Proceedings of 2009 APSIPA Annual Summit and Conference, Sapporo, Japan, October 4-7, 2009

A Fast Approach for Identifying Similar Features in Retrieval of JPEG and JPEG2000 Images

K.O. Cheng, N.F. Law and W.C. Siu
Center for Signal Processing, Department of Electronic and Information Engineering,
The Hong Kong Polytechnic University, Hong Kong.
E-mail: {encko, ennflaw, enwcsiu}@polyu.edu.hk

Abstract— **As digital images are often in compressed forms, image retrieval involves full decoding of images prior to feature extraction. The decoding process can be computation-expensive so feature extraction in compressed domain is desired. In this work, wavelet-based features are extracted as unified features for retrieval of JPEG and JPEG2000 images. A fast algorithm is proposed to approximately transform a JPEG image in the block-based discrete cosine transform (BDCT) domain to wavelet domain so that wavelet-based features can be extracted directly from JPEG images. Our proposed algorithm consists of a multiresolution reordering and a filter bank structure. The former is used to provide a rough approximation of wavelet subbands from BDCT coefficients in bandpass subbands in fine scales while the latter is used to provide an accurate approximation in bandpass subbands in coarse scales. Our theoretical analysis shows that the proposed algorithm can reduce the complexity by at least 79% when comparing with the straight forward approach that uses an inverse BDCT followed by wavelet transform. Besides the reduction in computational complexity, the experimental results demonstrate that our proposed conversion approach has higher retrieval performance than the pure multiresolution reordering approach.**

I. Introduction

To avoid manual annotation, subjective descriptions and limitation of vocabularies in describing visual content [1], content-based image retrieval (CBIR) has been studied as an alternative way for image access. It is particularly favorable to large image databases such as digital libraries. In CBIR, low-level features like color, texture and shape are extracted. Many existing feature extraction methods assume images in spatial domain. But in most cases, images are compressed. Therefore, additional operations are incurred to decompress the images to spatial domain for feature extraction.

In order to reduce complexity, recent research has been focused on feature extraction in compressed domain, [2, 3]. Nevertheless, the problem has not been completely solved due to the fact that images can be compressed in different formats, which involve different technologies. For example, JPEG uses block-based discrete cosine transform (BDCT) for decorrelation and energy compaction but another format JPEG2000 uses wavelet transform instead. This motivates analysis in similar features extraction from various domains such as JPEG and JPEG2000 [4, 5].

The extraction of similar features for JPEG and JPEG2000 compressed images usually exploits similarity of transform characteristics in frequency domain. For example in [5], multiresolution reordering [2] is used to partition DCT coefficients to follow the octave frequency decomposition of wavelet transform. As a result, similar wavelet-based features, which have been verified to provide promising performance in image retrieval [2, 6], can be extracted directly from both types of compressed images. In spite of high efficiency, the approximation accuracy of the multiresolution reordering approach is limited as only re-arrangement of DCT coefficients is performed. In this paper, we derive a filter bank structure that can convert BDCT to wavelet transform losslessly. In order to avoid substantial increase in computational complexity due to filtering operations, the proposed filter banks have been applied to estimate bandpass subbands in coarse scales only while the multiresolution reordering approach is performed for other subbands. As a result, the overall computational complexity is lower than the straight-forward conversion from BDCT to wavelet transform. At the same time, a good accuracy for the transform coefficients is maintained.

In the next section, a brief review on a subband-filtering model [5] is provided. In Section III, the fast method for conversion from BDCT domain to wavelet domain is proposed. Section IV shows the results of retrieval experiments. Finally, the paper is concluded in Section V.

II. Review on Subband-Filtering Model

The subband-filtering model proposed in [5] allows comparison between block-based discrete cosine transform (BDCT) and wavelet transform in filter responses. The model acts as a starting point for the proposed conversion method. For simplicity, only 1D space is considered. For an 1D signal $x[n]$, the DCT coefficient at the ith frequency for the kth block of length 2^{r_D}, $a_i[k]$, can be described by

$$a_i[k] = \sum_n x[n] u_i[n - k \cdot 2^{r_D}] \tag{1}$$

where $0 \le i < 2^{r_D}$ and $u_i[n]$ can be expressed as

$$u_i[n] = \begin{cases} w(i)\cos\left((2n+1)i\pi / 2^{r_D+1}\right), & 0 \le n \le 2^{r_D} - 1 \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

$$\text{where } w(i) = \begin{cases} 2^{-r_D/2}, & i = 0 \\ 2^{-(r_D-1)/2}, & 1 \le i \le 2^{r_D} - 1 \end{cases}. \tag{3}$$

In other words, $a_i[k]$ can be found by filtering $x[n]$ with $u_i[-n]$ and then downsampling by 2^{r_D}. For the 1D wavelet transform of level r_W, subbands can be obtained using the cascade structure that a two-channel filter bank is iteratively applied to the lowpass output in the previous level. Denote the lowpass and highpass filters in the two-channel filter bank by $\tilde{h}[-n]$ and $\tilde{g}[-n]$ respectively. The wavelet coefficients at level j ($1 \le j \le r_W$), $t_j[l]$, can be obtained by filtering and then downsampling in such a way that

$$t_j[l] = \sum_n x[n]\tilde{\psi}_j[n - l \cdot 2^j] \quad (4)$$

where $\tilde{\psi}_j[n]$ can be expressed in terms of $\tilde{h}[n]$ and $\tilde{g}[n]$ in z domain by

$$\tilde{\psi}_j(z) = \begin{cases} \tilde{G}(z), & j = 1 \\ \tilde{G}(z^{2^{j-1}})\prod_{k=0}^{j-2}\tilde{H}(z^{2^k}), & 2 \le j \le r_W \end{cases}. \quad (5)$$

Similarly, the scale coefficients at level r_W, s_{r_W}, are related to $x[n]$ as in (4) but with $\tilde{\psi}_j[n]$ replaced by another function $\tilde{\phi}_{r_W}[n]$. By using the subband-filtering models, the similarity between BDCT and wavelet transform can be measured based on the filters similarity in their equivalent filter banks.

III. Proposed Fast Unified Feature Extraction Algorithm

Although the multiresolution reordering approach in [5] utilizes the similarity in frequency partitioning and spatial locations, approximation is still deteriorated by the dissimilarity in filter responses expressed in the subband-filtering models. Thus, we implement the conversion from BDCT to wavelet transform in a single filter bank structure. Then, the structure is simplified and applied partially so as to reduce its complexity below the straight-forward approach which consists of inverse BDCT followed by wavelet transform. The new conversion structure is used to extract wavelet-based features directly from JPEG compressed images so as to compare without full decoding images compressed in JPEG2000 as well as in JPEG formats.

A. *Filter banks for Conversion from BDCT Domain to Wavelet Domain*

To simplify the discussion, we consider the BDCT and wavelet transform in 1D space because the results can be extended to 2D space by using their separable natures. Without loss of generality, assume that $r_D = r_W = r$. Based on the subband-filtering model, it can be shown that the original signal $x[n]$ can be reconstructed from its BDCT coefficients by,

$$x[n] = \sum_{i=0}^{2^r-1}\sum_k a_i[k]u_i[n - k \cdot 2^r]. \quad (6)$$

Substituting (6) into (4), we have

$$t_j[l] = \sum_{i=0}^{2^r-1}\sum_k a_i[k]\left(\sum_n u_i[n - k \cdot 2^r]\tilde{\psi}_j[n - l \cdot 2^j]\right). \quad (7)$$

By expressing l as $l = k' \cdot 2^{r-j} + l'$, where $k' \in Z$ and $l' = 0, 1, ..., 2^{r-j} - 1$, and with some manipulations, the wavelet coefficients can be rewritten as

$$t_j[k' \cdot 2^{r-j} + l'] = \sum_{i=0}^{2^r-1}\sum_k a_i[k]b_{i,j,l'}[k' - k] \quad (8)$$

where $b_{i,j,l'}[k] = \sum_n u_i[n + k \cdot 2^r]\tilde{\psi}_j[n - l' \cdot 2^j]$. (9)

Equation (8) actually computes the l'th polyphase component of $t_j[k]$, $\hat{t}_{j,l'}[k]$. Its second summation evaluates convolution between $a_i[k]$ and $b_{i,j,l'}[k]$. Hence, equation (8) is equivalent to

$$\mathbf{T}_j(z) = \mathbf{C}_j(z)\mathbf{A}(z) \quad (10)$$

where $\mathbf{T}_j(z) = [\hat{T}_{j,0}(z) \quad \hat{T}_{j,1}(z) \quad \cdots \quad \hat{T}_{j,2^{r-j}-1}(z)]^T$,

$\mathbf{A}(z) = [A_0(z) \quad A_1(z) \quad \cdots \quad A_{2^r-1}(z)]^T$ and

$\mathbf{C}_j(z) = \left[c_{l',i}^{(j)}(z)\right] = \left[B_{i,j,l'}(z)\right]$.

$\mathbf{C}_j(z)$ can be regarded as a polyphase matrix in a filter bank as shown in Fig. 1. Alternatively, a filter bank can be implemented to convert BDCT subbands into wavelet subbands with a filter $G_{i,j}(z)$ defined below

$$G_{i,j}(z) = \sum_{l'=0}^{2^{r-j}-1} z^{-l'}B_{i,j,l'}(z^{2^{r-j}}). \quad (11)$$

Due to similarity in calculation, the scale coefficients can also be obtained using the similar structure in Fig. 1.

Equations (8) and (9) suggest that in general, a wavelet subband at a given level may depend on DCT frequency components at any frequency and more than one DCT coefficient at a particular frequency may contribute to construct a wavelet subband at each sampled point. Our results thus reflect the deficiency of the multiresolution reordering approach in construction of wavelet subbands.

B. *Proposed Fast Algorithm for Conversion from BDCT Domain to Wavelet Domain*

In Section III.A, we have developed a filter bank approach

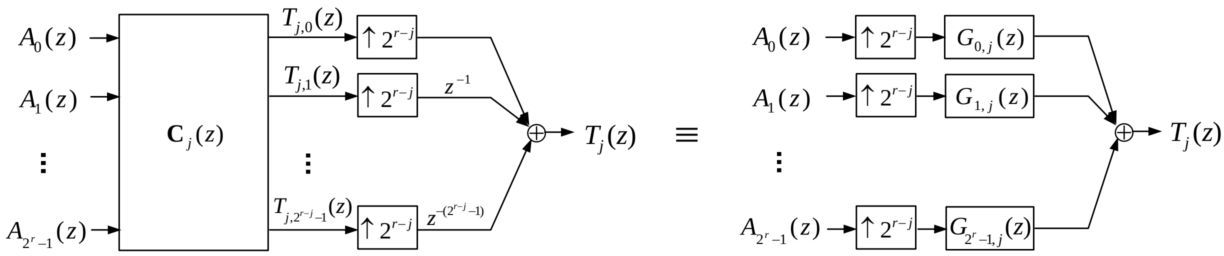

Fig. 1. The filter bank for calculating the j-th wavelet subband from BDCT subbands.

to convert BDCT coefficients into wavelet coefficients perfectly. However, the straight-forward approach can be implemented using fast algorithms for inverse DCT and wavelets [7, 8]. In order to reduce the complexity of our proposed method, partial conversion strategy has been adopted such that small filter coefficients in (9) are truncated and only a few DCT components are selected for calculation of each wavelet subband. These are reasonable because wavelets are localized in space and frequency. Besides, substantial reduction of computation complexity can be achieved by applying the filter bank structure only to bandpass subbands at coarse scales (i.e., the third level for 8×8 block) while the multiresolution reordering is performed for other subbands. This compromises the performance in approximation accuracy and computational complexity effectively because natural images mainly possess mid and low frequency contents and coarse bandpass subbands have small number of points only.

The overall computational complexity of the proposed fast conversion approach for JPEG images is summarized in Table I. In the calculation of computational cost, CDF 9/7 kernel is assumed in the wavelet transform. The straight-forward approach which adopts the fast inverse scaled-DCT [7] and lifting implementation of wavelet transform [8] is included for comparison. Even comparing with the straight-forward approach in the fast implementation, high percentage of reduction in additions and multiplications operations can be found as 91.4% and 79.4% respectively.

C. Feature Extraction and Similarity Measure

Using our proposed fast algorithm, BDCT coefficients of JPEG images are transformed to wavelet domain for further feature extraction. As for JPEG2000 images, they are represented in wavelet coefficients already. Thus, wavelet-based features are calculated as unified features for retrieval of images stored in the two formats. Energy feature based on L2 norm is extracted from each subband while seven normalized central moments are calculated from a significant map of each bandpass subband. Weighted similarity measurement used in WaveGuide [6] is employed to integrate distance measurements on the two different feature types.

IV. EXPERIMENTAL RESULTS

The retrieval performance of the proposed unified features was studied through experiments on 1800 images of nine classes used in [5]. In order to construct a database composed of both JPEG and JPEG2000 images, each of the images was used to generate one JPEG copy and one JPEG2000 copy.

TABLE I
COMPARISON OF NUMBERS OF ARITHMETIC OPERATIONS BETWEEN PROPOSED FAST CONVERSION METHOD AND STRAIGHT-FORWARD CONVERSION FOR AN IMAGE OF SIZE N×N

Operations	Proposed, C_p	Straight-forward, C_s	% of reduction, 100%×(C_s - C_p)/ C_s
Additions	$1.53N^2$	$17.72N^2$	91.4%
Multiplications	$1.80N^2$	$8.72 N^2$	79.4%

Hence, there were 3600 images in total. Each image was used as a query image once while the remaining images excluding the one, which was the same as the query image but compressed in different format, were used to construct a database in a retrieval system. Precision and recall [9] were calculated over different number of returned images. Meanwhile, they were averaged over all the query images for evaluation. The experiment was conducted for images in five compression ratios of 1.6, 5, 10, 20 and 40.

Figure 2 shows average precision against average recall at compression ratios of 1.6, 10 and 40. The results at compression ratios of 5 and 20 are omitted as the shapes of their precision-recall curves are similar to those at compression ratios of 1.6 and 40 respectively. The proposed algorithm had been compared with the straight-forward approach and the multiresolution reordering approach. The straight-forward approach imposes the upper bound on the retrieval performance because all JPEG images were transformed into wavelet domain without error for feature extraction. In general, the proposed algorithm outperforms the multiresolution reordering approach with largest improvement found at compression ratio equal to 10. For example, at this compression ratio and with the number of returned images set to 1, the proposed algorithm attains a precision of 0.9072 and a recall of 0.0023 while the multiresolution reordering approach has a precision of 0.8714 and a recall of 0.0022. When the number of returned images increases to the number of relevant images stored in the database, both the precision and the recall of the proposed algorithm are 0.4562 which are higher than those of the multiresolution reordering approach, 0.4076.

All of the three algorithms degrade in performance but with different degrees when the compression ratio increases. At low compression ratio, the three algorithms exhibit smaller difference. When compression ratio reaches 10, the performance of the proposed algorithm starts to drop but with smaller amount than that of the multiresolution approach. It is possibly caused by the fact that high frequency components of images are significantly distorted in JPEG compression when compression ratio is high. This implies that the distinguishing power of the wavelet-based features replies on those near the low frequency range. As the proposed algorithm has better approximation in bandpass subbands at the coarse level, its performance remains close to that of the straight-forward approach. When the compression ratio further increases, many high frequency coefficients would be quantized to zero in JPEG images. This means the distinguishing power of features from high frequency subbands obtained by the multiresolution reordering almost vanishes but the straight-forward approach would not as many wavelet coefficients at fine scales can be non-zero due to frequency leakage to low frequency components. As a result, the performance of the proposed algorithm becomes close to that of the multiresolution reordering approach. Despite that, the proposed algorithm should still significantly outperforms the multiresolution reordering approach in real situation because the typical range of compression ratio is around 10.

V. CONCLUSIONS

In retrieval, feature extraction in compressed domain has recently attracted a lot of research interests. In this work, similar features in JPEG and JPEG2000 compressed domains are explored. To reduce complexity, a fast conversion scheme from BDCT to wavelet transform is proposed. In the proposed scheme, multiresolution reordering is employed to roughly construct bandpass subbands at the fine scales and a lowpass subband while a filter bank structure is developed to approximate bandpass subbands at coarse scales to achieve a high accuracy. Since natural images usually have important features in low and mid-frequency range, the proposed method actually distributes heavier computational load to subbands with richer image characteristics. It has been shown that the proposed fast conversion method can reduce the number of arithmetic operations by at least 79.4% when comparing with the straight forward approach. Meanwhile, our retrieval experiments demonstrate that the proposed method results in higher retrieval performance than the pure multiresolution reordering approach especially in mid compression ratio.

ACKNOWLEDGMENT

This work is supported by the Center for Signal Processing, Department of Electronic and Information Engineering, the Hong Kong Polytechnic University, under the CERG Grant (PolyU 5215/08E) of the Hong Kong SAR Government. K.O. Cheng acknowledges the postdoctoral research fellowship provided by the University.

REFERENCES

[1] T. Sikora, “The MPEG-7 Visual Standard for Content Description – An Overview,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 6, pp. 696 – 702, June 2001.

[2] Y.-L. Huang and R.-F. Chang, “Texture Features for DCT-coded Image Retrieval and Classification,” *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol.6, pp.3013-3016, March 1999.

[3] F. Arnia, I. Iizuka, M. Fujiyoshi and H. Kiya, “Fast Method for Joint Retrieval and Identification of JPEG Coded Images Based on DCT Sign,” *IEEE International Conference on Image Processing*, vol.2, pp.II-229 – II-232, Sept. – Oct. 2007.

[4] J. Liu and H. Gu, “Image Retrieval in Various Domains”, *Computers & Graphics*, vol.27, no.5, pp.807-812, Oct. 2003.

[5] K. M. Au, N. F. Law and W. C. Siu, “Unified Feature Analysis in JPEG and JPEG 2000-compressed Domains”, *Pattern Recognition*, vol.40, no.7, pp.2049-2062, July 2007.

[6] K.C. Liang and C.-C.J. Kuo, “WaveGuide: A Joint Wavelet-Based Image Representation and Description System”, *IEEE Transactions on Image Processing*, vol. 8, no. 11, pp.1619 – 1629, Nov. 1999.

[7] E. Feig, “A Fast Scaled-DCT Algorithm”, *Proceeding of SPIE: Image Processing Algorithms and Techniques*, vol.1244, pp.2 – 13, 1990.

[8] W. Sweldens, “The Lifting Scheme: a Custom-design Construction of Biorthogonal Wavelets”, *Applied and Computational Harmonic Analysis*, vol.3, pp.186-200, 1996.

[9] D. Feng, W. C. Siu and H. J. Zhang, Multimedia Information Retrieval and Management: Technological Fundamentals and Applications, Springer-Verlag Berlin Heidelberg, 2003.

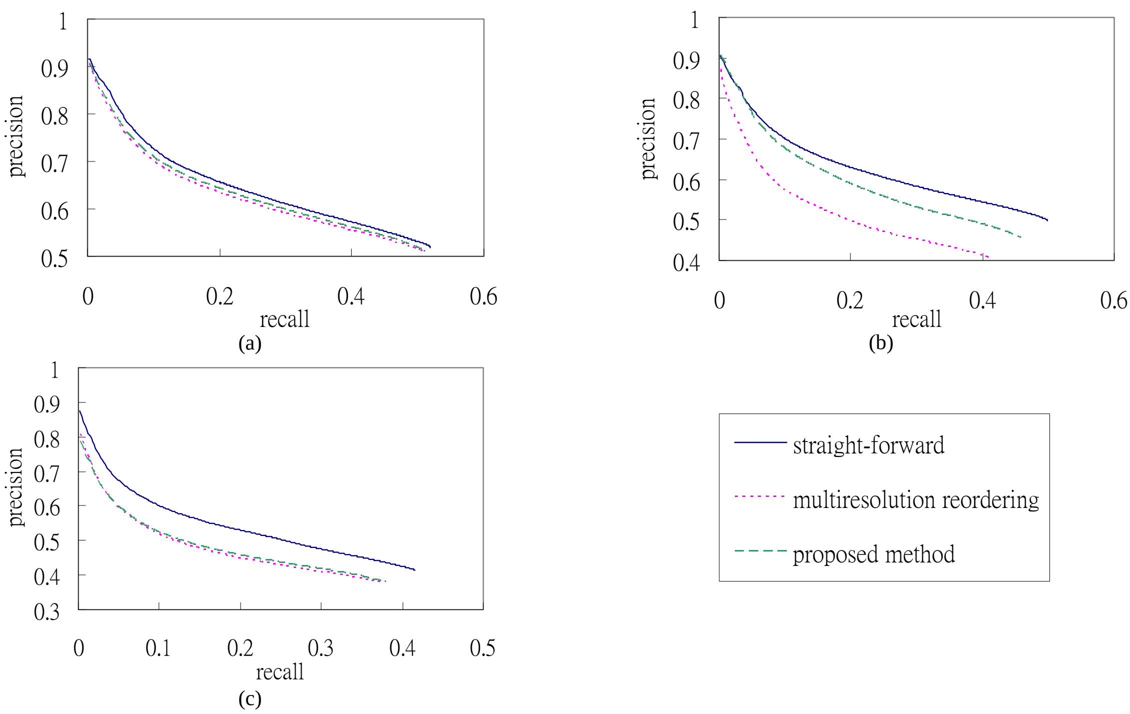

Figure 2. Average precision vs average recall at compression ratio (a) 1.6, (b) 10 and (c) 40.