W-transform Method for Feature-oriented Multiresolution Image Retrieval

Man Kam Kwong and Biquan Lin*

Mathematics and Computer Science Division
Argonne National Laboratory
Argonne, IL 60439-4844
E-mail: kwong@mcs.anl.gov, blin@mcs.anl.gov

Abstract

We present a local-feature-oriented image indexing and retrieval method based on Kwong and Tang's W-transform. Multiresolution histogram comparison is an effective method for content-based image indexing and retrieval. However, most recent approaches perform multiresolution analysis for whole images but do not exploit the local features present in the images. Since W-transform is featured by its ability to handle images of abitrary size, with no periodicity assumptions, it provides a natural tool for analyzing local image features and building indexing systems based on such features. In our approach, the histograms of the local features of images are used in the indexing system. The system not only can retrieve images that are similar or identical to the query images but also can retrieve images that contain features specified in the query images, even if the retrieved images as a whole might be very different from the query images. The local-feature-oriented method also provides a speed advantage over the global multiresolution histogram comparison method. The

*This work was supported by the Office of Scientific Computing, U.S. Department of Energy, under Contract W-31-109-Eng-38.

feature-oriented approach is expected to be applicable in managing large-scale image systems such as video databases and medical image databases.

1 Introduction

Image database management is important in the development of multimedia technology. Since enormous amount of digital images is likely to be generated within the next decades in order to integrate computers, television, VCR, cables, telephone and various imaging devices, it becomes an urgent task to build effective image indexing and retrieval systems so that images can be easily organized, searched, transmitted and presented.

Traditional image retrieval methods are built on indexing systems which use some keywords to describe image objects and attributes. This approach becomes ineffective in large-scale image database management. The reasons are as follows: First, it would be a tedious task to build such indexing systems since it requires enormous database entries to describe image objects and attributes and since the building processes cannot be automated. Second, it is difficult to design a scene description language (SDL) to describe the rich information contained in images; and it is almost impossible in some cases. The solution to overcome these difficulties is to build indexing and retrieval systems directly based on the content of images.

Recently, several content-based image retrieval methods have been studied. Among them, Pentland and *et al* [9] used Karhunen-Loeve (KL) or Word transform to obtain a compact description of the most significant features present in the image such as brightness, edges and textures; and then search images, find objects inside the images and compare textures of the images directly based on the KL transform coefficients. Since KL transform is considered to be a transform with the highest energy compactation, the KL transform based image indexing and retrieval method will effectively preserve most of the energy of the images. One drawback of this approach is its large amount of time consumption in the process of obtaining KL transform coefficients since KL transform is a computationally demanding transform. Lee and Dickinson [6] built a video indexing system by exploiting histograms of multiresolution decompositions of images. Histogram comparison is a method robust to image displacement, orientation and some amount of distortion. Multiresolution decomposition gives a hierarchical image structure so that one can manipulate images in different levels according to different requirements. The method combining these two important techniques is anticipated to be efficient. Their approach consists of performing multiresolution decomposition for the whole images but does not use the local image features. Since local image features are important for understanding and analyzing images. it is natural to extend the approach to design a local-feature-based image indexing and retrieval system.

The main difficulty in building a local feature oriented multiresolution analysis is the lack of suitable tools. Since most of existing wavelets such as Haar wavelet and Daubechies' wavelet [2] require that the size of images be a power of 2 and assume that periodical extension is possible beyond the boundaries of images. To satisfy the image size constraint, one needs to pad zeroes before performing multiresolution decomposition. Doing so not only introduces more storage requirements and computational burdens but also dilutes the features of the images. While it will not cause harm by assuming periodicity for the whole images, it is not reasonable to assume that every local areas are periodically extendable. These difficulties could be overcome if we had a wavelet tool which does not have size requirement or periodicity constraint.

Recently, Kwong and Tang [3] introduced W-transforms for multiresolution analysis. The W-transform is featured by its smoothness, its simple form of representation, its ability to handle arbitrary size of images and its exemption of periodicity assumptions. The last two properties make W-transform a natural tool for analyzing local image features and building indexing systems based on the local features. For those who are interested in experimenting with W-transforms by hands, the MATLAB version of the W-transform source code can be found in [4].

This paper presents a local-feature-oriented image retrieval method using Kwong and Tang's W-transforms. In this approach, the multiresolution histograms of the local feature images are used as index. The images with the specified features are retrieved by comparing the multiresolution histograms of the local features of the query images with the indexing histograms. Our numerical experiments of implementing this method indicated that the feature-oriented approach is practical and efficient.

The remainder of this paper is organized as follows: Section 2 introduces Kwong and Tang's W-tansforms for multiresolution analysis. Section 3 discusses multiresolution histogram comparison method for image analysis. Section 4 presents the local-feature-oriented image indexing and retrieval method. Section 5 gives some examples of numerical experiments. Finally, Section 6 concludes the discussion of this paper.

2 W-transforms and Multiresolution Analysis

First, we illustrate multiresolution decomposition and reconstruction of signals in the one-dimensional case. Let $f = (f_0, f_1, f_2, ..., f_{N-1})$ denote a one-dimensional signal. suppose G and H are $K \times N$ matrices and G^+ and H^+ are $N \times K$ matrices with the property that

$$G^+G + H^+H = I$$

where I denotes an $N \times N$ identity matrix. Then, the signal f can be decomposed into two components u and v via linear transforms

$$\begin{aligned} u &= Gf \\ v &= Hf \end{aligned} \tag{1}$$

and f can be reconstructed from the components u and v via

$$f = G^+u + H^+v. \tag{2}$$

For some specifically selected G and H, the components u and v are often referred to as lowpass component and highpass component. If we apply some matrices in the form similar to G and H to the lowpass component u, then we obtain lowpass component and highpass component of u. Repeating this process, we obtain a sequence of highpass components and a lowpass component. On the other hand, the original signal can be reconstructed from these components in the order reverse to the preceeding decomposition process. This is the basic idea of multiresolution decomposition and reconstruction.

There exist many matrices which can be used for such a decomposition and reconstruction. Some matrices G and H are associated with important transforms in signal processing. Here we give two examples, one is called Haar transform and the other is called Daubechies D_4 transform.

A Haar matrix pair G and H is a pair of $N/2 \times N$ matrices defined by

$$\mathbf{G} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & 1 & \ldots & 0 & 0 \\ & & & & \ldots & & \\ 0 & 0 & 0 & 0 & \ldots & 1 & 1 \end{bmatrix} \tag{3}$$

and

$$\mathbf{H} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & -1 & \ldots & 0 & 0 \\ & & & & \ldots & & \\ 0 & 0 & 0 & 0 & \ldots & 1 & -1 \end{bmatrix}. \tag{4}$$

The associated matrices G^+ and H^+ are defined by

$$\begin{aligned} G^+ &= G \\ H^+ &= -H. \end{aligned} \tag{5}$$

It is easy to verify that such G, H, G^+ and H^+ satisfy (2) so that it is possible to perform decomposition and reconstruction as (1) and (2).

The Haar transform is a transform with the simplest form. The Haar transform is also featured by its orthogonality, that is, $GH^T = HG^T = 0$. The orthogonality is important in some signal processing applications. For example, it can provides a complete representation of the original signal without redundancy and it will not introduce correlation of noise after performing transform for signals which are contaminated by independent noise.

The drawbacks of the Haar transform include its lack of smoothness and its size requirement for the original signals. The lack of smoothness means that its associated scaling function is not continuous so that a smooth signal will be mapped to a signal with sharp cusps. The size requirement means that the original signal is required to have a size of power of 2 to perform Haar transform. This size limitation makes the Haar transform not suitable for analyzing local feature images.

Daubechies [2] proposed a method to generate high-order orthogonal matrices which can be used to perform such decomposition and reconstruction. Specifically, she gave a pair of matrices with order 4, which is referred to as D_4 matrices in literature. We refer the reader to [2] and [3] for the detailed expression of such matrices. In contrast to the Haar transform, the scaling function associated with the D_4 transform is continuous. But the D_4 transform still requires that the original signal has a size of power of 2. In addition, it requires a periodic extension beyond the boundary of the original signal to obtain a complete reconstruction. Therefore, it is our opinion that the D_4 transform is not a good choice for analyzing local feature images either.

Now, we turn to the W-transform. Here we summarize the the results obtained in [3]. According to [3], an arbitrary vector $g = [g_1, g_2, g_3, g_4]$ and two constants c and d are first chosen. Then, use

$$h = [h_1, h_2, h_3, h_4] = [g_1/c, g_2/c, g_3/d, g_4/d] \tag{6}$$

together with g to construct W-matrices for signals with even or odd sizes. For a signal with an even size, the corresponding W-matrix pair G and H is

$$\mathbf{G} = \begin{pmatrix} g_1+g_2 & g_3 & g_4 & & & & & \\ & g_1 & g_2 & g_3 & g_4 & & & \\ & & & g_1 & g_2 & g_3 & g_4 & \\ & & & & & & \ddots & \\ & & & & & g_1 & g_2 & g_3+g_4 \end{pmatrix} \tag{7}$$

and

$$\mathbf{H} = \begin{pmatrix} h_1+h_2 & h_3 & h_4 & & & & & \\ & h_1 & h_2 & h_3 & h_4 & & & \\ & & & h_1 & h_2 & h_3 & h_4 & \\ & & & & & & \ddots & \\ & & & & & h_1 & h_2 & h_3+h_4 \end{pmatrix} \tag{8}$$

For a signal with an odd size, the corresponding W-matrix pair G and H is

$$\mathbf{G} = \begin{pmatrix} g_1+g_2 & g_3 & g_4 & & & & & & \\ & g_1 & g_2 & g_3 & g_4 & & & & \\ & & & g_1 & g_2 & g_3 & g_4 & & \\ & & & & & & \ddots & & \\ & & & & & g_1 & g_2 & g_3 & g_4 \\ & & & & & & & h_1 & h_2+h_3+h_4 \end{pmatrix}. \tag{9}$$

and

$$\mathbf{H} = \begin{pmatrix} h_1+h_2 & h_3 & h_4 & & & & & \\ & h_1 & h_2 & h_3 & h_4 & & & \\ & & & h_1 & h_2 & h_3 & h_4 & \\ & & & & & & \ddots & \\ & & & & & h_1 & h_2 & h_3 & h_4 \end{pmatrix}. \tag{10}$$

An interesting and useful property of such W-matrix pair is that its corresponding G^+ and H^+ can also be generated by a pair of vectors of length 4. Therefore, the original signal can be reconstructed by the decomposed local components. A specific set of such matrices given in [3] is built by $g = [-1, 3, 3, -1]$ and $h = [-1, 3, -3, 1]$.

We can see here that the W-matrices possess two interesting features among many others. First, observe that the W-matrices can be applied to signals with arbitrary sizes, even signals with odd sizes. This feature is not possessed by classic wavelet matrices such as Haar matrices and D_4 matrices given above. Second, it does not require periodic extension for the original signal. The transformed signal in some location is only related to the signals at its neighboring locations. Thus, according to the reason discussed above, it is suitable for analyzing local signals.

A two-dimensional signal (an image) can be expressed by a matrix. the 2D W-transform of a matrix is obtained by first applying the 1D W-transform to each column of the matrix and then applying the same W-transform to each row of the resulting matrix. This procedure generates four components (GG, GH, HG, HH) for a matrix. For example, the GG component is obtained by applying G to each column and each row of the matrix. For the component GG, we can also obtain its four components by applying W-transform to it. Repeating this process, we obtain a multiresolution decomposition of the original image, which can be used to reconstruct the original image by the reverse process. In the next section, we use these decompositions to analyze the original images and build a fast retrieval method for large-scale images.

3 Multiresolution Histograms for Image Storage and Retrieval

Multiresolution decomposition of images generated by a wavelet transform such as the W-transform has been applied to several image processing area. Among them, Coifman and Wickerhauser [1], Man and Tang [3] applied wavelet transform to image compression; Mallat and Hwang [8], Lin [5] restored images which are contaminated by noise by exploiting the properties of wavelet modulus and phase. This paper explores a new application of wavelet transform. We apply wavelet transform for large-scale image management and analysis.

Recently, RAID (Redundent Arrays of Inexpensive Disks) becomes popular to meet high-capacity storage requirements for multimedia technology. In RAID, a dual data set is stored to avoid unexpected data loss. For large-scale images, in our opinion, it will be more effective if the dual data set is stored as multiresolution decompositions of images instead of the exactly same images. The reasons are as follows: First, the original image can be quickly and completely reconstructed from the multiresolution decompositions. Therefore, saving the multiresolution decompositions of images is equivalent to saving the original images. Second, the multiresolution decompositions can be exploited for quick image retrieval. In the next paragraph, we give an example to illustrate the second point.

A typical medical image such as a mammogram has size of $2K \times 2K$ and it requires 10M bytes to store a single image. Usually, a radiologist needs to compare 10 to 100 images at one time. Suppose a radiologist requests 20 mammograms through Ethernet whose transmission rate is 1.25M bytes per second. The radiologist needs wait for about 4 to 5 minutes to get the requested images. Also, it is often the case that some or all of the retrieved images are not what really wanted. So the radiologist needs wait another several minutes to get another group of images. This situation can be solved by storing multiresolution decomposition of images. For example, if the images are decomposed up to scale 4, then the lowest components have size of 128. A strategy to solve the preceeding problem is retrieving the lowest components as index first and seeing if the images are exactly what wanted. If so, retrieve the original images or retrieve other components to reconstruct the original images. Since the time for retrieving index is 256 times less than retrieving the whole images, the response time for retrieving 20 indexing items will be reduced to 2 to 3 seconds, which are within tolerance level. Therefore, using this strategy can avoid the unnecessary waiting.

Histogram is a useful tool for image comparison. Let $\mathbf{f} = \{f(i,j), i,j = 1,2,...,N\}$ denote an image with pixel values from T_{min} to T_{max}. Let $T_{min} = t_0 < t_1 < t_2 < ... < t_K = T_{max}$. The histogram $\mathbf{h} = \{h(k), k = 0,1,2,...,K-1\}$ of $\mathbf{f}$ is defined as the number of (i, j)'s with $t_k \leq f(i,j) < t_{k+1}$. Let $\mathbf{h_1}$ and $\mathbf{h_2}$ be the histograms of images $\mathbf{f_1}$ and $\mathbf{f_2}$ respectively. The difference of histograms (DOH) of $\mathbf{f_1}$ and $\mathbf{f_2}$ is defined by

$$DOH = \sum_{k=0}^{K-1} |h_1(k) - h_2(k)|.$$

Histogram comparison is an efficient method for comparing images. Two images will have the same histogram if one of them is obtained by displacing or rotating the other one. Histograms of two images will not differ very much if one of the images is degraded from the other one due to some amount of noise. But for a $2K \times 2K$ sized image, it is still very costing to calculate its histogram. Since the lowest components of the multiresolution decomposition of an image gives a compact description about the the image features, comparing histograms of the lowest components should be able to distinguish images in most cases. If a $2K \times 2K$ image is decomposed to scale 4, the cost of calculating histogram of the lowest component is reduced by 256 times. Therefore, multiresolution decomposition provides a method for quick image comparison, and multiresolution histograms can be used as index items for quick image retrieval.

However, this approach have two major drawbacks. First, it cannot search such images that are significantly different from the query image as a whole but contain local features identical with or similar to those in the query image. On the other hand, the local-feature-based search is desirable in some applications. For example, in computer aided diagnosis (CAD), radiologists often compare local image features rather than the whole images. Second, The multiresolution histogram comparison method still requires a large amount of computational time to compute histograms for an image database which contains a lot of large size images. For example, for each $2K \times 2K$ mammogram, we still need to compute histogram for the coarsest component with size of 128×128 if we decompose the image up to scale 4. Suppose that the database contains 10,000 such mammograms. You can imagine the computational cost for computing histograms for all coarsest components. In this paper, we develop a local-feature-based indexing and retrieval method to overcome these drawbacks.

4 Feature-oriented Indexing and Retrieval Algorithms

In this section, we present a local-feature-oriented image indexing and retrieval algorithm using Kwong and Tang's W-transform. The main difference between our algorithm and the algorithm in [6] is that our algorithm exploits local features but the Lee and Dickinson's [6] algorithm requires computing multiresolution histogram for the whole images.

The feature-oriented image indexing system is built by the following steps:

1. Select image features which are either Region of Interest (ROI)'s or regions used to identify the image.
2. Take W-transforms for each selected feature from scale 1 to scale J, where J usually takes an integer value between 3 to 5.
3. Compute histograms for multiscale decomposed feature images and save them as indices.

The algorithm for retrieving images where there exist the same or similar features as those in the query image is as follows:

1. Click features in the query image.
2. Take W-transforms for the selected features from scale 1 to J.
3. For each scale from J to 1,
   - compute histograms for the multiscale decomposed feature images;
   - compute DOH between the histogram of the query multiscale feature images and the histogram in the index;
   - select image if the DOH's are less than given thresholds.
4. Display the retrieved images.

The threshold in Step 3 of the retrieving algorithm is used to measure the degree of similarity: the less the threshold is, the more similar the feature contained in the retrieved image and that in the query image are. (If the threshold is zero, only the images with the exactly same features as in the query image will be retrieved.) In our numerical experiments, it was found histograms of multiscale feature images will not differ very much if we click a feature which has some displacement from that in the indexing system. Therefore, a slightly relaxed threshold is recommended to compensate the histogram error caused by little displacement even if one tries to retrieve the images with the same features.

There are several different loop strategies which can be used in Step 3 of the retrieving algorithm, depending how much accuracy is required. One may compare all DOH's or just the first one or two DOH's from coarse scale to fine scale to decide whether or not to terminate searching. In our numerical experiments, it was found that for most cases, the desired images can be found by comparing DOH for the coarsest scale if the threshold is carefully selected.

5 Experimental Results

We conducted simulation experiments to test the local-feature-oriented indexing and retrieval method using MATLAB for am MRI image database.

First, we build a feature-oriented index for the image database using the algorithm described in the previous section. The process of building index is an automatic process: we only need to pick the features in the image and the program will build the index itself. This is an important advantage of a content based approach over an SDL based approach.

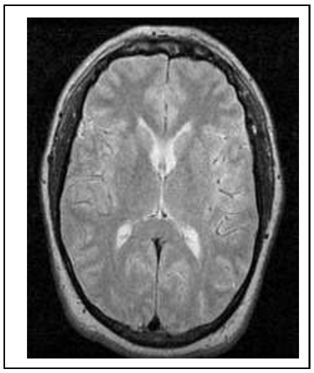
Figure 1 to 4 illustrates the retrieving process. Figure 1 is a typical MRI image with the size of 256×256 and 8 bit gray scales. Figure 2 is a selected feature inside the Figure 1 and is consisting of 60×80 pixels. Here we see that the size of selected image is not required to be a power of 2. Figure 3 is the multiresolution decomposition of Figure 2 using $W-$transform up to scale 3. Figure 4 is the histogram of the lowest frequency part of Figure 3. The histogram is compared with the indexing histograms and those images will be displayed if the difference between the query histogram and indexing histogram is less than some given threshold.

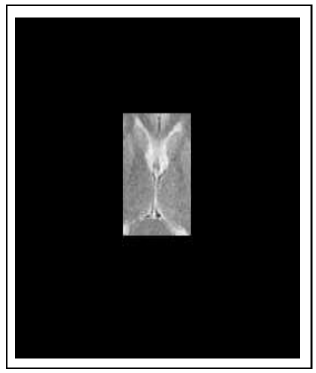
Selecting threshold is crucial in the retrieving algorithm: if the threshold is selected too small, the program will not tolerate any displacement of the selected features; if the threshold is selected too large, many undesired images will be retrieved. We tested this observation by examples. Figure 5 is average accumulative number of retrieved images for different DOH thresholds based on 4 sets of 22 feature images. Here we see that if a threshold is set to be larger than 300, two false images will be retrieved on average. A test on a larger image database is to be done in the future.

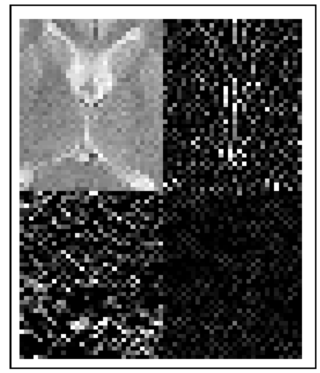
6 Conclusions

In summary, the feature-oriented image indexing and retrieval method not only can retrieve images that are similar or identical to the query images but also can retrieve images that contain features specified in the query images, even if the retrieved images as a whole might be very different from the query images. The feature-oriented method also provides a speed advantage over the global multiresolution histogram comparison method. This approach provides a promising application on managing large-scale image systems such as video databases and medical image databases. How to build the technique into different hierarchical storage systems remains as a future research topic.

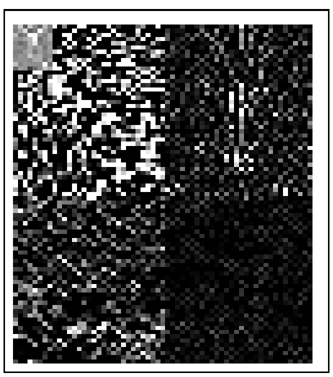
Figures

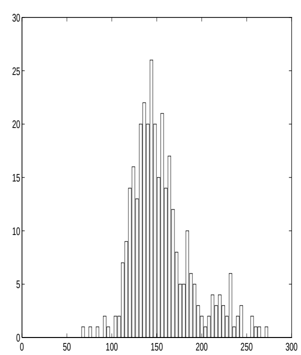


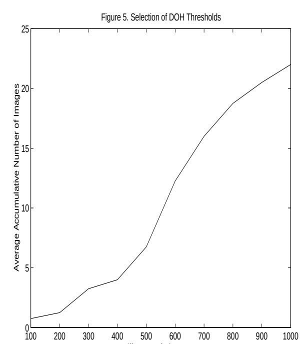




1. MRI Query Image   2. MRI Feature Image   3. L1 Decomposition







4. L3 Decomposition   5. Histogram (L2)   6. Threshold Selection

References

[1] Coifman, R. R., and Wickerhauser, M. V., *Entropy based algorithms for best basis selection,* IEEE Trans. on Information Theory, vol. 38, No. 2, pp.712-718, Mar., 1992.

[2] Daubechies, I., *Ten Lectures on Wavelets*, CBMS-NSF Series Appl. Math., SIAM, 1991.

[3] Kwong, Man Kam, and Tang, Peter P. T., *W-matrices and Nonorthogonal Multiresolution Analysis of Finite Signals of Arbitrary Length*, Argonne National Laboratory Preprint Series MCS-P449-0794, 1994.

[4] Kwong, Man Kam, *MATLAB Implementation of W-Matrix Multiresolution Analyses*, Argonne National Laboratory Preprint Series MCS-P449-0794, 1994.

[5] Lin, Biquan, *Wavelet Phase Filter for Denoising in Tomographic Image Reconstruction*, PhD Thesis, Illinois Institute of Technology, Dec., 1994.

[6] Lee, J., and Dickinson, B. W., *Multiresolution Video Indexing for Subband Coded Video Databases,* SPIE Proc. of Storage and Retrieval for Image and Video Databases II (1994), 162–169.

[7] Mallat, S. G., *A theory for multiresolution signal decomposition: The wavelet representation,* IEEE Trans. on Pattern Analysis and Machine Intelligence 11 (1989), 674–693.

[8] Mallat, S. G., and Hwang, W. L., *Singularity Detection and Processing with Wavelets,* IEEE Trans. on Information Theory, vol. 38, No. 2, pp.617-645, Mar., 1992.

[9] Pentland, A., Picard, R., Davenport, G., and Haase, K., *Video and Image Semantics: Advanced Tools for Telecommunications,* IEEE Multimedia (1994), 73-75.