논문 2005-42TE-1-7

새로운 멀티플 디스크립션 선택적 부호화 방식

(A new multiple description selective coding scheme)

이 종 배*

(Jong Bae Lee)

요 약

채널의 잡음 영향을 극복하기 위하여 새로운 멀티플 디스크립션 기법을 제안한다. 본 기법은 임베디드 부호화기법을 채택하고 있으며, 부대역 계수가 한번에 비트 플레인 단위로 부호화 되는데 이때 비적응적 산술부호화 방식을 사용한다. 각 영역의 중요도에 따라 관심영역을 부호화할 때, 다른 부분에 비해서 패스 횟수를 늘려서 부호화를 행함으로써 관심영역이 다른 부분에 비해서 고화질로 부호화되도록 한다. 또 채널 에러를 극복하기 위해서 부호화 과정에서 소스 데이터에 제어된 중복성을 허용하는 멀티플 디스크립션 기법을 채택한다. 본 방식은 고압축률을 지향하고, 특정한 영역이 다른 부분에 비해서 중요한 시스템의 경우에 보다 중요한 특성을 나타낸다.

Abstract

A new multiple description selective coding is proposed to overcome noisy channels. Our algorithm adopts an embedded coding scheme, in which subband coefficients are encoded one bit plane at a time using a non-adaptive arithmetic encoder. According to the importance ratio for each region, we code interest regions with more passes than background in order to reconstruct interest regions with higher quality. To overcome channel errors, we adopt multiple description scheme which adds controlled amounts of redundancy to the original data during the compression process. Proposed algorithm achieves better quality compared with other algorithms especially in the circumstances where very low bit rate coding is required and some regions are more important than other regions.

Keywords: Multiple description, selective coding, embedded coding, interest region, noisy channel

I. Introduction

Most of networks used today are based on an exchange of packets of data. When congestion occurs because of data moving from a higher to lower capacity link, the network is often forced to drop some of the packets. When there is no feedback channel that could lead to retransmission of lost packets, the decoder must recover meaningful information from the received packets alone. Recently, the generalized multiple description coding(MDC) frame work has received considerable attention and many new algorithms have been proposed. It can efficiently combat packet loss without any retransmission, thus not only guarantee the demand of real-time services, but also relieve the network congestion. Embedded coding schemes based on MDC[1,2] are the examples. Although they achieve graceful degradation of image quality in the presence of increasing description loss, it improves the overall quality without taking into account the content of image. In some situations, some interest regions(for example, Tank or airplane in a tactical scene or chairperson in video conference system) often attracts more interests than background does. But the algorithms[1,2] do not consider

the importance of interest regions. So a new multiple description selective coding is proposed which differentiates bit rate between interest regions and other regions according to an importance ratio which

* 정회원, 명지전문대학 컴퓨터 정보과
(Computer information and Science Dept. MyongJi College)
접수일자: 2004년11월1일, 수정완료일: 2005년11월29일

defines the amount of importance. Suggested method allocates higher passes to the interest region compared with background, thus improving the image quality especially for the interest region. Suggested method extends embedded scheme to the generalized MDC framework.

Our method uses an embedded coding because it allows for convenient determination of the importance of each data byte. By progressively ordering the data, the embedded coding scheme sends the globally most relevant information first. It produces an embedded bit stream such that its later bits refine earlier. Therefore, the data coded earlier are more important to the image quality the later data. But embedded coding scheme is very susceptible to packet loss regardless of high performance and low complexity. So our method adds multiple description coding concept in order to overcome packet loss in noisy channels. Also this scheme can recover interest regions better image quality because it considers the importance of interest regions.

Ⅱ. Embedded coding scheme based on multiple description coding

Embedded coding scheme is proposed by some researchers[1-3]. The method can stop at any required rate or image quality, which is the feature of embedded coding, so it is suitable for progressive image transmission.

The essential idea of multiple description is to generate multiple independent descriptions of the source such that each description independently describes source with a certain fidelity. When more than one description is available, they can be combined to enhance quality. Each description has the same error sensitivity and the same importance. One of the examples[1] is an embedded scheme which combines multiple coding with the Set Partitioning in Hierarchical Trees(SPIHT) algorithm. It codes SPIHT bit stream into multiple description, and each description has a variable number of fully coded original trees and several sets of partially coded redundant trees corresponding to different spatial locations in the image. Each spatially coded redundant tree is the copy of wavelet coefficient tree of different coding pass. Because coding to a lower pass corresponds to sending only the most significant data, the more important parts of each tree are sent in more descriptions than less important parts. But algorithm[1] does not define interest regions specifically, it can not recover these regions with required quality well.

In order to solve this problem, an embedded coding scheme with MDC considering the importance of regions[4] is proposed. This scheme that is based on multiple description coding with SPIHT algorithm first makes zero trees for an image frame. Next it assigns redundant tree coding rates to all trees according to their relative importance. The more important the tree is, the higher the coding rate of redundant tree assign to it. In order not to increase the bandwidth, the overall source coding and redundancy should be constant. The main idea of the strategy is allocating most redundancy to an interest region(ROI) in order to make strong protection to this region, while allocating lower coding rates to other non-ROI trees according to their relative importance. But embedded coding scheme usually defines quantization step size according to the maximum value of pixels in a special Wavelet level, it is not easy to control many zero tree regions. That is, even in case of non-ROI zero tree regions, some parts are very complex and they need more bits than other parts.

Ⅲ. Suggested scheme

Suggested scheme aims at very low bit ratecoding equal to or less than 0.3 bpp(bit per pixel). In case of some situations such as tactical scenes, channel capacity is very limited and it requires very low bit rate coding scheme. Also in such circumstances, some objects are considered more important than other parts(background). For example, tank or airplane is more important than sky, forest or fields in a tactical scene and a chairperson is more important than a table or a door in a video conferencing system. We also assume that target is relatively small compared with background. There exist many algorithms for object segmentation. Most of current systems only use global features such as overall color histogram and texture moments. Another algorithm that uses relevance feedback was proposed[5] to incrementally supply more information, but they may still fail due to the lack of higher-level information about what exactly was of interest.

Often better capture of image object is achieved by the user rather than the computer. So [6] proposed a method which combined user defined Region-of-interest and spatial layout for content-base image retrieval, which achieves more accurate relevance feedback and leads to a more powerful search engine. Our scheme is also a semi-automatic method[5] which combines image characteristics, and user-defined interest regions. Specifically it first separate interest regions using image threshold and region merging, and finally a user confirms and defines interest regions.

In order to represent ROI boundary in an image frame, we can think about pixel-based approaches or quadtree segmentation scheme[7]. But such methods need too many bits to represent the boundary of a object, and also computational burden is to high to implement our progressive coding scheme. We represented object region with rectangle like in fig. 1 for minimizing redundancy for object indication(four byte for one object). We also separated interest regions from a given frame based on macroblock resolution for implementing hierarchical coding(e.g., Wavelet coding) easily.

In order to model packet loss in the network, we use two-state Markov model[8]. So we adopted multiple description coding for compensating packet loss during data transmission. In order to differentiate the image quality between background and interest regions, We coded ROI and background with Wavelet coding scheme.

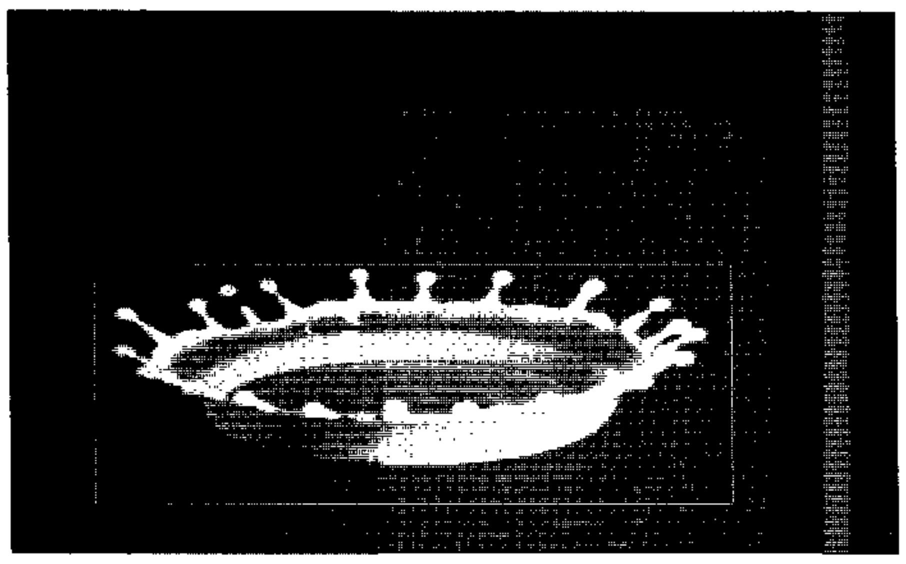


그림 1. Splash 영상과 관심영역
Fig. 1. Splash image and interest region.

Suggested scheme basically follows Progressive wavelet image coding called EPWIC (Embedded Predictive Wavelet Image Coder)[3]. EPWIC uses the two-parameter model of equation to describe each subband and it is based on a separable QMF decomposition using 9-tap symmetric filters[9].

The encoded bit stream begins with a header containing the dimensions of the image frame(in case of ROI, the start position, width and height) and the number of pyramid levels. After quantizing coefficients, we compose bitplanes with them. We ordered subband planes according to a greedy algorithm, which chooses the bitplane with the highest ratio of encoded variance to encoded size(in bits). The sign bit for each coefficient is sent just after the fist "on" bits, as in [10]. For each subband, the parameters (s,p) and (σ, n) defined in [3] are estimated and encoded as 8-bit quantities. They are next sent in the encoded bit stream just before the first bitplane of their corresponding subbands.

Using the four encoded parameters, the receiver constructs a numerical table of the joint probability density for each parent-child pair[3]. For each child bit, we calculate the probability that it is one by numerically integrating over the appropriate rectangular region of the joint distribution. A non-adaptive arithmetic encoder[11], which takes advantage of the computed probability values, is then used to encode the actual bit values. Finally, the arithmetic encoding of the bitplane, and the smaller

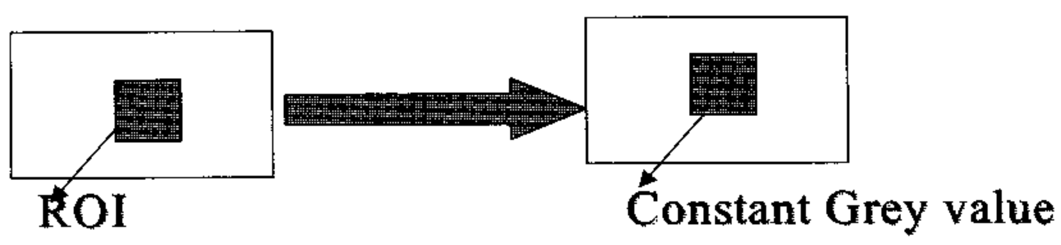


그림 2. 중복패킷에서 배경부호화를 위한 그레이레벨
Fig. 2. Grey level value for Background coding in a redundant packet.

representation is inserted into the encoded bit stream. We assume maximum level depth of frame is 5 and in case of interest regions, it can be reduced according to the minimum block size. We first coded the entire image frame and next interest regions separately.

In case of a redundant packet, the first half of it is used for background and the remainder of it is used for interest regions. When we coded background, we changed interest regions to constant grey value like in Fig. 2 for further data reduction.

Ⅳ. Channel Model

We use the two-state Markov model[4] to model packet loss in the network. This model is able to capture the dependence between consecutive states. For packet losses, we assume that the transmitted packets are represented by an indicator array $\{S_i\}_{i=1}^{M}$, where $S_i = 1$ indicates that the i_{th} packet has arrived successfully and 0 otherwise. The current state S_i of this stochastic process (shown in fig. 3 depends only on its previous state S_{i-1}. The transition probabilities between these two states are defined as in Eq. (1) and state transition matrix is defined as in Eq. (2).

$$P_{10} = P[S_i = 0 \mid S_{I-1} = 1]$$
$$P_{01} = P[S_i = 1 \mid S_{i-1} = 0] \qquad (1)$$

$$P = \begin{bmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{bmatrix} = \begin{bmatrix} 1-P_{01} & P_{01} \\ P_{10} & P_{11} \end{bmatrix} \qquad (2)$$

The maximum-likelihood (ML) estimators of P_{01} and P_{10} given an indicator array $\{S_i\}$ can be expressed as in Eq. (3).

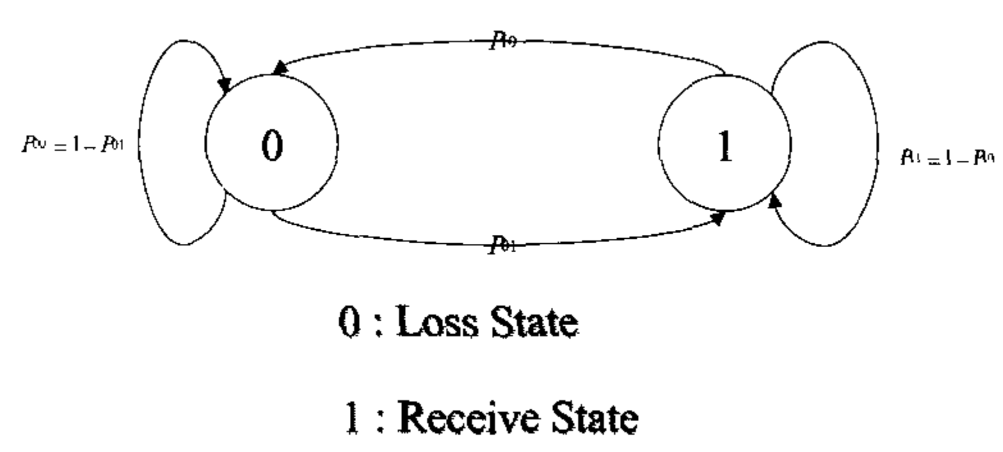


그림 3. 이진 Markov 모델
Fig. 3. two-state Markov model.

$$\hat{P}_{01} = \frac{m_{01}}{m_0},\ \hat{P}_{10} = \frac{m_{10}}{m_1},\ P_L = \frac{\hat{P}_{10}}{(\hat{P}_{01} + \hat{P}_{10})} \qquad (3)$$

where m_0 and m_1 are the number of zeros and ones in the indicator array, m_{01} is the number of times in the indicator array when 1 follows 0, m_{10} is the number of times in the indicator array when 0 follows 1, $m_0 + m_1 = M$, where M is the length of the indicator array, and P_L is the probability that the current state is the loss state.

Ⅴ. Simulation results

We tested proposed method using Splash grey level image(image size is 512 ×512). Actually, existing PSNR is not suitable to evaluate the system performance because it does not consider the importance of special interest regions. So we define a new criterion measure N-PSNR which considers importance ratio M like Eq. (4). N-PSNR can evaluate the performance better than existing PSNR in the circumstances where special regions are more important than other regions.

$$N-PSNR = \frac{(MI+B)\times 255^2}{10\log(M\sum_{i\in I}\sigma_i^2 + \sum_{i\in B}\sigma_i^2)} \qquad (4)$$

where M is the importance ratio, I is the area of interest region, B is that of background and σ_i^2 is mean square error between original pixel and decoded pixel.

If M is 1, N-PSNR is exactly same as the conventional PSNR because importance between interest region and background is same.

Suggested method(Method-3) is compared to three other methods, embedded coding scheme without no multiple description coding(Method-1), embedded coding with packet repetition(Method-2), and original tree + redundant zero trees(Method-4) [4].

Fig. 4 represents simulation result of one frame which uses 4000 bytes(compression ratio is approximately 65 : 1). In this case, our algorithm uses 2000 bytes to total frame, 2000 bytes to interest

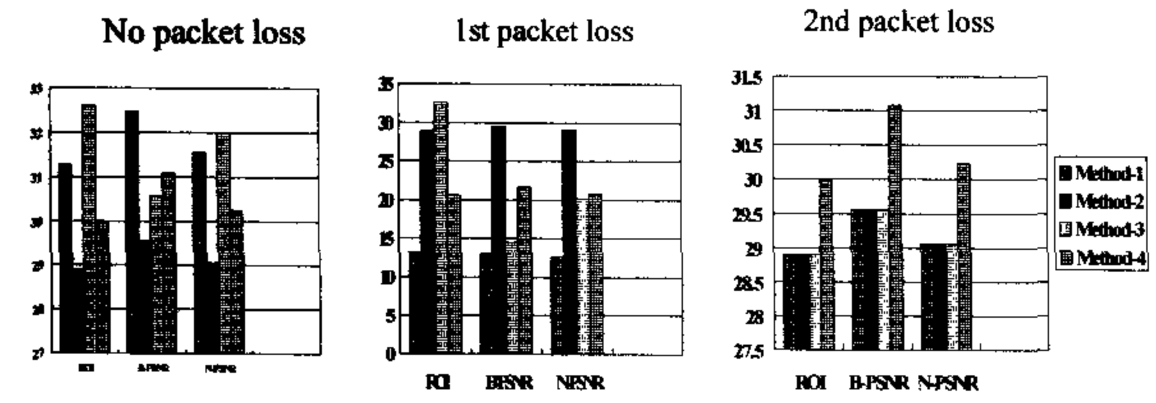


그림 4. 4000 바이트를 사용한 실험결과(단독프레임)
Fig. 4. Simulation result of 1 frame using 4000 bytes.

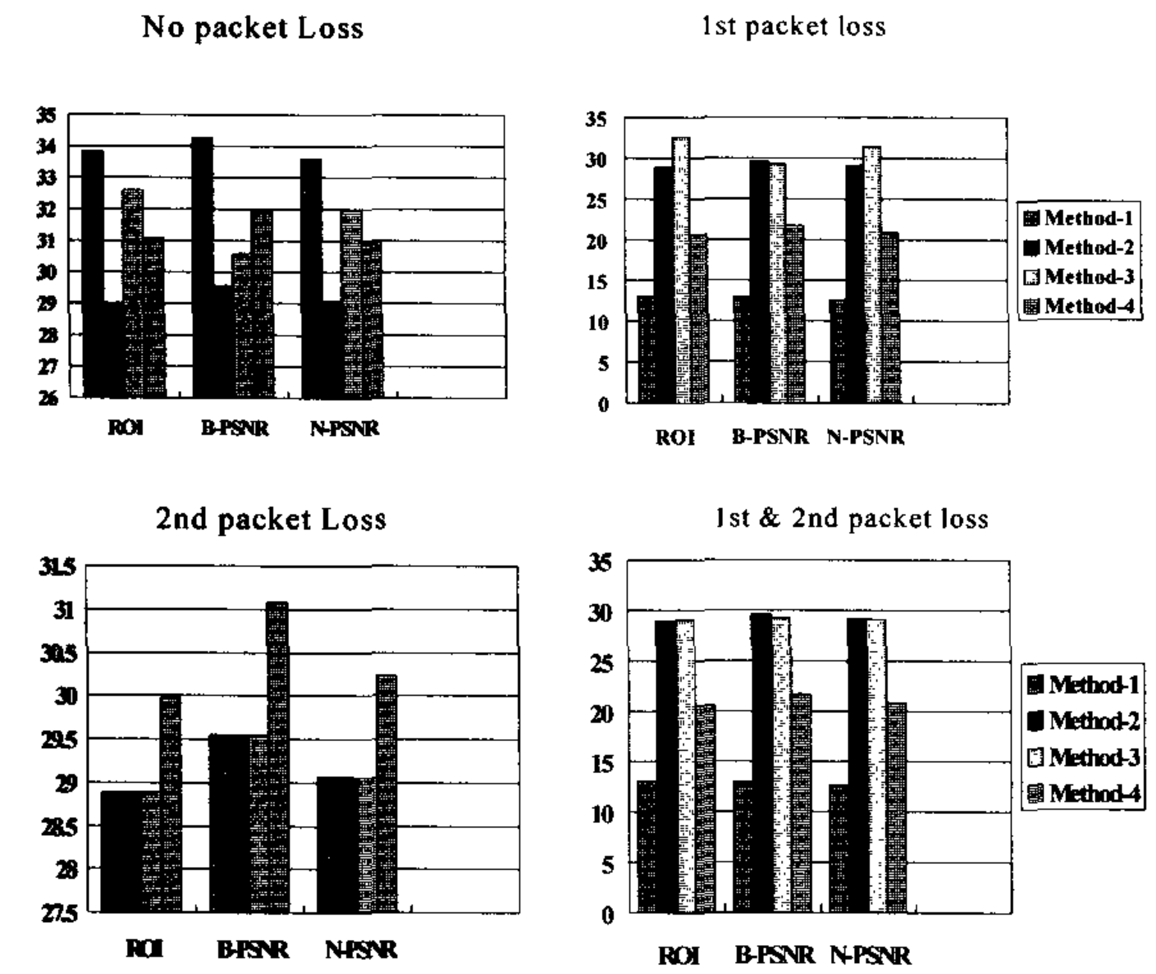


그림 5. 6000 바이트를 사용한 실험결과(단독프레임)
Fig. 5. Simulation result of 1 frame using 6000 bytes.

region(the region surrounded by rectangular region in fig. 1) and M is set to 10. In fig. 4, ROI represent conventional PSNR of interest region, B-PSNR means conventional PSNR. Without packet loss, the performance of our method is better than that of other methods if we adopt N-PSNR. If we lose the first packet, Method-2 performs best because Method-2 uses repetition method, which means one packet exactly represent other packets. If we lose the second packet, the performance of Method-4 is a little better than that of other methods. If our scheme do not consider multiple description scheme, the performance of our scheme does not always better than that of other methods.

Fig. 5 represent simulation result of one frame which uses 6000 bytes(compression ratio is approximately 43 : 1). In this case, our algorithm uses 2000 bytes to total frame, 2000 bytes to interest region and other 2000 bytes to redundant packet(1000 bytes to interest region and 1000 bytes to background). Without packet loss, Method-4 represents better result

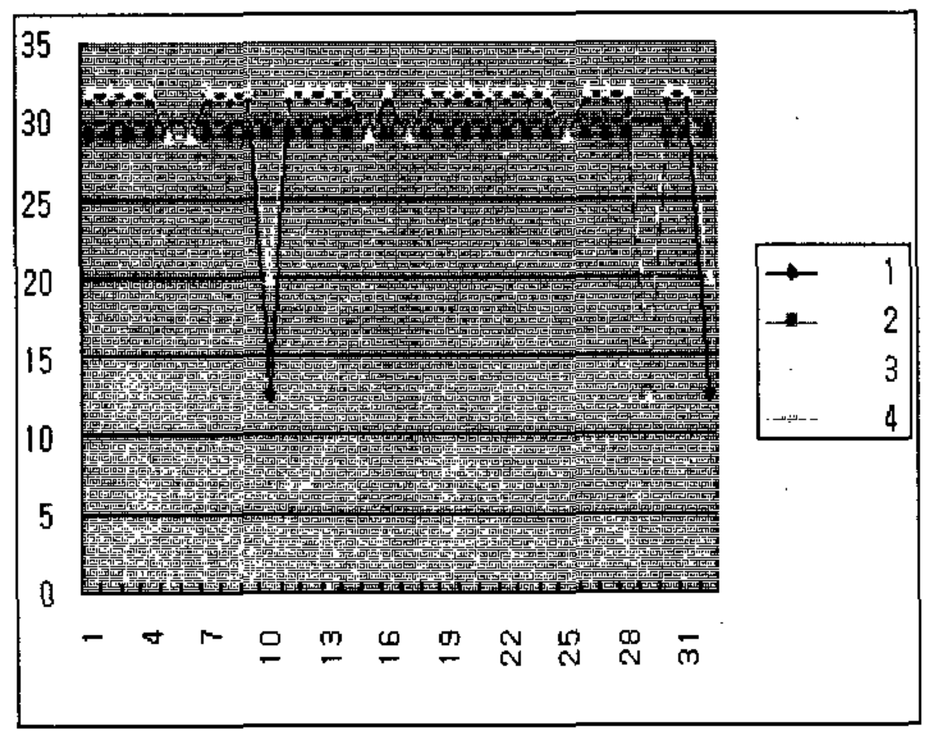


그림 6. 4000 바이트를 사용한 실험결과(멀티프레임)
Fig. 6. Simulation result of multi frame using 4000 bytes.

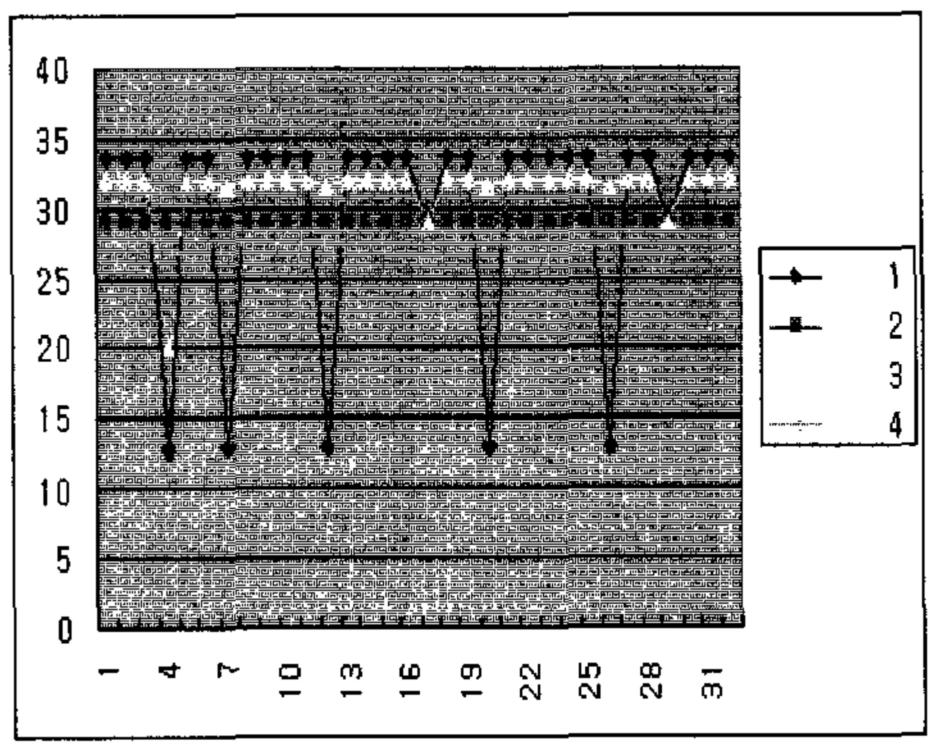


그림 7. 6000 바이트를 사용한 실험결과(멀티프레임)
Fig. 7. Simulation result of multi frame using 6000 bytes.

than other methods if we adopt N-PSNR. If we lose the first packet, our method is better result than other methods because our method can use a redundant packet to recover background. If we lose the second packet, Method-4 gives a little better result than other methods. If we lose packet-1 and packet-2, Method-2 and suggested method gives better result than the other two methods. If we do not have packet loss, Method-1 is better result than other methods, but the result is poor when packet loss happens. On the contrary, Method-2 and Method-4 is not so good result in case that packet loss does not happen. Suggested algorithm performs relatively well in both cases where packet losses happens or packet loss does not happen.

Fig. 6 represents the simulation result of multi frames which uses 4000 bytes/ frame and consider channel errors. As we mentioned in chap. 4, we assume that the channel follows a two-state Markov model($P_{00} = 0.9, P_{11} = 0.2$). The result tells our scheme performs well in case first packet is successfully received. But the first packet is lost, performance is not so satisfactory compared with other methods.

Fig. 7 represents the simulation result which uses 6000 bytes/frame and Method-2, Method-3 and Method-4 adds redundant packet(In case of Method-2, packet-3 is exactly same as packet-1). The performance of our scheme is better than others even if we lose packet-1 because we can recover the total frame using redundant packet. Simulation results of Fig. 7 shows that our scheme represents better results than other methods through all sequences. But one exception is that the quality of 4th frame is not so satisfactory because we could not receive packet-3 as well as packet-1. If we prevent this problem, we have to add more redundant packets, which sacrifices bit rate.

Ⅵ. Conclusion and further study

we suggested a new multiple description selective coding scheme which allocates higher passes to the interest region compared with background, thus improving the image quality especially for the interest region. Our algorithm extends embedded scheme to the generalized MDC framework. Suggested multiple description coding with embedded coding scheme could overcome packet loss, and also recover interest regions better image quality because it considered the importance of interest regions. Simulation results show that our scheme has better results especially in the circumstances where very low bit rate requires and special parts are more important than other parts.

In future research, we will control the amount of redundant packets according to the transmission channel status. Also we will extend this algorithm to video sequences.

Reference

[1] A.C. Miguel, etc.,"SPIHT for generalized multiple description coding," ICIP-1999, pp. 842-846.

[2] A. E. Mohr, etc.,"Generalized multiple description coding through unequal forward error protection", Proceeding of ICIP, 1999, No. 3,pp. 277-299.

[3] R. W. Buccigrossi and E. P. Simoncelli, "Progressive wavelet image coding based on a conditional probability model", ICSSAP, Apr., 1997, pp.2957-2960.

[4] Song Xiao, etc., "A new robust Multiple Description Coding Method based on region of interest",AINA '03, 17th international conference March 2003, pp. 501-504.

[5] Y. Rui, etc. ,"Relevance Feedback: A Power tool for interactive content-based image retrieval", IEEE trans. On Circuits and Video tech., Special Issue on Segmentation Description, and retrieval of Video Content,8(5), Sep. 1998.

[6] Qi Tian, etc.,"Combined user defined region-of-interest and spatial layout for image retrieval", ICIP-2000, pp.746-749.

[7] J. B. Lee and S. D. Kim,"Moving target extraction and image coding based on motion information", IEICE Trans. Fundamentals, Vol. E78-A, Jan. 1995, pp.127-130.

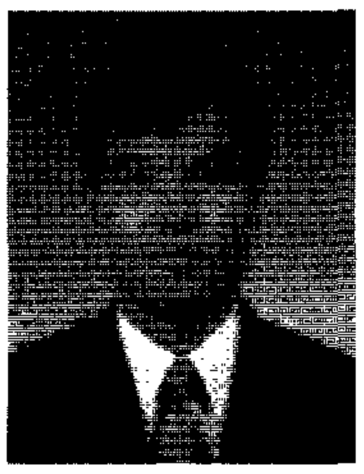
[8] E.N. Gilbert, "Capacity of a burst-noise channel", BellSyst. Tech. J., Vol. 39, No.5, 1960, pp. 1253-1265.

[9] B. Natarajan,"Filtering random noise from deterministic signals via data compression", IEEE Tras. Signal Processing,43(11), Nov., 1995.

[10] E. Scwartz, etc.,"Implementation of compression with reversible embedded wavelets", in Proc. SPIE, 1995.

[11] T. Cover and J Thomas, Elements of Information Theory, Wiley, Newyork, 1991.

저 자 소 개



이 종 배(정회원)

1985년 서울대학교 공과대학 제어계측학과 학사

1987년 한국과학기술원 전기 및 전자 공학과 석사

1996년 한국과학기술원 전기 및 전자공학과 박사

1987년~1996년 LG전자 연구원

1996년~1997년 특허청 영상기기 심사관

1997年~현재 명지전문대학 부교수

<주관심분야 : 정보처리, 영상통신>