SCALABLE CODING AND PROGRESSIVE TRANSMISSION OF CONCENTRIC MOSAIC USING NONLINEAR FILTER BANKS

King-To Ng and Shing-Chow Chan

Department of Electrical and Electronic Engineering
The University of Hong Kong
{ktng, scchan}@eee.hku.hk

Heung-Yeung Shum

Microsoft Research, China

hshum@microsoft.com

ABSTRACT

This paper studies the scalable coding and progressive transmission of concentric mosaic to support interactive applications over LAN or Internet. Concentric mosaic is an effective 3D image based representation of static scene. A typical concentric mosaic might consist of thousands of images, which poses significant problem in digital storage and transmission. A new multiresolution decomposition for supporting progressive transmission of concentric mosaic is proposed. Instead of using the propular 9/7 wavelet filterbank, a nonlinear perfect reconstruction filter bank with lower arithmetic complexity is employed. It also considerably simplifies the random access operation of the slit images during rendering. By encoding the subband signals into different layers, a scalable compressed bit stream of the concentric mosaic is obtained. Therefore progressive transmission of concentric mosaic, using a combination of these layers, to support devices with different capabilities become possible.

1. INTRODUCTION

Image based rendering (IBR) has recently emerged as a simple yet powerful photo-realistic representation of real world scenes [1,3-7]. Its basic principle is to render new views of a scene using rays that were previously captured in densely sampled pictures of the scene. It has plenty of applications including virtual walkthrough, electronic games, medical simulation, visualization and many others, which requires virtual camera motion. IBR is also an excellent alternative to conventional 3D model building if what we want is just to re-render novel views at different viewpoints. It not only provides superior image quality than 3D model building for real world scenes, but also requires much less computational power for rendering, regardless of the scene complexity. Central to IBR is the plenoptic function [2], which describes all of the radiant energy that can be perceived by the observer at any point in space and time. In its most general form, it is a 7-dimensional function allowing one to reconstruct any novel view at any point in space and time. If the viewpoint is fixed and only the viewing directions and camera zoom can be altered, the plenoptic function simply becomes a 2D panorama (cylindrical [5] or spherical [6]). Due to the difficulties in capturing and storing the complete plenoptic function, many simplifications have been proposed. This includes the 4D Light Field [3] and the 4D Lumigraph [4], where the plenoptic function of a static scene under the constraint that the object or the viewer could be constrained within a 3D bounding box is captured.

Recently, Shum and He [7] had proposed a 3D IBR representation of static scene called the concentric mosaics. It captures the plenoptic function when the user is constrained to move inside a planar circle. Due to its reduced dimension, the construction of a concentric mosaic is relatively simple and the amount of data recorded is also significantly less than that required for Light Field and Lumigraph. A simple method for constructing a concentric mosaic is to take a series of images at different angles when a single camera, placed at the end of a round-swinging beam, is rotated in a circle. A typical concentrate mosaic might consist of thousands of images at a resolution of (320×240). For example, the 'Lobby' concentric mosaic [7] consists of 1350 images at a resolution of (320×240), which requires a total storage of 297 mega bytes. Because of this reason, the compression of IBR, and in particular concentric mosaic, has received considerable attention recently.

Different compression schemes were developed to explore the correlation both within and across the image frames of the concentric mosaics [7-10]. Unlike conventional image compression, individual lines of each mosaic image have to be rapidly decoded to support real-time rendering. This requirement usually leads one to employ simple technique such as vector quantization (VQ) with a fixed length code word in compressing concentric mosaic [7] and other IBR representation. Unfortunately, the compression ratio of VQ is usually limited (e.g. VQ in [7] achieves a compression ratio of only 12:1). More recently, the authors have proposed MPEG2-like codec [8] for compressing concentric mosaics, while taking the random access problem into account. A compression of ratio of 65 can be achieved with good image reconstruction quality. A 3D wavelet approach has also been proposed for compressing concentric mosaics [9]. It achieves a better compression ratio at the expense of more complicated decoding procedure.

In this paper, we study the scalable coding and progressive transmission of concentric mosaic to support interactive applications over LAN or Internet. As mentioned earlier, the amount of storage associated with concentric mosaic can be very large, occupying hundreds of mega bytes. This poses significant problems in supporting interactive applications in bandlimited channel such as the Internet. To avoid sending the whole concentric mosaic, progressive transmission operating in a "on demand" basis can be used to improve the interactive speeds by reducing amount of data to be transmitted, and hence the delay associated with transmission and decoding. A very popular and effective method for supporting progressive transmission of images and video is by means of multiresolution signal decomposition. In multiresolution signal decomposition [11], the signal to be encoded is decomposed say by a perfect reconstruction filter bank into a series of subband signals with different resolution or importance. These subband signals are then encoded to form either an embedded bitstream or layers with increasing importance or resolution. In this paper, we propose a new multiresolution decomposition for concentric mosaic using a nonlinear perfect reconstruction filter bank. The main reason of using a nonlinear perfect reconstruction filter bank, instead of the commonly used 9/7 wavelet filterbank [12], is that its arithmetic complexity is very low and it considerably simplifies the random access operation of the slit images during rendering, as we shall see later in Section 3. Interested readers are referred to [16,17] for applying 9/7 wavelet filter to multiresolution video coding and their performances. By encoding the subband signals into different layers, a scalable compressed bit stream of the concentric mosaic is obtained. Therefore progressive transmission of concentric mosaic, using combinations of these layers, to support devices with different capabilities becomes possible.

The paper is organized as follows. A brief introduction to concentric mosaics and its rendering scheme are introduced in Section 2. The multiresolution decomposition of concentric mosaic using the nonlinear filter bank is presented in Section 3. Experimental results are presented in Section 4. Finally, we summarize our works in Section 5, the conclusion.

2. CONCENTRIC MOSAICS

2.1. Capturing and Rendering a Novel View

Concentric mosaics are a set of manifold mosaics [14] constructed from slit images taken by cameras rotating on concentric circles. Figure 1 shows how a mosaic image, CM_k, is obtained by recording the slit images, r_k, that are tangent to the circle CO_k, when a camera is moving along the circle. A simpler method, called the normal setup, is to capture an image, instead of a slit image, using an outward facing camera when it is moving along a larger circle, CO_n. It can be seen that rays of the inner circles, say r_k, are captured as part of these images. This considerably simplifies the capturing process because only one circular motion is necessary. The entire concentric mosaic is indexed by three parameters: the radius of the mosaic, the rotational angle of the camera, and the vertical elevation of the slit image. Therefore, it is very efficient to retrieve the appropriate slit images from the mosaic images to construct a novel view at any position inside the circle. Consider the rendering of a novel view at a point P with a polar co-ordinate (R,θ) measured from the center of the CM as shown in Figure 2. The ray $\overline{PV_i}$ is not captured at the novel view point P. Since the circular region is a free space, we can use the ray previous captured at point v_i in the concentric mosaic CM_k. Similarly, the ray $\overline{PV_j}$ can be retrieved from the point v_j in the concentric mosaic CM_j. Therefore, the novel view at P can be completely constructed from the concentric mosaics. In practice, however, only a small subset of the rays is stored in the concentric mosaics. For those rays that are not recorded in the concentric mosaics, they have to be approximated from adjacent ones that have previously been recorded. Interested readers are referred to [7] for more details.

2.2. Compression and the Random Access Problem

As mentioned earlier, concentric mosaic can be compressed by a MPEG2-like codec supporting random access to individual slit image in a mosaic picture [8]. This is illustrated in Figure 3, showing how random access to a slit image can be performed. The mosaic image is divided into non-overlapping blocks of size (16×16 macroblock). Here, the blocks are scanned vertically so that pixel data of each vertical line are contained in a group of consecutive blocks. In order to retrieve the pixel data of a line, only the compressed data of the blocks containing that line have to be located and decoded. Locating this data from the compressed data stream by searching the headers of the blocks can be very time consuming, especially for real-time rendering. To overcome this problem, a set of pointers to the starting locations of the vertical group of blocks in the compressed data is first determined and stored in an array prior to rendering. In performing a multiresolution decomposition of concentric mosaic, one must also take into account this random access problem. Otherwise, the decoding speed can be very slow.

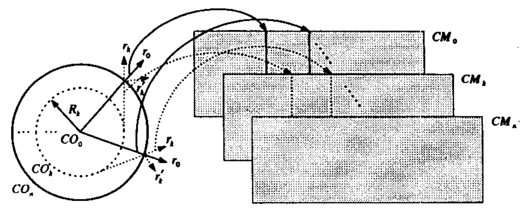

Figure 1. Construction of concentric mosaic from one circle: along normal.

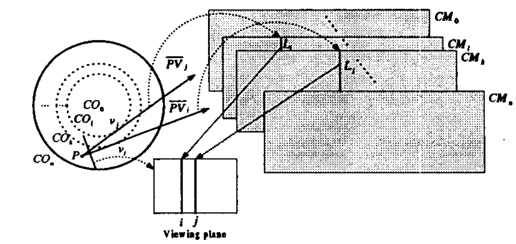

Figure 2. Rendering a novel view with concentric mosaic.

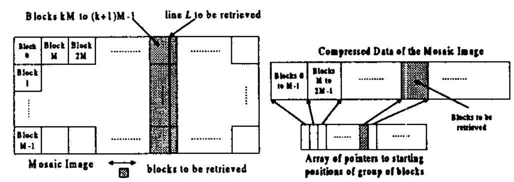

Figure 3. Random access for each line in a mosaic image.

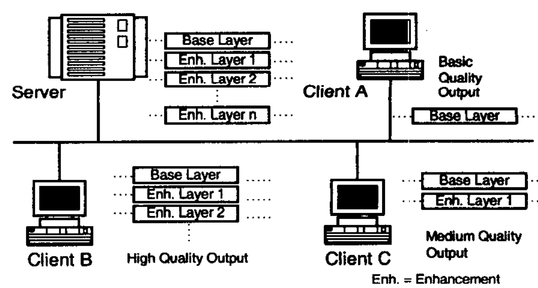

Figure 4. Scalable bitstreams for supporting users with different capabilities.

3. MULTIRESOLUTION REPRESENTATION OF CONCENTRIC MOSAIC

3.1 Progressive Transmission

In order to deliver the interactive virtual walkthrough experience offered by concentric mosaics, the compressed data might be transmitted to the users through the Internet or LAN. In applications where the channel has limited or rapidly changing bandwidth such as Internet, it is undesireable to keep user waiting for a long time while transmitting the whole data stream. By using multiresolution technique, we can decompose the mosaic into different resolutions and encode them into different layers for transmission. Figure 4 illustrates this concept in a LAN environment. The base layer consisting of the lowest resolution of the mosaics and hence the smallest data size can be transmitted very quickly to the receiver to render a low quality novel view. Hence the user can start the application as soon as possible. For those terminals or users with limited capabilities, they can still make use of the base layer for rendering without any difficulties. Due to the division of the compressed data into different layers, the upper layers can be transmitted progressively afterwards. The quality of the mosaic is refined during the running time until eventually the highest resolution or quality is reached.

3.2 Multiresolution Decomposition Using Nonlinear Filter Banks

In this paper, we propose a new multiresolution decomposition for concentric mosaic using a nonlinear perfect reconstruction filter bank. The main reason of using a nonlinear perfect reconstruction filter bank, instead of the commonly used 9/7 wavelet filterbank [12] is that it has a lower arithmetic complexity and lower support so that random access of the slit images during rendering can be considerably simplified. Figure 5 shows one level of decomposition using this non-linear filter bank [13]. For simplicity, a notation similar to the conventional filter bank systems is adopted, namely LL, LH, HL and HH bands. The low-low (LL) band is simply obtained by downsampling the original image by a factor of two both horizontally and vertically, and it consists of pixels at positions $(2i, 2j)$. The HH band consists of prediction residuals, $e(2i+1, 2j+1)$, when pixels at positions $(2i+1, 2j+1)$ are predicted from its four quantized neighbours at positions $(2i, 2j)$, $(2i+2, 2j)$, $(2i, 2j+2)$, and $(2i+2, 2j+2)$. Similarly, HL and LH bands consist of prediction residuals when the pixels at $(2i+1, 2j)$ and $(2i, 2j+1)$ are predicted from their respective neighbours. Different prediction schemes can be used in this simple scheme. The following is a simple but effective scheme which is adopted from [13] : For 4 input samples $\{p_1, p_2, p_3, p_4\}$, discard the minimum and maximum of $\{p_1, p_2, p_3, p_4\}$. The output is equal to the mean of the remaining two elements. This operation though very simple is able to enhance image edges using only the subsampled image data. It can also be viewed as a post-processing operation for improving the visual quality of the reconstructed image. The operation is particular simple which involves only 4 comparsions and 2 additions per pixel. Similar decomposition can be applied to the LL band creating a multiresolution decomposition of the image, like an octave wavelet filter bank. Another reason for using this non-linear filter banks is that the reconstruction of a given pixel at a given resolution only requires its four adjacent neighbors, unlike their conventional linear counterparts. This property greatly simplifies the rendering of concentric mosaic because only very few additional samples are needed to reconstruct a given pixel. Furthemore, the decomposition, as we have seen previously, does not require any multiplications or floating point numbers.

In this paper, the base layer sequences is chosen as the LL band of the non-linear filter bank which can be compressed using the MPEG2 based encoder that we have described in Sect. 2.2. We use one I-frame and seven B-frames in a Group of Picture as in our previous work [8]. For simplicity, we shall consider one level of decomposition here. Generalization to include more levels, to increase the scalability of the data streams, can be done similarly. The HL, LH, and HH bands are then compressed by a MPEG2 like coder, treating each of them as the residuals of typical P-frames. Each of them will form an

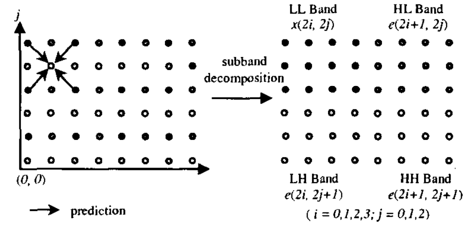

Figure 5. One level of decomposition using the nonlinear filter bank. Samples marked with • are used to predict the samples marked with • , ◦ and • . e(• , •) are the prediction residuals.

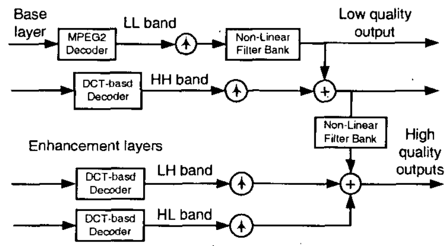

Figure 6. Decoder of the proposed multiresolution schemes for concentric mosaics using the non-linear filter bank.

Figure 7. A typical images of the decompressed mosaics sequences with base layer only. (Luminance component)

additional enhancement layer for progressive transmission. Figure 6 shows the decoder of the proposed scalable coder for concentric mosaics using the non-linear filter banks. Another point worth mentioning for the non-linear decomposition is that it can significantly improve the performance of the coder at very low compression ratio. It is because at very low compression rate, all the high frequency subbands will be dropped in order to reduce the information to be encoded as well as those overheads (such as headers) associated with commonly used video coding standards. Therefore, it is possible to make better use of the available bits to achieve a more reasonable image quality even at very low bit-rate. Those users who have more resources or bandwidth can decode all the enhancement layers to obtain a high quality output.

4. EXPERIMENTAL RESULTS

The "lobby" concentric mosaic described in previous sections is compressed using the proposed scalable coding algorithm. For simplicity, only one level of decomposition is used in our coder. As all the mosaic images will be used to render the novel views, it is important to ensure an acceptable quality throughout the images. Because of this reason, no rate control algorithm is applied and a uniform quantizer with a scale factor Q = 8 and 12 is used in the MPEG2 encoder for both the base and the enhancement layers. Table 1 shows the compression performances of the different layers. For Q = 12, the compressed data size of the base layer is only 2.144 MB while the mean PSNR of the reconstructed mosaic images is 33.23dB (Luminance component in quarter resolution). In the decoder, Figure 6, the decoded low resolution image is interpolated using the simple reconstruction scheme in section 3.1. Figure 7 shows a typical reconstructed mosaic images using the base layer and nonlinear interpolation. Although the mean PSNR of the mosaics images is only 26.37 dB, it can be seen that the visual quality is quite good and its compression ratio is 138.5. The edges are fairly sharp and there are no significant ringing artifacts which is frequently encountered in linear interpolation with long interpolation filters. It also found that the possible aliasing is not visually disturbing. Other more sophisticated post-processing schemes can be employed to further improve the image quality. Also, more levels of decomposition can be used to generate even lower resolution concentric mosaic for severely bandlimited channels.

5. CONCLUSION

A new multiresolution decomposition for supporting progressive transmission of concentric mosaic is presented. Instead of using the propular 9/7 wavelet filterbank, a nonlinear perfect reconstruction filter bank with lower arithmetic complexity is employed. Furthermore, due to the shorter support of the nonlinear filterbank, the random access of the slit images required for rendering is greatly simplified. By encoding the subband signals into different layers, a scalable compressed bit stream of the concentric mosaic is obtained. This can be used together with progressive transmission to support devices with different capabilities in viewing concentric mosaic over bandlimited channel.

ACKNOWLEDGMENT

This work is supported by the Hong Kong Research Grants Council, Area of Excellent in Information Technology (AOE IT) project.

Layers	Data Size (MB)	Mean PSNR (dB)		
		Y	U	V
Base (Low Resolution)	2.144	33.23	40.54	39.66
Base	-	26.37	26.63	26.71
HH	2.531	30.71	40.38	39.54
HL&LH	2.682	31.59	40.43	39.59

(a) Uniform quantizer Q = 12 for all layers

Layers	Data Size (MB)	Mean PSNR (dB)		
		Y	U	V
Base (Low Resolution)	3.222	34.71	41.73	40.74
Base	-	26.51	26.67	26.75
HH	3.900	31.31	41.51	40.58
HL&LH	4.436	32.44	41.59	40.68

(b) Uniform quantizer Q = 8 for all layers

Table 1. Compression performance of different layers.

REFERENCES

[1] L. McMillan and G. Bishop, "Plenoptic modeling: An image-based rendering system," in *Computer Graphics (SIGGRAPH'95)*, pp.39-46.
[2] E. H. Adelson and J. Bergen, "The plenoptic function and the elements of early vision," in *Computational Models of Visual Processing*, pages 3-20. MIT Press, Cambridge, MA, 1991.
[3] M. Levoy and P. Hanrahan, "Light field rendering," in *Computer Graphics Processings (SIGGRAPH'96)*, pp. 31-42, August 1996.
[4] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen, "The lumigraph," in *Computer Graphics (SIGGRAPH'96)*, pp. 43-54.
[5] S. E. Chen, "QuickTime VR – an image-based approach to virtual environment navigation," in *Computer Graphics (SIGGRAPH'95)*, pp. 29-38, August 1995.
[6] R. Szeliski and H. Y. Shum, "Creating full view panoramic image mosaics and texture-mapped models," in *Computer Graphics (SIGGRAPH'97)*, pp.251-258, August 1997.
[7] H. Y. Shum and L. W. He, "Rendering with Concentric Mosaics," in *Computer Graphics (SIGGRAPH'97)*, pp. 299-306, August 1999.
[8] H. Y. Shum, K. T. Ng and S. C. Chan, "Virtual Reality Using The Concentric Mosaic: Construction, Rendering And Data Compression," in *Proc. ICIP*, Vol. 3, pp. 644-647, 2000.
[9] L. Luo, Y. Wu, J. Li, and Y.Q. Zhang, "Compression of concentric mosaic scenery with alignment and 3D wavelet transform," *SPIE Image and Video Communications and Processing*, SPIE 3974-10, San Jose, CA, Jan. 2000.
[10] W. H. Leung and T. Chen, "Compression with mosaic prediction for image-based rendering applications," in *Proc. IEEE Intl. Conf. On Multimedia and Expo*, New York, July 2000.
[11] J. M. Shapiro, "Embedded Image Coding Using Zerotrees of Wavelet Coefficients," *IEEE Trans. Signal Processing*, Vol. 41, No. 12, Dec. 1993, pp. 3445-3462.
[12] M. Antonini, et al, "Image coding using wavelet transform," IEEE Trans. on Image Processing, vol. 1, no. 2, April 1992.
[13] R. L. de Queiroz and D. A. F. Florencio, "A non-linear filter bank for image coding," *Proceedings of the 38th Midwest Symposium on Circuits and Systems*, 1995, pp. 1115 –1118, vol. 2.
[14] S. Peleg and J. Herman, "Panoramic mosaics by manifold projection," in *IEEE CVPR'97*, pp. 338-343, June 1997.
[15] ITU-T Rec. H.262|ISO/IEC 13818-2,"Generic Coding of Moving Pictures and Associated Audio Information: Video," Nov. 1994.
[16] K. T. Ng, S. C. Chan and T. S. Ng, "A multi-resolution two-layer video codec for networking applications," in *Proc. 1996 Third Int. Conf. on Signal Processing*, Beijing, Oct. 1996, pp. 1071-1074.
[17] K. T. Ng, S. C. Chan and T. S. Ng, "An adaptive multiresolution modification of H.263 video coding algorithm," in *Proc. First International Conference on Information, Communicatons & Signal Processing* ICICS'97, Singapore, Sept. 1997, Vol I, pp. 288-291.