 best matching images for the shape matching stage to 1. That is, the query passes the WIPE system only if it contains less than 10% predefined yellow or else it does not match any objectionable image within the top 15 best matching images. Otherwise, it is rejected by the WIPE system. In System Configuration B, we set the two thresholds to 10% and 3. This allows many more benign images to pass the WIPE system correctly, but somewhat more objectionable images are mistakenly allowed. Choosing a more constraining filter, i.e., more protective for children, will cause a greater loss of valid images, but that trade-off may still be worthwhile for some parents. The algorithm can be set up so that the parent user can set the level of protection.

Figure 5 and 6 show typical images being mistakenly marked by the WIPE system. Most of the few failures for marking objectionable images happen when the query contains textual noise and/or a frame, as in Figure 6.a.

V Conclusions and Future Work

In this paper, we have demonstrated an efficient shape-based indexing and matching system using Daubechies' wavelets developed by us for screening objectionable images.

It is possible to improve the search accuracy by fine-tuning the algorithm, e.g., using neural network, using a perceptually-comparable color space, adjusting weights for different matching steps, or adding more complicated preprocessing to eliminate textual noise or frames. The accuracy may also be improved by fine-tuning the training database.

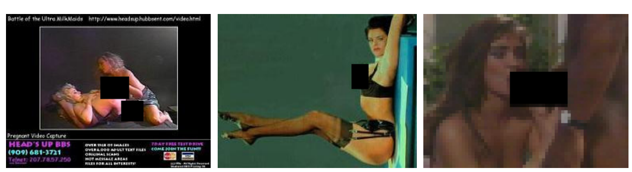
Fig. 4. Animal images are sometimes mistaken for objectionable images. Some areas of objectionable images are blackened and blurred.

(a) *(b)* *(c)*

Fig. 5. Typical benign images being marked mistakenly as objectionable images by WIPE. (a) too much yellow (b) hard to tell (bathing) (c) partially undressed.

It is also possible to make the searching faster by developing a better algorithm for storing and matching the feature vectors. Significant speed-up is also possible if a more extensive statistical analysis can be utilized. The algorithm can also be modified to execute in parallel on multi-processor systems. Experiments with our algorithm on a video database system could be another interesting study.

Finally, we are working on applying this technique to shape-based search of regular image databases containing, for example, merchandise catalog images and museum art images. It is also possible to extend this technique to texture-based search.



(a) *(b)* *(c)*

Fig. 6. Typical objectionable images being marked mistakenly as benign images by WIPE. (a) frame and textual noise (b) dressed but objectionable (c) image too dark and of extremely low contrast. Some areas of objectionable images are blackened and blurred.

VI Acknowledgements

We would like to thank Jia Li of Stanford University Electrical Engineering Department for useful discussions and valuable help in implementing the algorithm. We would also like to thank the Stanford University Libraries and Academic Information Resources (SULAIR) for providing computer equipment during the development and testing process.

References

1. Turgut Aydin et al, Multidirectional and Multiscale Edge Detection via M-Band Wavelet Transform, *IEEE Transactions on Image Processing*, Vol. 5, No. 9, pp. 1370-1377, 1996.
2. Dana H. Ballard, Christopher M. Brown, *Computer Vision*, Prentice-Hall, Inc., New Jersey, 1982.
3. Ingrid Daubechies, Orthonormal bases of compactly supported wavelets, *Communications on Pure and Applied Mathematics*, 41(7):909-996, October 1988.
4. Ingrid Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, 1992.
5. C. Faloutsos et al, Efficient and Effective Querying by Image Content, *J. of Intelligent Information Systems*, 3:231-262, 1994.
6. G. B. Folland, *Fourier Analysis and Its Applications*, Pacific Grove, Calif., 1992.
7. Margaret Fleck, David A. Forsyth, Chris Bregler, Finding Naked People, *Proc. 4'th European Conf on Computer Vision*, 1996.
8. David A. Forsyth and Margaret Fleck, Finding Naked People, *journal reviewing*, 1996.
9. David A. Forsyth et al, Finding Pictures of Objects in Large Collections of Images, *Proceedings, International Workshop on Object Recognition*, Cambridge, 1996.
10. Amarnath Gupta and Ramesh Jain, *Visual Information Retrieval*, Communications of the ACM, vol.40 no.5, pp 69-79, 1997.
11. Rafael C. Gonzalez, Richard E. Woods, *Digital Image Processing*, Addison-Wesley Publishing Co., 1993.
12. M. K. Hu, Visual Pattern Recognition by Moment Invariants, *IRE (IEEE) Trans. Info. Theory*, Vol. IT-8, pp. 179-187.
13. C. E. Jacobs, A. Finkelstein, D. H. Salesin, Fast Multiresolution Image Querying, *Proceedings of SIGGAPH 95, in Computer Graphics Proceedings, Annual Conference Series*, pp.277-286, August 1995.
14. Gerald Kaiser, *A Friendly Guide to Wavelets*, Birkhauser, Boston, 1994.
15. Yves Meyer, *Wavelets: Algorithms & Applications*, SIAM, Philadelphia, 1993.
16. W. Niblack et al, The QBIC project: Query image by content using color, texture and shape, *Storage and Retrieval for Image and Video Databases*, pages 173-187, San Jose, 1993. SPIE.
17. R. W. Picard, T. Kabir, Finding Similar Patterns in Large Image Databases, *IEEE ICASSP*, Minneapolis, Vol, V., pp.161-164, 1993.
18. A. Pentland, R. W. Picard, S. Sclaroff, Photobook: Content-Based Manipulation of Image Databases, *SPIE Storage and Retrieval Image and Video Databases II*, San Jose, 1995.
19. M.J. Swain and D. H. Ballard, Color Indexing, *Int. Journal of Computer Vision*, 7(1):11-32, 1991.
20. Martin Vetterli, *Wavelets and Subband Coding*, Prentice Hall, N.J., 1995.
21. James Ze Wang, Gio Wiederhold, Oscar Firschein, Sha Xin Wei, Wavelet-Based Image Indexing Techniques with Partial Sketch Retrieval Capability, *Proceedings of the Fourth Forum on Research and Technology Advances in Digital Libraries (ADL'97)*, Washington D.C., May 1997.
22. James Ze Wang, Michel Bilello, Gio Wiederhold, Textual Information Detection and Elimination System for Secure Medical Image Distribution, *submitted for conference publication*, March 1997.
23. Gio Wiederhold, Digital Libraries, Value, and Productivity, *Com. ACM*, Vol.38 No.4, April 1995, pages 85-96.