한국통신학회 1987년도 춘계학술발표회 논문집(1987. 5. 23) 87408

프레임간 및 양갈래 탐색 벡터 양자화기를 혼합한 영상 부호화 시스템

* 금낙연, ** 최종수
*,** 중앙대학교 전자공학과

A Hybrid Interframe/BTVQ Image Coding System

Kum Nak Youn, Choi Jong Soo
Dept. of electric Eng. Chung-ang University

electronic Eng.

Abstract

A new efficient coding system which can transmit video conference or videophone signals at a 64Kbps is proposed. In addition to the interframe and CRC (Conditional Replenishment Coding) system, BTVQ (Binary Tree-searched Vector Quantizer) and RLC (Run Length Coding) methods are incorporated. Double buffer memory is used for simple control of channel symbol transmission and memory underflow. And also, buffer memory overflow is easily controlled by the thresholds of a MAD (Moving Area Detector).

1. Introduction

When we attempt to transmit image signals through a digital transmission line,we meet with a serious problem that bandwidth expands enormously. And so, two different efforts have been attemped to solve the problem.One is to replace the existing digital transmission line with a high quality one such as optical fiber which is capable of transmitting much more information. The other is to lower the necessary bit rates to transmit image signals to the minimum level without loss of fidelity by compressing them which contain much redundancies.

In any case, to lower the bit rates is very desirable one on the economical viewpoint. And so, various image compression techniques have been reported since 1970's,they can be classified three groups[1]. One is transform coding methods [1,2,3], another is predictive coding methods [1,4], and the other is vector quantization methods [1,5,6,7].

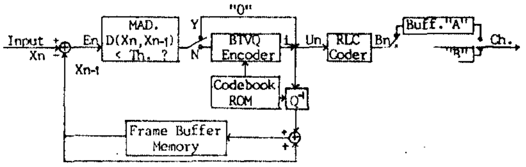
Considering the above method's advantages,in this paper, we construct a system which can compress videophone or video conference signals to the 64 Kbps level by incorporating an interframe coding which has the highest reduction rate among redundancy **reduction** techniques, conditional picture element replenishment coding which is very useful in the case of low-to-moderate movement images, BTVQ which is suboptimal but requires far less searching, and RLC.

We simulate the system using inside and outside training sequence images to verify a stability of the system. And the performance is evaluated on the basis of SNR, average distortion, memory requirements, and computational coplexity.

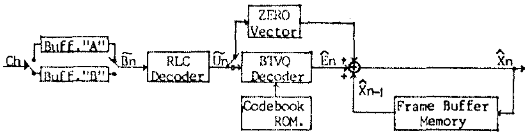
This paper is orgainzed 4 chapters: ch. 1 deals with an introduction, ch. 2 shows the entire block diagram of the system proposed in this paper and provides detail explanation, ch. 3 represents the results and methods of computer simulation, and ch. 4 concludes this paper.

2. Hybrid Coding System

The basic block diagram of the hybrid interframe/BTVQ coding and decoding system proposed in this paper is presented in Fig. 1.



(a) Encoder



(b) Decoder

Fig. 1. Block diagram of the Hybrid Coding System.

At the transmitting end, first, interframe difference signals are calculated using the previous frame buffer memory values and led to the moving area detector. Moving area detector, then, **decides** whether vector quantize or not using given thresholds. If a vector turns out to be a moving area,it passes through the BTVQ and an output index of the BTVQ is led to the RLC. If the vector proves to be **a non-moving area,an index "0" is delivered to the** RLC. Run length coded signals are, then, stored in buffer memory "A" for the odd frames and in buffer memory "B" for the even frames. When buffer memory "A" is written from the **transmitting** end, buffer memory "B" is read out to the receiving end and vice versa.

At the receiving end, decoded signals can be reproduced through the reverse operation described above. The following 5 sections give somewhat more detail explanations of the system.

2-1. Interframe Coding

These methods use the characteristics that image signals have much redundancies in the time domain. And so, they code only the difference signals between **the current** input vector and frame buffer memory value. As a result, the first order entropy (Gray level) reduces and the number of bits to be transmitted reduces accordingly [8,9,10,11]. Among various redundancy reduction **techniques these methods** have the highest redundancy reduction rate [9, 10]. Specifically, in the case of low-to-moderate movement circumstances such as video conference or videophone, these methods can provide better performances.

In **this paper, we use** a (192*128) frame buffer memory with 256 gray level to calculate the difference signals between actual input signals and the values of the buffer memory. The buffer memory values are refilled with **the** reproduction vector which corresponds to the channel symbol of the BTVQ when an input vector passes through the BTVQ, but remain to the previous values when an input vector proves to be a non-movement area.

After a frame is executed, the buffer memory can be considered as the decoded output unless we consider the channel errors.

2-2. Conditional Replenishment Coding

This technique is mainly companied with the interframe coding system. Conventional CRC system codes selectively the pixels which show large difference between two succeeding images, and transmit the difference value along with **the address** which indicates the position of the pixel. In high compression system, however, the address bits become nuisance in bits rate reduction.

Considering the statistical characteristics of the videophone or video conference images, in this paper, we incorporate a VQ method, and assign one bit "0" to the vector considered non-movement area and vector quantize to the moving area. A bit "1" is added to the MSB (Most Significant Bit) of the index of BTVQ to distinguish between the vector quantized one and not. MAD (Moving Area Detector) decides whether vector quantize or not using given thresholds, and the thresholds are controlled by the occupancy of the buffer memory.

2-3. BTVQ

BTVQ is a kind of VQ methods.General VQ provides high reduction rate, and can accomplish a performance arbitrarily close to lower bounds provided by Rate-Distortion theory [7] in a given rate by enlarging **vector dimension [1,5,6]. But it has some** shortcomings as follows.

First,it takes too much time to search a codeword which gives minimum distortion among codewords stored in a codebook. Second, it needs much memories **to store the codebook. Third,the decoded image** shows blocky effect in general. Forth, the system can be unstable for outside training sequence images [1].

But the interframe/BTVQ presented in this paper can overcome all of them except the memory requirements problem. The system is suboptimal but requires far less searching time, and also stability and blocky effect problems are improved [6,12].

The structure of the BTVQ codebook is depicted in Fig. 2.

The number of codewords CW in this paper is taken to 512 for simulation purpose. More codewords, however, would be appropriate for practical purposes. The exepansion of codewords does not affect largely on bit rate in this method. To design the codebook,splitting method [1,5,6,13] were used and parameters were taken using trial and error as 0.001 for threshould and 1.0 for delta.

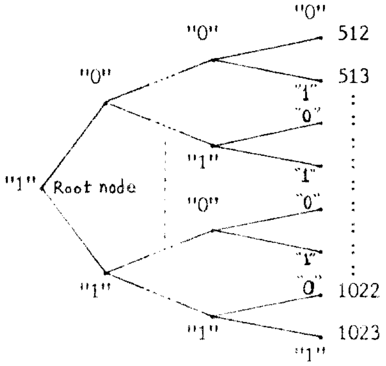


Fig. 2. The Structure of BTVQ Codebook.

The distortion caused by reproducing an input vector X by a reproduction vector $\hat{X}$ is represented as $D(X, \hat{X})$, and squared error distortion is used for this system as in equation 1.

$$D(X, \hat{X}) = \sum_{i=1}^{k} |x_i - \hat{x}_i|^2 , \quad (1)$$

and vector dimension $k = 64$.

Therefore the rate of the BTVQ is

$$R = \log_2 CW \text{ bits/vector}, \quad (2)$$
$$= \log_2 512 = 9 \text{ bits/vector},$$
$$r = R/k \text{ bits/pixel}, \quad (3)$$
$$= 9/64 = 0.140625 \text{ bits/pixel},$$

computational complexity is

$$K = 2 * R \text{ distortions}, \quad (4)$$

and memory requirements is

$$M = 2 * CW - 1 \text{ vectors}. \quad (5)$$

For the full search VQ,

$$K = CW \text{ distortions}, \quad (6)$$
$$M = CW \text{ vectors}. \quad (7)$$

Although the memory requirement is increased almost doubled, the number of distortion computations is drastically reduced.

2-4. Run Length Coding

To transmit the "0" bits somewhat more efficiently, **we adopt the RLC [14,15,16]. The indexes of** the BTVQ are run length coded with fixed rate by simply adding "1" to the MSB of the indexes. The "0" bits being generated frequently in this system are also run length coded with fixed rate for simplicity. We considerd 4 bits with leading "0" indicating nonmovement area. Fig. 3 dipicts the bits allocation.

'1'	B	T	V	Q	I	N	D	E	X

(a) For the Inedexes of BTVQ.

'0'	R	L	C

(b) For the Non-movement Area.

Fig. 3. Bits Allocation of the RLC.

2.5 Buffer Memory

In this system,the rate of data generation varies with the amount of movement.and so,a buffer is necessary to transmit the data with fixed rate. The buffer memory has a double memory structure in this system for easy control of underflow and data transmission. They are constructed of two 2.2 Kbit memories.

While the data are written into buffer memory "A", the data in the buffer memory "B" are read out. In the succeding frame, the buffer memory "B" are written while the buffer memory "A" are read out. When the data generation rate is less than the rate of 2.2 Kbit/frame, the data fill up only a part of the buffer memory. Thus, insignificant data which had formerly been written remain in the rest part the memory. Although the whole data in the 2.2 kbit memory are transmitted, only the significant data are decoded at the transmitter. Therefore no buffer memory underflow occurs in this system[9].And also,the buffer memory overflow is controlled by changing the thresholds of moving area detector according to the buffer memory occupancy.

3. Methods and Results of Simulation.

In the simulation, efforts are mainly devoted to the design of codebook of the BTVQ.We have to have a training sequence to design codebook. Therefore, we first construct the training sequence by selecting subblocks which satisfy given conditions from interframe difference images made from two consecutive or far away more than one frame images.

The images used for making the training sequence are 16 frame consecutive images named ELL101 - ELL116. The images are of size (192 * 128). The difference images are subdivided into (8*8) subblocks. And the subblocks constitute a vector each To be the training sequence,the vector must satisfy a condition that average distortion is greater than 3 or number of pixels, which are greater than 12, is greater than 8. The values were taken subjectively by author. In this way, we constitute 4096 training sequence vectors.

Using the training sequence, then, we design the 512 level codebook by splitting method mentioned in section 2-3. Fig. 2 shows the structure of the binary tree codebook, where the sequence of node indexes searched by the BTVQ become a channel symbol of the input vector.

Thereafter, we simulate the system for evaluation of performance using inside and outside training sequence images. The performance is evaluated on the basis of SNR, average distortion per pixel presented in equation 8 and 9 respectively.

$$\mathrm{SNR} = 10 * \log_{10} \frac{E(\|X\|^2)}{E[D(X,\hat{X})]} \tag{8}$$

$$\mathrm{AVGDIS} = \frac{1}{KL} \sum_{j=1}^{L} \sum_{i=1}^{K} D\,(\,xi - \hat{x}i\,) \tag{9}$$

, **where L is # of bubblocks**

The results are summarized in table 1, and the resulted images not only for inside and outside training sequences but also for scene changes are shown in Fig. 4.

Table 1. The Results of Simulation.

	SNR	Avgdis/pel
Inside Train.	24.72	4.84
Outside Train.	27.44	3.69
Scene Change	8.12	34.03

(a-1) Original **(a-2) Buffer** **(a-3) Result**

(b-1) Original **(b-2) Buffer** **(b-3) Result**

(c-1) Original **(c-2) Buffer** **(c-3) Result**

Fig. 4. The Resulted Images for the
(a) Inside Training Sequence image.
(b) Outside Ttaining Sequence Image.
(c) Secene Change.

4. Conclusion

This paper has described a hybrid interframe/BTVQ coding system which is capable of transmitting videophone or video conference signals at a 64 Kbps rates. By incorporating interframe conditional replenishment coding, BTVQ, and other techniques, this system has realized a reduction down to the practical level.

Simulation results have shown that the system is well suited for low-to-moderate movement images. Furthermore, the results on the outside training sequences verify that the system is stable. But in the case of scene change, a few frames have to be executed to recover a reasonable picture quality. To reduce the time some more codewords should be designed.

We conclude that the system is very useful for practical purposes provided universal and some more codewords.

References

1. R. M. Gray, " Vector Quantization," IEEE ASSP Mag. vol. 1, pp. 4-29,Apr. 1984.
2. P. A. Wintz," Transform Picture Coding," Proc. IEEE, vol. 60, pp.809-819,July 1972.

3. T. J. Natarajan and N. Ahmed, " On Interframe Transform Coding, " IEEE Trans. Commun. vol. COM-25, pp. 1323-1329, Nov. 1977.
4. W. Zschunke," DPCM Picture Coding with Adaptive Prediction,"IEEE Trans. Commun. vol. COM-25, pp.1295-1302, Nov. 1977.
5. Y. Linde, A. Buzo, R. M. Gray, " An Algorithm for Vector Quantizer Desgn," IEEE Trans. Commun., vol. COM-28, pp. 84-95, Jan. 1980.
6. R. M. Gray and Y. Linde, " Vector Quantizers and Predictive Quantizers for Gauss-Markov Sources,"IEEE Trans. Commun., vol. COM-30, pp. 381-389, Feb. 1982.
7. L. D. Davisson, " Rate-Distortion Theory and Application," Proc. IEEE,vol. 60, pp. 800-808, July 1972.
8. K. Iinuma, Y. Iijima, T. Ishiguro, H. Kankeko, and S. Shigaki," Interframe Coding for 4-MHz Color Television Signals," IEEE, Trans.Commun. vol. COM-23, pp. 1461-1466,Dec. 1975.
9. N. Mukawa, H.Kuroda,and T Matsuoka," An Interframe Coding System for Video eleconferencing Signal Transmission at a 1.5 Mbit/s Rate,"IEEE Trans. Commun. vol.COM-32, pp.280-287, March 1984.
10. H. Yasuda, H. Kuroda, H. Kawanishi,F. Kanaya, and H. Hashimoto, "Transmitting 4-MHz TV Signals by Combinational ifference Coding,"IEEE Trans.Commun. vol.COM-25,pp.508-516, May 1977.
11. B. G. Haskell, P. L. Gordon, R. L.Schmidt,and J. V. Scattaglia, " Interframe Coding of 525-Line, Monochrome Television at 1.5 Mbitss," IEEE Trans. Commun. vol. COM-25, pp.1339-1348, Nov. 1977.
12. Y. Matsuyama and R. M. Gray, " Universal Tree Encoding for Speech,"IEEE Trans.Inform.Theory, vol. IT-27, pp. 31-40, Jan. 1981.
13. M. J. Sabin and R. M. Gray,"Product Code Vector Quantizers for Waveform and Voice Coding, "IEEE Trans.Acoust.,Speech, Signal Processing, vol. ASSP-32, pp.474-488,June.1984.
14. H. Tanaa and A. Leon-Garcia, "Efficient Run-length Encodings," IEEE Trans. Inform. Theory, vol. IT-28, pp. 880-890, Nov. 1982.
15. W. H. Chen and W. K. Pratt, " Scene Adaptive Coder," IEEE Trans. Commun.vol COM-32,PP.225-232, March 984
16. R. C. Gonzalez and P. Wintz, " Digital Image Processing,"Addison-Wesley Publishing company, inc.Advanced Book Program,Massachusetts, 1979.

*** This research was conducted as a commission from the ETRI.**
The authers would like to thank the related persons.