HAVE 2005 – IEEE International Workshop on
Haptic Audio Visual Environments and their Applications
Ottawa, Ontario, Canada, 1-2 October 2005

Lifted Wavelet-Based Image Dataset Compression with Column Random Access for Image-Based Virtual Environment Navigation

Kehua Jiang , Eric Dubois
School of Information Technology and Engineering, University of Ottawa
800 King Edward Ave., Ottawa, Ontario, K1N 6N5, Canada

***Abstract* — *A lifted wavelet-based compression scheme with effective image-column random access for coding image datasets is proposed. This image dataset compression scheme is based on the approach of a lifted wavelet transform combined with embedded block entropy coding. This approach is adapted to compressing the image datasets used for image-based virtual environment navigation and facilitating a column random access mechanism. A specific data structure is designed to achieve efficient compression and accommodate column random access. The code stream syntax is constructed to efficiently assemble bit streams corresponding to the designed data structure. The two-level index tables in the code stream are used to assist random access. This image dataset compression scheme features a proper compromise between the coding efficiency and the random access flexibility, and thus ensures fast interactive image rendering. It reduces the decoding delay, simplifies the implementation of a random access mechanism and easily facilitates real-time image decoding. Experimental results of the software implementation of this image dataset compression scheme show its superior coding performance over the standard JPEG and JPEG 2000 schemes.*

***Keywords* — *Wavelet transform, lifting scheme, image dataset compression, random access, embedded coding.*

I. INTRODUCTION

An image-based virtual environment navigation system can provide users with realistic and immersive exploration experiences of real-world environments by making use of image-based rendering (IBR) techniques. IBR takes image datasets rather than conventional geometric models as primitive inputs to synthesize photo-realistic novel views. The image datasets constitute the fundamental database for image-based virtual environment navigation. They involve a large amount of data. Compression of the image datasets is crucial to reducing the required storage space and transmission bandwidth and putting image-based virtual environment navigation into practical use.

Wavelet transform proves to be an effective and efficient technique for image and video sequence compression. The ability to provide spatial, temporal and quality scalabilities makes wavelet transform more attractive for some applications with scalable coding requirements. However, wavelet-based compression schemes have complex implementation structures due to the use of filter banks for wavelet analysis and synthesis. The development of the lifting scheme [9] greatly alleviates this problem. It leads to faster, memory-saving, in-place wavelet decomposition and reconstruction and makes wavelet transforms more feasible for practical applications. For this reason, the lifted wavelet transform is adopted in JPEG 2000 [4] as a standard coding technique for still image compression. The wavelet transform needs to combine with an effective entropy coding approach to form an efficient compression scheme. Embedded coding is a desired feature of the entropy coding algorithms to obtain the target output bit stream rates and control the distortion of the compressed images. Shapiro [7] developed the embedded zero-tree wavelet algorithm (EZW). It is based on the hierarchical subband decomposition of the wavelet transform, and generates a fully embedded bit stream by assembling codes in order of importance. Said and Pearlman [6] offered an alternative exposition of the EZW principles. They proposed a new extension and implementation of progressive embedded encoding through set partitioning in hierarchical trees (SPIHT).

Compression of the image datasets used for image-based virtual environment navigation is quite different from the conventional problem of generic image and video sequence compression. Some special requirements for compressing the image datasets need to be satisfied. The interactive mode of image-based virtual environment navigation requires fast image synthesis enabled by efficient random image access and decoding. Some compression schemes for image datasets have been proposed based on the wavelet transform. Luo et al. [5] presented a wavelet-based compression scheme for coding concentric mosaic image datasets. To support random access to the coded image dataset and minimize the relevant computational expense, Wu et al. [11] proposed the progressive inverse wavelet synthesis (PIWS) scheme. As an effort to further exploit the cross-image redundancy and improve the coding efficiency of the wavelet-based compression scheme, a smart rebinning scheme is presented in [12]. Girod et al. [3] applied wavelet transform for coding light field image datasets. They developed a novel approach with a disparity-compensated lifting scheme. EZW and SPIHT are important embedded coding techniques widely applied in wavelet-based compression schemes. They exploit the inter-subband redundancy to enhance the compression efficiency. However, this restricts the random access flexibility to coded image datasets. Taubman [10] developed an image compression algorithm of embedded block coding with optimized truncation (EBCOT). It partitions the decomposition coefficients in each sub-band into blocks and provides a potential of random access to code blocks.

A lifted wavelet-based compression scheme with effective image-column random access for coding image datasets is proposed in this paper. The proposed scheme follows a JPEG 2000-like approach based on the framework of a lifted wavelet transform combined with embedded block entropy coding. This approach is adapted to compressing the image datasets used for image-based virtual environment navigation and facilitating a column random access mechanism. The random access to image datasets at the image column level is the key requirement for compressing image datasets for efficient interactive image rendering in quite a number of image-based rendering techniques, such as rendering with concentric mosaics [8], plenoptic stitching [1] and image rendering based on panoramic image datasets [2]. A specific data structure is designed to achieve efficient compression and accommodate column random access. The code stream syntax is constructed to efficiently assemble bit streams corresponding to the designed data structure. The two-level index tables in the code stream are used to assist random access. A tradeoff between the compression efficiency and random access flexibility is reached. The proposed compression scheme reduces the decoding delay, simplifies the implementation of a random access mechanism and easily facilitates real-time image decoding due to its independent intra-image coding feature, although some coding efficiency is sacrificed because no inter-image redundancy is exploited. This image dataset compression scheme has a simple implementation structure compared to some other schemes like the modified MPEG family.

This paper is organized as follows. After this introductory section, Section II gives a discussion of the lifted wavelet-based compression scheme with image-column random access for image-based virtual environment navigation. Simulation results of a software implementation of the proposed scheme applied to compressing sample image datasets are provided in Section III, compared with the results of JPEG and JPEG 2000 schemes. Finally, some conclusions are drawn in Section IV.

II. LIFTED WAVELET-BASED COMPRESSION SCHEME WITH COLUMN RANDOM ACCESS

The proposed image dataset compression scheme is based on an efficient lifted discrete wavelet transform algorithm combined with embedded block entropy coding. This scheme is adapted to facilitate spatial random access required by image dataset compression. A block diagram depicting the structure of the lifted wavelet-based image dataset compression approach is shown in Fig. 1. Raw format image datasets which are intended to be used as basis

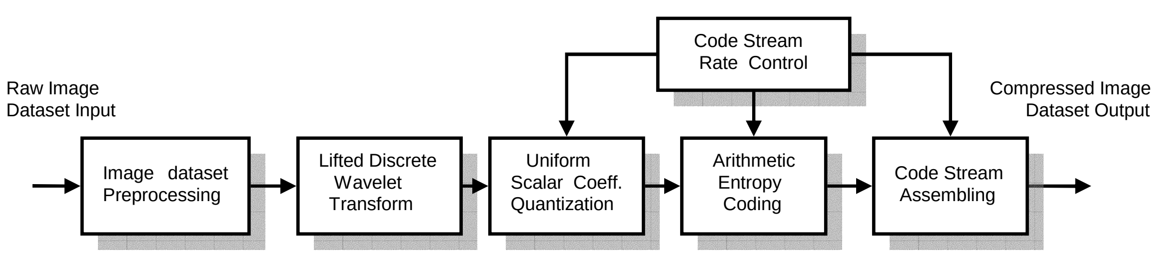


Figure 1. Block Diagram of the Lifted Wavelet-Based Image Dataset Compression Approach

images to synthesize novel views are taken as the input of the compression scheme. The preprocessing process performs operations on input image datasets for image conversions. DC level shifting is carried out, if necessary. Inter-component transformation is conducted to exploit the redundancy between image components. An efficient wavelet kernel is applied to transform image samples into decomposition coefficients. This wavelet transform is factorized into lifting operations. The resulting decomposition wavelet coefficients are quantized in the uniform scalar quantization process. Then, the quantized coefficients are partitioned into rectangular blocks of proper size. Each quantized coefficient block is independently encoded with an embedded arithmetic entropy encoder. The bit streams from all code-blocks are assembled into a code stream to produce the compressed image dataset output. The rate of the output code stream is controlled by changing the values of the quantization step sizes and truncating the bit streams of code-blocks.

A. Lifted Bi-orthogonal Daubechies 9/7 DWT

The discrete wavelet transform is a powerful tool for image compression due to its capability of efficient image representations with inherent scalabilities. In this proposed compression scheme, a two-dimensional separable discrete wavelet transform is applied for each individual basis image in the image datasets by utilizing the bi-orthogonal Daubechies 9/7 wavelet transform kernel for its high performance in image compression. The wavelet transform is factorized into a lifting operation structure for fast computation and memory saving. The lifting operation of a one-dimensional forward bi-orthogonal Daubechies 9/7 wavelet decomposition is depicted in Fig. 2. The input image samples are represented by S_i (i: $0, 1, 2, \ldots$) on the left side. The high-pass and low-pass wavelet decomposition coefficients are output on the right side. The actual output of high-pass coefficients are modified by a weight $K_h = 1.230\ 174\ 105$, and the low-pass coefficients are modified by $K_l = 1 / K_h$. The lifting coefficients are represented by W_i (i: 1~4). Each level of wavelet decomposition results in four sub-bands of decomposition images with halved resolution in both the horizontal and vertical directions. The input image samples are extended periodically and symmetrically beyond image boundaries to ensure the lifting operation working properly at the boundaries. In order to further decorrelate the decomposition coefficients, Mallat or dyadic decompositions are utilized for additional decompositions in the low sub-bands. With N levels of wavelet decomposition, the compressed image can be reconstructed at $N+1$ different resolution levels.

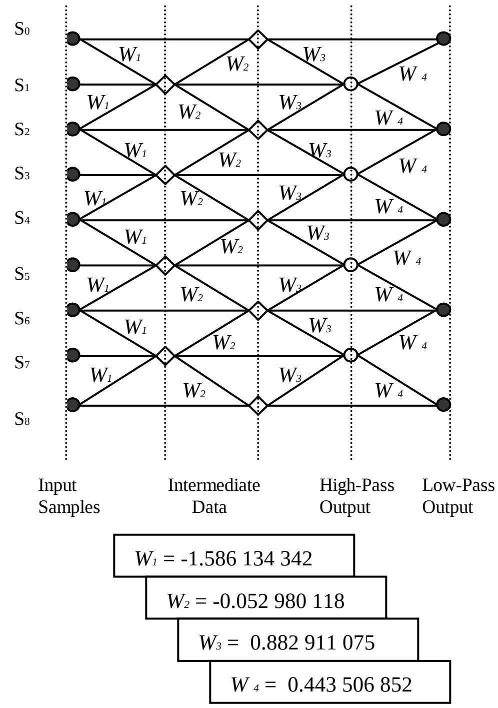


Figure 2. Lifted 9/7 Forward Wavelet Transform

B. Uniform Scalar Quantization

The decomposition wavelet coefficients in each subband are quantized by using a uniform scalar quantizer with a central dead-zone. The image samples are assumed to follow a Laplacian probability distribution. The uniform scalar dead-zone quantization is suitable for this feature of image sample distribution based on the requirement of the rate-distortion optimization. The width of the dead-zone is twice the quantization step size. The quantizer step size for each subband is decided based on the contribution of the quantization noise in the subband to the overall mean-squared error (MSE) . The more the quntization noise in the subband contributes to the overall MSE, the smaller the quantizer step size is for quantizing the decomposition coefficients in that subband. The quantizer scales the decomposition coefficients to quantization index values.

C. Data Structure for Facilitating Column Random Access

The quantized decomposition coefficients in each subband are partitioned into code-blocks. These code-blocks are organized into a hierarchical structure for achieving efficient coding performance and facilitating the required random access. A uniform size is chosen for code-blocks in all subbands at all resolution levels. A certain number of horizontally connected code-blocks in the same subband constitute a code-block set. The width of code-block sets is same in all subbands at the same resolution level. Code-block sets at different resolution levels are allowed to involve different number of code-blocks. Code-block sets with the same horizontal position in the same subband constitute a block-set cluster. A cluster group consists of all the block-set clusters with the same horizontal position in all subbands at the same resolution level. Bigger sizes of code-blocks result in a more efficient representation of decomposition coefficients, but constraint the flexibility of random access to partial image data. The selection of the code block size is a tradeoff balancing the coding efficiency and the random access flexibility. The choice of the code-block set size is usually aimed at making the block-set clusters at different resolution levels corresponding to the same size in the original image to easily facilitate partial spatial image random access at the image column level.

D. Embedded Entropy Block Coding

An embedded entropy encoder is used to independently generate a progressive bit stream for each code-block. Subbands are independently encoded from each other. Although some coding efficiency is lost because no inter-subband redundancy is exploited, this increases the flexibility to combine the bit streams of code-blocks in different orders in the code stream according to the requirement of applications. Binary arithmetic bitplane coding is applied in the embedded entropy encoder. The binary symbols in each bitplane are encoded together with their probability estimations. The probability estimation of each binary symbol is adaptively updated according to its history and present state. Three sub-bitplane passes are involved in the encoding process of each bitplane. These three passes are performed in a fixed order according to their contributions to the distortion reduction. The multiple-pass bitplane encoding accommodates more efficient and more accurate bit rate control. Truncating the code stream to control the rate is allowed at the boundaries of the sub-bitplane passes. Encoding one block results in one arithmetic codeword. Subband code-block coding creates embedded bit streams and assists the localized random access and scalable coding.

E. Code Stream Syntax

All bit streams from the code-blocks are assembled to form a single output code stream. The code stream syntax is constructed to efficiently organize bit streams based on designed data structure. The code stream of an image dataset starts with an image dataset header followed by coded data of basis images in the image dataset. The image dataset header involves information of the image dataset, the camera information, global encoding parameters and the setup information for starting the decoding process. An index table indicating the start points of each basis image in the code stream is also involved in the image dataset header. There is an image header for each basis image in the coded image data part of the code stream. The coding parameters for the whole basis image are put in the image header. Also, an index table indicating the start points of all the cluster groups in each basis image is involved in the image header. The coded image data consisting of arithmetic codewords from all the coded-blocks in a basis image follows the image header. The two-level index tables are used to assist image column random access.

F. Random Access and Selective Decoding

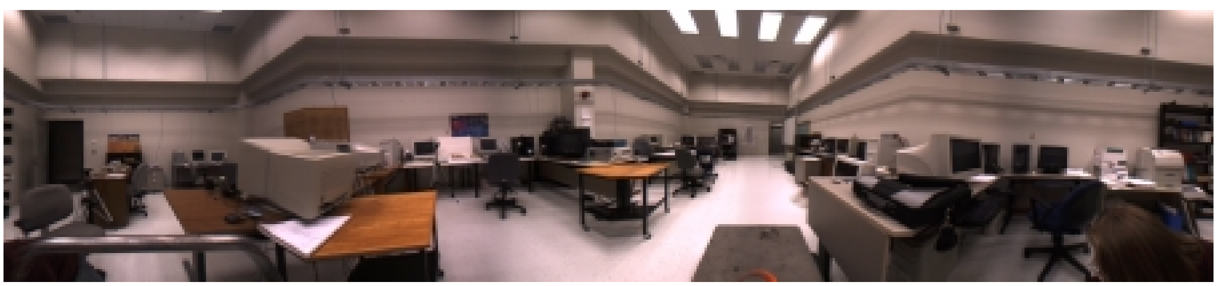
A mechanism of image-column random access to the image dataset in the code stream is accommodated in the scheme. A tradeoff between the compression efficiency and the random access flexibility is reached. In order to decode an image column at a certain required resolution level, all the relative cluster groups at all lower resolution levels need to be accessed and decoded. For the bi-orthogonal Daubechies 9/7 wavelet kernel adopted in this scheme, five columns of high-pass decomposition coefficients and four columns of low-pass decomposition coefficients at one level lower than the required resolution level need to be accessed and decoded to reconstruct one column image at the required resolution level. For each column of coefficients at the lower resolution level, five columns of high-pass coefficients and four columns of low-pass coefficients at one level further lower need to be reconstructed. This process continues until the lowest resolution level is reached. Usually the maximum number of resolution levels involved in such a process is 5 or 6. It is important to correctly locate to relevant columns in the cluster groups at different resolution levels to reconstruct the image column at the required level. This is the key consideration in designing the data structure and choosing the size of different structure elements in this scheme. The total number of accessed coefficient columns is bigger in the case that the constructed columns are on or near the cluster boundaries.

III. SIMULATION RESULTS

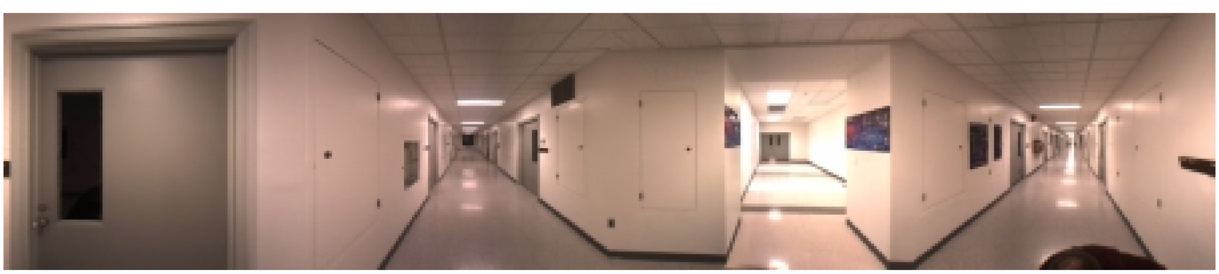
The proposed image dataset compression scheme is implemented in a software codec. The sample image datasets used for testing the image dataset compression scheme and obtaining simulation results are two cubic panorama image datasets: Lab and Corridor. Each cubic panorama image consists of six images corresponding to six sides of a cube. The cubic panorama image datasets can be used to synthesize novel views in image-based virtual environment navigation based on specially developed rendering techniques. From image dataset compression point of view, we connect the six images in a cubic panorama to form a rectangular image. It is the combination of a top view, a bottom view and a horizontal view. After the conversion, the Lab image dataset consists of 128 RGB images of resolution 3072x512, and the Corridor image dataset consists of 912 RGB images of resolution 3072x512. The horizontal views of the Lab and Corridor image datasets are shown in Fig. 3. The simulation results of the proposed image dataset compression scheme are obtained by working on these two image datasets.

Currently, the majority of image acquisition equipments use JPEG format to record the captured images in compressed form in the storage media, and JPEG 2000 as the new image compression standard is expected to be widely applied in the future. Therefore, the results of JPEG and JPEG 2000 used for compressing the sample image datasets are also provided to be compared with the results of the proposed image dataset compression scheme.

Table 1 shows the results of different compression schemes used for coding the Lab image dataset. The coding performances are compared by setting the code stream rates (bits/pixel) of different schemes at the same value and obtaining average peak signal-to-noise ratio (PSNR) results for comparison. The results show that the proposed lifted wavelet-based image dataset compression scheme denoted by WIDC in the tables outperforms the other two schemes. Especially, it provides a significant performance im-



(A) The Lab Image Dataset



(B) The Corridor Image Dataset

Figure 3. Horizontal Views of the Sample Image Datasets in Use

provement compared with the JPEG scheme. Table 2 shows the results of coding the Corridor image dataset. The range of the code stream rates is extended at both the lower rate end and the higher rate end. Again the results show that the proposed scheme provides better PSNR performance than that of the other two schemes besides supporting the random access mechanism for image dataset compression.

Table 1. Compression Performance Comparison for the Lab Image Dataset

Rates (bits/pix.)	Average PSNR in dB		
	JPEG	JPG2000	WIDC
1.23 bpp	41.23	44.30	44.55
0.69 bpp	38.63	41.43	41.64
0.48 bpp	36.85	39.69	39.82

Table 2. Compression Performance Comparison for the Corridor Image Dataset

Rates (bits/pix.)	Average PSNR in dB		
	JPEG	JPG2000	WIDC
1.41 bpp	46.01	47.20	47.41
0.33 bpp	40.73	42.65	42.79
0.23 bpp	38.42	41.76	41.92

IV. CONCLUSIONS

In this paper, a lifted wavelet-based compression scheme with effective image-column random access for coding image datasets is proposed. The image dataset compression scheme follows a JPEG 2000-like approach based on the framework of a lifted wavelet transform combined with embedded block entropy coding. It features a proper compromise between the coding efficiency and the random access flexibility, and thus ensures real-time interactive image rendering. Although some coding efficiency is sacrificed because no inter-image redundancy is exploited with independent intra-image coding, it reduces the decoding delay, simplifies the implementation of a random access mechanism and easily facilitates real-time image decoding. The proposed compression scheme is significantly less complex than some other schemes because no motion estimation and compensation are needed. It can satisfy to a large extent almost all the requirements for image dataset compression except for extremely high compression ratios. Experimental results of the software implementation of this image dataset compression scheme applied to compressing sample image datasets showed that this proposed scheme achieves superior coding performance over the standard JPEG and JPEG 2000 schemes.

ACKNOWLEDGEMENTS

We wish to acknowledge the funding from LORNET, an NSERC research network, and would like to thank Alan Brunton and Mark Fiala for providing the sample image datasets.

REFERENCES

[1] D. G. Aliaga and I. Carlbom, “Plenoptic Stitching: A Scalable Method for Reconstructing 3D Interactive Walkthroughs,” *Proc. ACM Annu. Conf. Computer Graphics and Interactive Techniques (SIGGRAPH’01)*, Los Angeles, CA, USA, Aug. 2001, pp. 443-450.

[2] S. E. Chen, “QuickTime VR – An Image-Based Approach to Virtual Environment Navigation,” *Proc. ACM Annu. Conf. Computer Graphics and Interactive Techniques (SIGGRAPH’95)*, Aug. 1995, pp. 29-38.

[3] B. Girod, C.-L. Chang, P. Ramanathan and X. Zhu, “Light Field Compression Using Disparity-Compensated Lifting,” *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP’03)*, Hong Kong, China, Apr. 2003, pp. IV760-763.

[4] ISO/IEC JTC1/SC29/WG1, “Information Technology --- JPEG 2000 Image Coding System, Part 1: Core Coding System,” ISO/IEC 15444-1 and ITU-T Rec. T.800 (JPEG2000), Aug. 2000.

[5] L. Luo, Y. Wu, J. Li and Y.-Q. Zhang, “Compression of Concentric Mosaic Scenery with Alignment and 3D Wavelet Transform,” *Proc. SPIE Image and Video Communications and Processing Symp.*, San Jose, CA, USA, Jan. 2000, vol. 3974, pp. 89-100.

[6] A. Said and W. A. Pearlman, “A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees,” *IEEE Trans. Circuits and Syst. for Video Technol.*, vol. 6, no. 3, pp. 243-250, June 1996.

[7] J. M. Shapiro, “Embedded Image Coding Using Zerotrees of Wavelet Coefficients,” *IEEE Trans. Signal Processing*, vol. 41, no. 12, pp. 3445-3462, Dec. 1993.

[8] H.-Y. Shum and L.-W. He, “Rendering with Concentric Mosaics,” *Proc. ACM Annu. Conf. Computer Graphics and Interactive Techniques (SIGGRAPH’99)*, Los Angeles, CA, USA, Aug. 1999, pp. 299-306.

[9] W. Sweldens, “The Lifting Scheme: A Construction of Second Generation Wavelets,” *SIAM J. Math Anal.*, vol. 29, no. 2, pp. 511-546, Mar. 1998

[10] D. Taubman, “High Performance Scalable Image Compression with EBCOT,” *IEEE Trans. Image Processing*, vol. 9, no. 7, pp. 1158-1170, July 2000.

[11] Y. Wu, L. Luo, J. Li and Y.-Q. Zhang, “Rendering of 3D-Wavelet Compressed Concentric Mosaic Scenery with Progressive Inverse Wavelet Synthesis (PIWS),” *Proc. SPIE Visual Communications and Image Processing Symp.*, Perth, Australia, June 2000, vol. 4067, pp. 31-42.

[12] Y. Wu, C. Zhang and J. Li, “Smart Rebinning for the Compression of Concentric Mosaic,” *IEEE Trans. on Multimedia*, vol. 4, no. 3, pp. 332-342, Sept. 2002