The Journal of Natural Science,
Pai Chai University, Korea
Volume 5(2):33-38, 1992

멀티미디어 텔레컨퍼런스를 위한 새로운 영상 압축 기술

김 용 호* · 장 종 환**
* 배재대학교 국제산업대학 전자계산학과
** 배재대학교 국제산업대학 국제자원개발학과

A New Image Compression Technique for Multimedia Teleconferences

Yong-Hoh Kim* and Jong-whan Jang**
* *Dept. of Computer Science, Pai Chai University*
** *Dept. of Information and Communication Engineering, Pai Chai University*

텍스처럴 리전의 러프니스와 사람의 시각 시스템의 특성에 기초하여 세그멘테이션을 수행하는, 멀티미디어 텔레컨퍼런스를 위한 새로운 텍스처 세그멘테이션-베이스 영상 코우딩 기술을 제안한다. 세그멘테이션은, 텍스처의 영역이, 지각된 콘스탄트 인텐시티와 스무드 텍스처 및 러프 텍스처의 세가지 텍스처 클래스로 분류되도록 프랙탈 디멘전을 쓰레쉬호울딩하여 이루어진다. 각 세그먼트 바운더리와 각 텍스처 클래스를 위한 효과적인 코우딩 기술을 개발하여 높은 압축률과 좋은 영상 품질을 갖는 영상 코우딩 시스템을 달성하고, 이 기술의 코우딩 효율을 잘 확립된 기술 (디스크릿 코사인 트랜스폼(DCT) 영상 코우딩)의 코우딩 효율과 비교한다.

A new texture segmentation-based image coding technique which performs segmentation based on roughness of textural regions and properties of the human visual system (HVS) is presented for multimedia teleconference. The segmentation is accomplished by thresholding the fractal dimension so that textural regions are classified into three texture classes; perceived constant intensity, smooth texture, and rough texture. An image coding system with high compression and good image quality is achieved by developing an efficient coding technique for each segment boundary and each texture class. We compare the coding efficiency of this technique with that of a well established technique (discrete cosine transform (DCT) image coding).

Keywords : Multimedia, Texture, Segmentation, Fractal, Image

1. Introduction

A multimedia teleconference is an information exchange system for people to conduct meetings effectively without leaving their offices. Sitting in front of their personal computers (PCs) or workstations, conferees can see each other via real-time motion videos on their multimedia PCs or workstation displays, talk and listen to all the conversation via real-time audio, and watch presentation via on-line Electronic Blackboard. The effectiveness of a multimedia teleconference should be evaluated based on how much it can compress particularly image data among multimedia services; data, audios, and images[10]. Images in a multimedia teleconference require storage of a very large number of bits and large transmission bandwidth.

Classical image coding technique [5] solely guided by information theory often produced low compression rates. The main reason is due to the fact that what the human eye sees and how it sees are not taken into account. Techniques attempting to overcome these limitations by using properties of the human visual system have been applied to image coding to achieve high compression with small visual quality loss. One such technique is segmentation-based image coding technique. Unfortunately, there are limitations with existing segmentation-based image compression techniques. The main limitation is due to the fact that the image data have been segmented into regions of constant intensity [1-4]. In highly textured areas, a good representation of the texture requires many small segments. However, in order to get low bit rate, the number of segments must be limited and thus the quality is degraded.

To overcome the texture representation problem in this paper by proposing a methodology for segmenting an image into texturally homogeneous regions with respect to the degree of roughness as perceived by the HVS. The fractal dimension, the expected value, and the just noticeable difference are the measures used to characterize the texture information. It is of interest to compare the coding efficiency of this technique with that of DCT image coding technique. A comparison of the compression ratio(CR)/SNR table for the two techniques is given.

2. A texture segmentation-based image codec

A block diagram of the proposed texture segmentation based image coding system is presented in Figure 1. Figure 1(a) shows a block of the transmitter including three main stages; the preprocessor, the segmenter, and the encoder.

In a segmentation-based image algorithm, the number of segments and the number of bits representing the textures of the segments are directly proportional to the bit rate of the coded image. Thus, the main purpose of the preprocessor which is the first stage of the proposed transmitter in Figure 1(a) is to alter the image in such a way that fewer segments and textures are produced by the segmenter, but without degrading the visual quality of the segmented image. After preprocessing, the image data is segmented into texturally homogeneous regions with respect to the degree of roughness as perceived by the HVS. The segmentation is accomplished by thresholding the fractal dimension. The last stage in the transmitter is the mixed encoding of the segments of each class and their boudaries.

For boundary coding, accurate representation of the boundary is necessary to describe the location of the region boundary because of the HVS sensitivity of the edges. We choose an errorless coding scheme to represent the boundaries. A binary image representing the boundaries is created. Then, the binary data is encoded using an adaptive arithmetic code since it has been found to be superior to Huffman code, runlength code ,and crack code.

For regions which belong to perceived constant intensity, only the mean intensity values need to be transmitted to describe the textures of the regions. In this case, lossy compression has already taken place since we are approximating each region texture with a constant value. We do not wish to introduce any further compression so a lossless adaptive arithmetic code is again employed to achieve further compression. Since a mean intensity requires 8 bits, the mean values are converted into an $8 \times N$ binary array, where N is the number of segments belonging to perceived constant regions.

Regions belonging to smooth texture and rough texture are not directly encoded. To get higher compression, these regions are modeled first using the 1-D first order polynomial functions. The coefficients of the polynomial functions are encoded because the variance of the coefficients is less than that of the original data. An adaptive arithmetic code is used to encode the coefficients.

At the receiver given Figure 1(b), two types of coded information come into the mixed decoder, boundary information and region texture information. The boundary decoder must regenerate the boundaries for the decoded image. The region decoder must fill in the missing texture information within each region.

3. A discrete cosine transform image codec

To compare this texture segmentation-based image coder with classical image coding techniques, two different types of imagery are coded by the proposed algorithm and transform coding. The first is a head and shoulder image. This image is typical of teleconferencing applications and one which the previously proposed segmentation-based compression techniques are best suited. The second is a complex and natural outdoor image with many edges and highly textured areas. The existing segmentation-based compression techniques[5-7] do not work well for the second image.

The Discrete Cosine Transform (DCT) was used in these experiments rather than the Discete Fourier Transform (DFT) or any of the many other transforms that have been investigated for image coding[8-9]. The advantage of the DCT is that it is easy to evaluate because it does not use complex-number arithmetic, there are fast algorithms for calculating it, and it is more efficient at signal decorrelation than most other transforms.

A block diagram of the DCT encoder and decoder is given in Figure 2. In the encoding process given in Figure 2(a) the input image is processed in blocks, the size of which is a parameter that must be selected. As the block size is increased the transform packs the information more efficiently, but, in the algorithm used here, the overhead of transmitting the side information also increases. Thus, there will be an optimum block size at a given rate. However, the choice of block size is more likely to be determined by the complexity of calculating the transforms, and it is almost certain to be a power of two.

Each block is transformed by the forward DCT (FDCT) into a set of values referred to as DCT coefficients. One of the values is referred to as the DC coefficent and the others as the AC coefficients. Each of the coefficients is then quantized using one of corresponding values from a quantization table determined by one of the table specifications shown in Figure 2(a). No default value for quantization table is specified in this specification; applications may specify values which customize picture quality for their particular image characteristics, display devices, and viewing conditions.

After quantization, the DC coefficient and the AC coefficients are prepared for entropy encoding. The previous quantized DC coefficient is used to predict the current quantized DC coefficient, and the difference is encoded. The quantized AC coefficients undergo no such differential encoding, but are converted into a one-dimensional zig-zag sequence.

The quantized coefficients are then passed to an entropy encoding procedure which compresses the data further. One of two entropy coding procedures can be used, as described. If Huffman encoding is used, Huffman table specifications must be provided to the encoder. If arithmetic encoding is used, arithmetic code conditioning table specifications must be provided. In this paper, Huffman table is chosen because it is widely used rather than arithmetic coding.

Figure 2(b) shows the main procedures for all DCT-based decoding processes. Each step shown performs essentially the inverse of its corresponding main procedure within the encoder. The entropy decoder decodes the zig-zag sequence of quantized DCT coefficients. After dequantization the DCT coefficients are transformed to a block by the inverse DCT (IDCT).

4. Performance comparison

To compare this texture segmentation-based image coder with DCT image coder, two different types of imagery are coded. The two test images are shown in Figure 2. The first is a head and shoulder image. This image is typical of teleconferencing applications and one which the previously proposed segmentation-based image compression techniques are best suited. The second is a complex and natural outdoor image with many edges and highly textured areas. Each image consists of 256×256 gray levels. The 8×8 block size is used for two techniques. D1 and D2 are used as 2.035 and 2.363 respectively for the class type images. The decoded images of the two test images are shown in Figure 4 using the proposed image coding algorithm. Comparison of CR (compression ratio)/SNR table for the two techniques is given in the table 1. It shows that the proposed texture segmentation-based image coder performs better than the DCT image coder in terms of SNR.

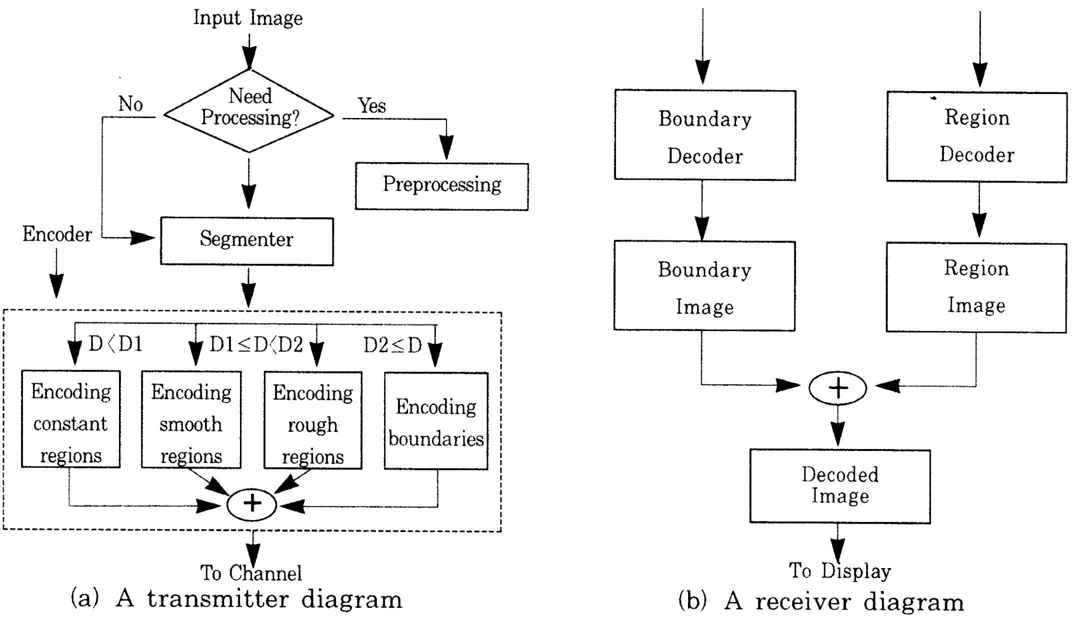

Fig. 1. The proposed texture segmentation-based image coding system

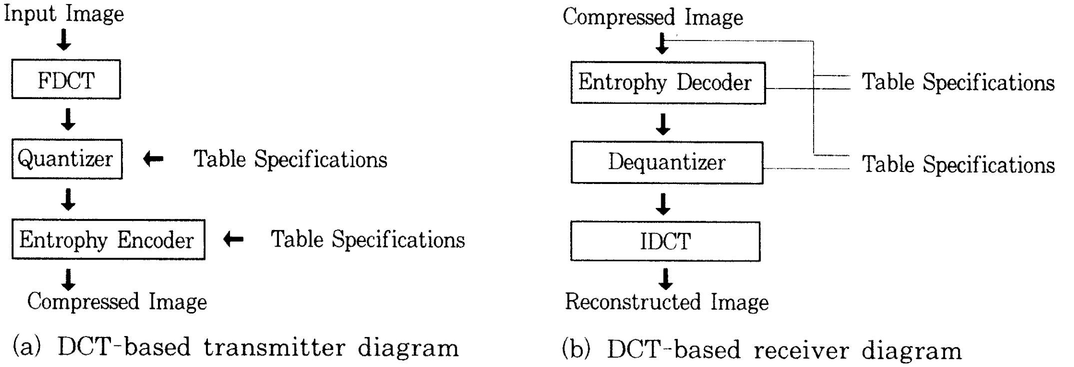

Fig. 2. A DCT-based image coder system

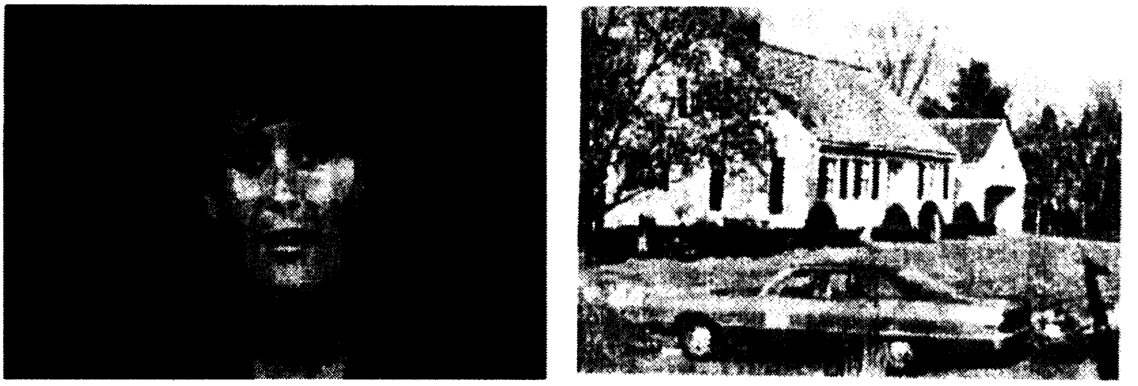

Fig. 3. Original Test Images. Each image is 256 ×256 pixels with 256 gray levels. Miss USA and House are on the left and right respectively.

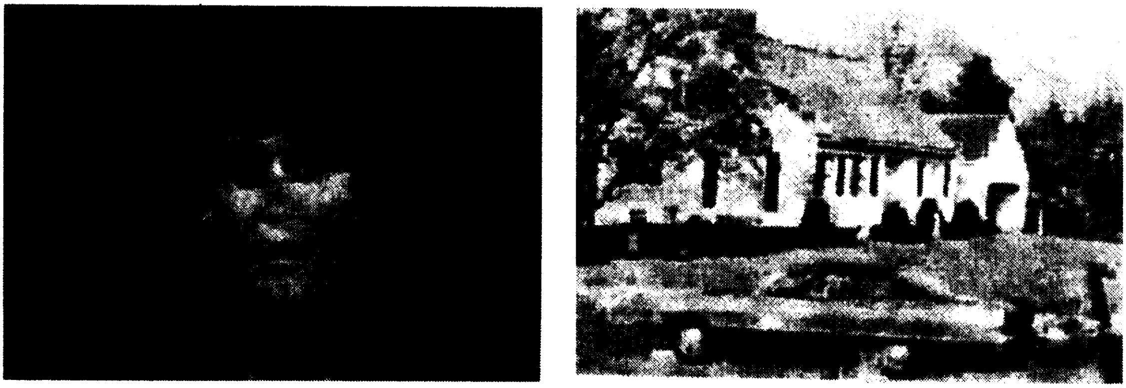

Fig. 4. The decoded images of the test images with D_1 = 2.035 and D_2 = 2.363 The compression ratios of the decoded images of Miss USA and House are 60 and 25 respectively.

CR	Miss USA		House	
	DCT	Proposed	DCT	Proposed
8	26.5	27.3	14.0	15.5
10	25.6	26.4	13.2	14.8
20	24.5	25.9	10.8	12.0
40	19.4	21.2	8.5	9.7
80	15.2	18.1	6.3	8.5
100	13.1	15.2	5.8	6.9

Table 1. Comparison of CR/SNR for the DCT and the proposed image coding system. CR stands for compression ratio.

References

1. M. Barsley, Fractals everywhere, Academic Press, Inc., 1988.
2. M. J. Biggar, O. J. Moris, and Constantinides, Segmented-image coding : Performance comparison with the discrete cosine transform, Proc. IEEE, Part F, Vol. 135 (2), pp. 121-132, April 1988.
3. T. N. Cornsweet, Visual perception, Academic Press, 1971
5. A. E. Jacquin, A novel fractal block-coding technique for digital images. Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing, pp. 2225-2228, April 1990.
6. J. W. Jang and S. A. Rajala, Texture segmentation-based image coder incorporating properties of the human visual system, Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing, May 1991.
7. M. Kunt, A. Ikonmopoulous, and M. Kocher, Second-generation image coding technique, Proc. IEEE, Vol. 73, pp. 549-574, April 1985.
8. M. G. Musmann, P. Pirsch, and H. J. Grallert, Advances in Picture Coding, Proc. IEEE, Vol. 73, pp. 523-548, April 1985.
9. A. N. Netraviali and J. O. Limb, Picture coding: A review, Proc. IEEE, Vol. 68, pp. 366-406, March 1980.
10. Tak-Shing Yum, Mon-Song Chen, and Yiu-Wing Leung, "Bandwidth allocation for multi-media telecon-fences," ICC, pp. 27.5.1-27.5.7, 1991.