Fast Wavelet Histogram Techniques for Image Indexing

M. K. Mandal, and T. Aboulnasr
School of Information Technology and Engineering
University of Ottawa, Ottawa, Canada - K1N 6N5
E-mail: mandalm@marge.genie.uottawa.ca

S. Panchanathan
Department of Computer Science and Engineering
Arizona State University, Tempe, Arizona, USA
Email: panch@asu.edu

Abstract

Image histograms have become very popular in image and video indexing applications because of their low complexity. However, histogram-based techniques are not efficient in retrieving texture images. Recently, a histogram technique that exploits the directional properties of wavelet transform has been proposed in the literature. Although, this technique provides a good retrieval performance for texture images, its complexity is very high. In this paper, we propose fast wavelet histogram techniques that provide a superior performance at a substantially reduced complexity.

1. Introduction

Image and video indexing techniques have become important with the recent advances in image and video compression standards (MPEG-4, MPEG-7). These techniques store and retrieve images based on their contents [1, 2]. The block chematic of a typical image archival and retrieval system is shown in Fig. 1. These techniques have applications in several areas including multimedia information systems, digital libraries, movie industry, and video on demand. Traditional databases use keywords as labels to quickly access large quantities of text data. However, representation of visual data with text labels needs a large amount of manual processing. In addition, the retrieval results might not be satisfactory when the query is based on features not abstracted by the associated keyword.

Histogram-based techniques have become popular in indexing applications due to their low complexity [3]. Here, the histogram of a query image is compared to the histograms of all candidate images in a database. A subset of images with least difference of image histograms (DOIH) is then retrieved. This technique generally performs well for natural images. However, it fails to provide a good indexing performance for texture images, since different texture images can have very similar histograms.

Smith *et al*. [4] have recently proposed a wavelet histogram technique (WHT) where an image is first decomposed using wavelets [5]. A multi-dimensional histogram is calculated from these band-images, which is then employed as an index. We note that the computational complexity of WHT is very high. In addition, it does not provide a good indexing performance for natural images.

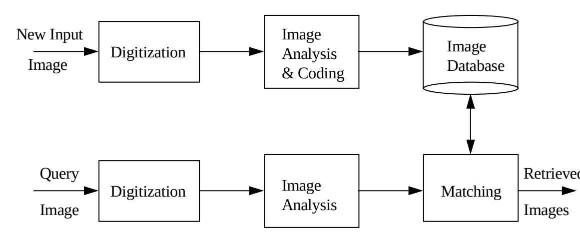


Fig. 1. Schematic of an image storage and retrieval system

In this paper, we propose fast wavelet histogram techniques (FWHT) which provides a good retrieval efficiency at a similar, or reduced complexity.

2. Wavelet histogram techniques

In this section, we provide a brief review of the wavelet histogram technique proposed by Smith *et al.* [4]. Here, the image is first decomposed to M stages using wavelets. The wavelet bands form a pyramid of M levels, where the level-k bands are the highpass bands after the k-th stage decomposition. A three-stage DWT decomposition is shown in Fig. 2, where level-1 bands consists of $\{A_7, A_8, A_9\}$, level-2 bands consists of $\{A_4, A_5, A_6\}$, and level-3 bands consists of $\{A_0, A_1, A_2, A_3\}$. *Intensity wavelet bands* (IWB) are then generated using the magnitudes of DWT coefficients from highpass bands of all levels. Each of these IWB's is upsampled to the full-size image by inserting zeros and subsequently passed through appropriate filters to obtain a texture channel. In [4], a simple pixel replication filter has been used for upsampling and filtering. The entire process for a three-stage wavelet decomposition is illustrated in Fig. 2. Here, nine texture channels are generated from nine highpass DWT bands. A texture point is then defined as a 9-D vector by considering texture channel values from the same location of all nine bands. Thus for a $M \times N$ image, there will be MN 9-D vectors. Each element of the 9-D vector is thresholded to two levels – high (1) and low (0). A wavelet histogram (with 512 bins) of all texture points, with 9-D thresholded vectors, is created and subsequently used as an index of the image.

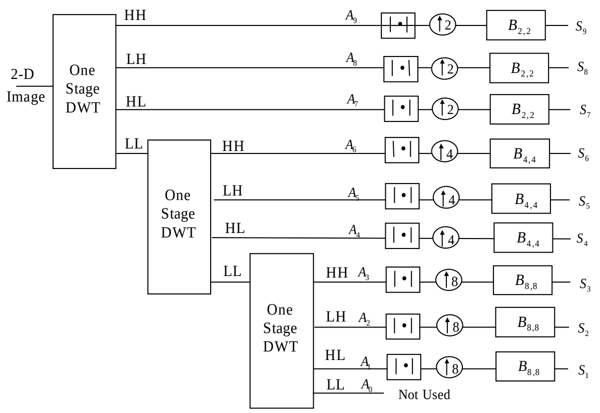


Figure 2. Wavelet histogram generation [4]

The computational complexity of wavelet histogram generation is shown in Table 1 (scheme-0, third column). It is observed that the total complexity is approximately $(3T+1)*MN$ where T is the number of channels. We note that the complexity of feature generation (*i.e.*, the computation of histogram) for DOIH technique is approximately MN operations. Hence, the complexity of WHT technique is significantly higher compared to DOIH technique. The higher complexity is mainly due to the fact that all highpass bands are to be upsampled, and thresholded. In addition, the multi-channel nature of the entire operation increases the complexity by a factor of T, the total number of channels.

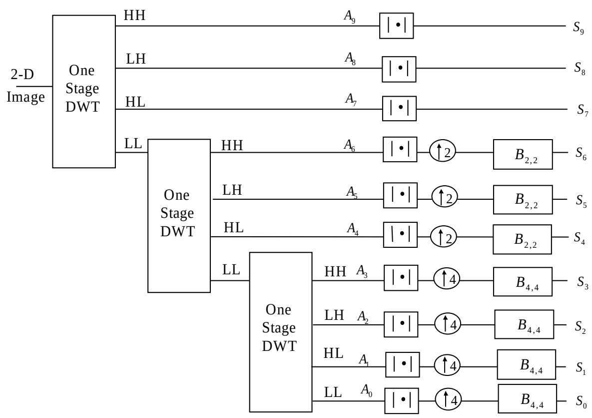
3. Proposed fast wavelet histogram techniques

The WHT technique [4] upsamples and filters all highpass wavelet bands such that each upsampled band contains the same number of coefficients as present in the original image (which can be considered as level-0 wavelet band). In other words, the level-k bands are upsampled by a factor of $2^k \times 2^k$. We note that there is significant redundancy present in the upsampled and subsequently filtered images. Hence, there is a potential to reduce this redundancy without degrading the performance.

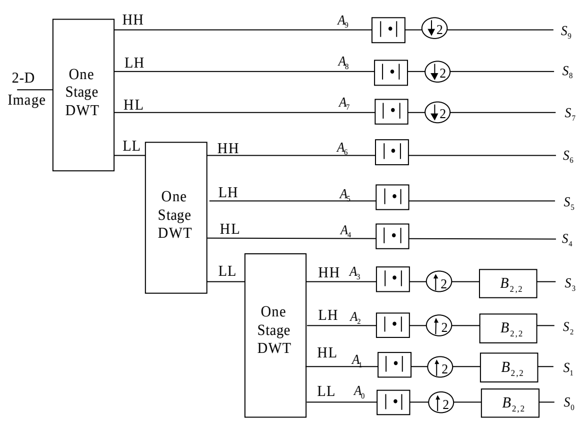
In this section, we propose fast WHT techniques (FWHT) exploiting the multiresolution nature of wavelet representation. In scheme-k FWHT technique, various highpass wavelet bands are appropriately upsampled or downsampled to have the same number of coefficients as in a level-k band, $k \in \{1,2,...M\}$ for M-stage decomposition. In addition, the zeroth band (*i.e.*, lowpass subband at level-M) is also included in WH generation. Fig. 3(a) shows histogram generation for scheme-1 FWHT technique with a 3-level decomposition. Here, the coefficients from level-1 bands are not altered, while the coefficients from level-2 and level-3 bands are upsampled by a factor of 4 (2x2), and 16 (4x4), respectively. A texture point is then defined as a 10-D (since 10 bands have been employed) vector consisting of values from the same location of all ten wavelet channels. Each element of the 10-D vector is thresholded to two levels – high (1) and low (0). Thus for a $M \times N$ image, there will be *MN/4* 10-D vectors. A wavelet histogram with 1024 bins is created from these vectors.

WH generation for scheme-2 FWHT is shown in Fig. 3(b). Here, the coefficients from level-2 bands are not altered while the coefficients from level-1 bands are downsampled by a factor of 2x2 and the coefficients from level-3 are upsampled by a factor of 2x2. Following a procedure similar to the scheme-1 FWHT, *MN/16* 10-D vectors are created. A wavelet histogram with 1024 bins is then created from these vectors. Finally, the wavelet histogram for scheme-3 FWHT (which has not been shown) can be generated by downsampling the coefficients from level-1, and level-2 bands by factors of 16 (4x4) and 4 (2x2), respectively and following the above procedure.

Histograms corresponding to selected FWHT techniques are shown in Fig. 4. It is observed that wavelet histograms generally have high peaks at a regular interval. The histograms are sparse at higher levels because of lower number of feature points. In addition, increase in the number of bins from 512 (because of 9-bands) to 1024 (because 10-bands) resulting from the inclusion of the zeroth band corresponds to a sparser histogram. The indexing performance of histogram-based techniques is generally not robust when the histograms are sparse. We now propose a modified version of FWHT techniques to address this issue. Here, we employ only the first few of the low-resolution wavelet bands for WH generation. The histograms of the modified FWHT are more smooth and dense since there are fewer histogram bins. Fig. 4 (c) shows the wavelet histogram, having 128 histogram bins, for modified scheme-2 FWHT using seven wavelet bands (A_0 - A_6). In addition to the increased smoothness of histogram, the complexity of this modified FWHT is lower than FWHT.



(a)



(b)

Figure 3. Schematic of a) scheme-1 and b) scheme-2 fast wavelet histogram generation (FWHT).

The information loss resulting from ignoring some subbands may lead to a degradation in the indexing performance. However, we note that out of the ten bands, three bands provide horizontal information, three bands provide vertical information, and three bands provide diagonal information. Since information along a particular direction is provided by at least three subbands (although at different scales), some of the higher resolution bands may be ignored without performance degradation.

4. Computational complexity of FWHT

In this section, we provide an estimate of the computational complexity of the FWHT techniques. There are two types of complexities involved - i) complexity of feature vector generation, and ii) complexity of feature vector comparison. We note that the feature vector may be generated only once, and stored along with each image. On the other hand, the complexity of feature comparison is involved each time a retrieval is performed.

The complexity of WH generation for scheme-k ($k \in \{0,1,2\}$) FWHT technique is provided in Table 1. We note that WHT can be considered as scheme-0 FWHT. Further, for scheme-0 and scheme-1 only upsampling is required, while for scheme-2, both upsampling and downsampling are required. In all schemes, upsampling and filtering has been implemented by simple pixel replication. It is observed from Table 1 that the complexity of scheme-0 (column 3) and scheme-1 (column 4) is proportional to the number of bands and output DWT coefficients. For scheme-2 FWHT, downsampling has been implemented as a weighted average of 2x2 or 4x4 DWT coefficients. Thus, the complexity of generating an output pixel by downsampling is greater than the corresponding complexity of generation by upsampling. Hence, the former has been given a larger weight (=4) in row 3/column 5 of Table 1. We note that the complexity of scheme-1, scheme-2, and scheme-3 (which is not shown in Table 1) FWHT techniques is 0.23, 0.08, and 0.04 times the complexity of WHT technique, respectively. This reduction in complexity of WH computation is mainly due to two factors: i) the reduction in the number of effective DWT coefficients for histogram calculation, ii) reduction in the number of bands (only 7 bands out of 10 bands) employed for upsampling and downsampling.

Table 2 compares the complexity of WH generation for scheme-2 FWHT using 7 and 10 bands. It is observed the complexity for the case of 7 bands is approximately half the complexity of the 10-band case.

The run-time complexity of FWHT techniques is similar to any DOIH/WHT technique, and is proportional to the number of histogram bins.

5. Performance of the Proposed Techniques

In this section, we evaluate the performance of the proposed techniques. We employ the *retrieval efficiency* as the performance criterion [6]. This is defined as follows: for each image i , in a database of size K , we manually list the similar images found in the database. Let, N_i , $1 \le i \le K$, be the number of such images. We then apply an indexing technique for a query image- q , and retrieve the first $(N_q + \tau)$ images. Here, τ is a positive integer, and is used as a tolerance for retrieval. If n_q is the number of successfully retrieved images, the efficiency of retrieval can then be defined as:

Table 1

Complexity (in operations) of computing wavelet histogram in scheme-{0,1,2). P: Total number of channels to be upsampled and filtered, Q: Total number of channels to be downsampled, T : Total number of channels, MN : Number of image pixels.

Main Modules	Sub-Modules	Approximate no. of Operations		
		Scheme-0	Scheme-1	Scheme-2
Upsampling & Thresholding	Absolute Value Calculation	$T*MN$	$P*MN/4$	$P*MN/16$
	Thresholding	$T*MN$	$P*MN/4$	$P*MN/16$
Downsampling & Thresholding	Absolute Value Calculation	-	-	$4*Q*MN/16$
	Thresholding	-	-	$Q*MN/16$
Histogram Calculation	Feature Point Generation	$T*MN$	$T*MN/4$	$T*MN/16$
	Histogram Calculation	MN	$MN/4$	$MN/16$
Total		$(3T+1)*MN$	$(2P+T+1)*MN/4$	$(2P+5Q+T+1)*MN/16$
Total (typ. value)		1.8e+6[a]	4.1e+5[b]	1.4e+5[c]

[a] for T =9, M =256, N =256, [b] for P =7, T =10, M =256, N =256, [c] for P =4, Q =3, T =10, M =256, N =256

$$\eta_R = \frac{\sum_{q=0}^{K} n_q}{\sum_{q=0}^{K} N_q}$$

Here, we have employed two databases of images. The first database, IDB1, contains 44 images of (size 248×256) wide varieties, including faces, natural scenes, animals, birds. Each image has five derivatives corresponding to i) normal, ii) translated rightward, iii) translated leftward, iv) rotated clockwise, v) rotated anti-clockwise. Thus, we have a total of 220 images in IDB1. The second database, IDB2, contains 50 Brodatz textures of size 256×256. Each texture has four similar images, *i.e.*, the total number of images is 200. A tolerance τ of 5 has been used in all cases. Daubechies 8 tap minimum phase wavelet has been used for WH generation. The histograms are compared in L^1 metric.

Table 2: Complexity of scheme-2 FWHT technique. *No. of Bands* (B) includes the zeroth band. P: Total number of bands to be upsampled, Q: Total number of bands to be downsampled, MN : Number of image pixels.

No. of Bands	Complexity (OP/image)	Typical Value[1]
7	$(2P+B+1)*MN/16$	3.0e+3
10	$(2P+5Q+B+1)*MN/16$	4.9e+4

[1]for P =4, Q =3, T =10, M =256, N =256

The performance of DOIH, WHT, and FWHT techniques on IDB1 and IDB2 databases is shown in Table 3. It is observed that WHT technique provides a performance of 85.20% on IDB2 database. The retrieval efficiency of WHT on IDB2 database is superior to that of DOIH technique. This is because the texture images in IDB2 have a strong directional property that is captured by the WHT technique. On the other hand, the DOIH technique captures only the global description of an image in spatial domain and is hence not as efficient. For the general database IDB1, the DOIH and WHT techniques provide a performance of 92.3% and 70.11%, respectively. Since the IDB1 database contains images with a few large objects, the image histogram is generally influenced by these objects. Hence, histograms of images containing similar objects are very similar. For this reason, DOIH technique provides a good performance. However, WHT technique provides a lower retrieval efficiency since the directional information is less prominent in these images.

Table 3: Effieicncy (η_R) versus complexity (ξ_G) of DOIH, WHT and FWHT (scheme-2) feature generation. FWHT employs 7 bands for IDB1 and 10 bands for IDB2 (zeroth band is included in both cases).

	IDB1 (General image)		IDB2 (Texture image)	
Technique	ξ_G (in op/image)	η_R (in %)	ξ_G (in op/image)	η_R (in %)
DOIH	6.5e+4	92.30	6.5e+4	72.00
WHT	1.6e+6	70.11	1.8e+6	85.20
FWHT	6.3e+4	80.68	1.3e+5	86.99

The performance of FWHT techniques on both IDB1 and IDB2 databases is shown in Fig. 5. For the IDB2 database, we observe the following: i) best performance is achieved with scheme-2, ii) WH generated with all wavelet bands provide the best performance, and iii) the inclusion of zeroth band does not seem to influence the retrieval efficiency. On the other hand, for the IDB1 database, we observe the following: i) best performance is achieved at level-1, ii) best performance is achieved around 7 bands, and iii) the inclusion of zeroth band significantly improves the retrieval efficiency.

The above observations confirm the expectations based on the properties of the images in the two databases. Images in IDB2 database have strong directional features and hence employing all highpass subbands provides the best result. Inclusion of the zeroth band is not important for this database. However, for the IDB1 database, the zeroth subband provides crucial information about the low frequency texture information. Hence, inclusion of this band is very important. In addition, the subbands at the highest resolution are not crucial for the images in IDB1. The directional information achieved by a few subbands of lower resolution is generally sufficient.

The above experimental results suggest that the complexity of the wavelet histogram generation can be reduced significantly while retaining (or even improving) the performance level. Table 3 shows the complexity and performance WHT and FWHT techniques for IDB1 and IDB2 database, respectively. It is observed that for IDB2, a complexity reduction by a factor of 10 is possible without degrading the performance. On the other hand, for IDB1 database, a complexity reduction factor of 30 is possible with improvement in performance.

If the WHs are stored in a database, the complexity of WH comparison for retrieval is directly proportional to the number of histogram bins. In this case, a FWHT with seven-band (with 128 bins) is eight times faster than a FWHT with ten bands (1024 bins).

6. Conclusions

Fast wavelet histogram techniques have been proposed in this paper to improve performance of image indexing systems. For natural images, the FWHT techniques provide a superior performance compared to WHT, at a reduced complexity, due to the inclusion of the zeroth wavelet band. On the other hand, for texture images, FWHT techniques provide a performance comparable to that of the WHT technique at a reduced complexity. The proposed techniques exploit the multiresolution property of DWT and are particularly useful when a wavelet-based coding scheme is also used to compress the images in the database.

References

[1] F. Idris and S. Panchanathan, "Review of image and video indexing techniques," *Journal of Visual Communication and Image Representation*, June 1997.

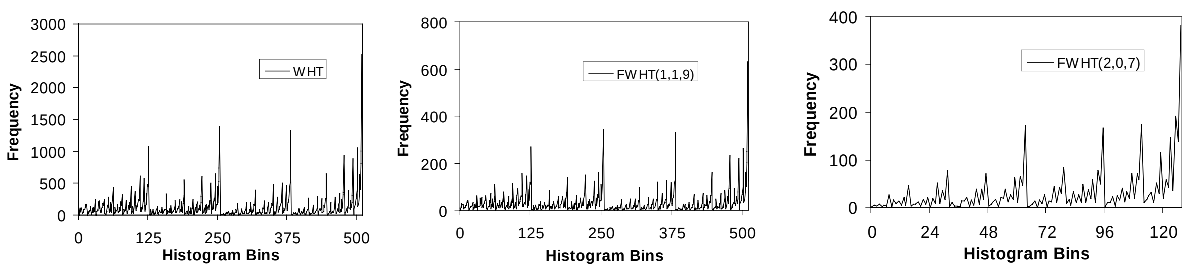
[2] B. Furht, S. W. Smoliar and H. Zhang, *Video and Image Processing in Multimedia Systems*, Kluwer Academic Publishers, 1995.

[3] M. J. Swain and D. H. Ballard, "Color indexing," *International Journal of Computer Vision*, Vol. 7, No. 1, 1991.

[4] J. R. Smith and S. F. Chang, "Automated binary texture feature sets for image retrieval," *Proc. of ICASSP*, Atlanta, May 1996.

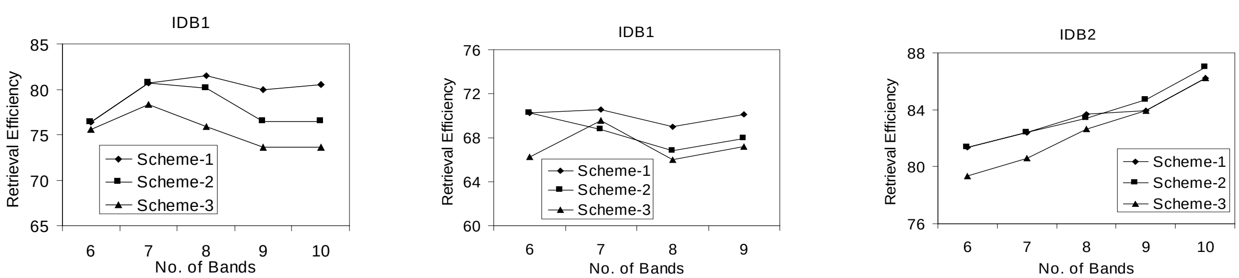
[5] J. Woods, *Subband Image Coding*, Kluwer Academic Publishers, 1991.

[6] B. M. Mehtre, M. S. Kankanhalli, A. D. Narasimhalu and G. C. Man, "Color matching for image retrieval," *Pattern Recognition Letters*, Vol. 16, pp. 325-331, March 1995.



(a) (b) (c)

Figure 4. Typical wavelet histograms. a) histogram for WHT. (b) & (c) histograms for FWHT. FWHT(p,q,r) refers to WH calculated from subband [q,(q+r-1)], with scheme-p.



(a) (b) (c)

Figure 5. Comparison of indexing performance of various wavelet-histogram techniques. "No. of bands" refers to the number of bands employed for WH generation. (a) & (c) includes zeroth band, (b) does not.