NON-CAUSAL INTERPOLATIVE PREDICTION FOR B PICTURE ENCODING

Tomoya HARABE, Akira KUBOTA, and Yoshinori HATORI

Tokyo Institute of Technology,
Interdisciplinary Graduate School of Science and Engineering,
Department of Information Processing,
226-8502 4259 G2-911 Nagatsuta Midoriku Yokohama, Japan
E-mail: harabe.t.aa@m.titech.ac.jp

ABSTRACT

This paper describes a non-causal interpolative prediction method for B-picture encoding. Interpolative prediction uses correlations between neighboring pixels, including non-causal pixels, for high prediction performance, in contrast to the conventional prediction, using only the causal pixels. For the interpolative prediction, the optimal quantizing scheme has been investigated for preventing coding error power from expanding in the decoding process. In this paper, we extend the optimal quantization sceme to inter-frame prediction in video coding. Unlike H.264 scheme, our method uses non-causal frames adjacent to the prediction frame.

Keywords: non-causal interpolative, B-picture, image processing, optimal quantization scheme

1. INTRODUCTION

Predictive coding is examined as a type of highly efficient coding, using spatially and temporally high-correlation methods, such as differential pulse code modulation (DPCM) and motion compensated prediction (MC). However these conventional predictive coding techniques use only causal signals. As for time-varying image, it is known that there is strong correlation interpixel or interframe. Therefore, predictive coding using pixels that will be encoded in future, can the prediction performance improve [1]. We called this method "non-causal interpolative prediction in this paper. For this interpolative prediction, the optimal quantizing scheme (OQS) has been proposed to suppress the coding error power expantion by minimizing quantization errors [2], [3]. B-picture "Bi-Directional Predictive Picture, is based on encoding method the interframe prediction coding: B-picture is encoded using I (Intra coded) picture or P (Predictive coded) picture as the reference frames. The compression can be as lowest as possible the highest increasing the number of B-pictures, however, there is a problem that the prediction error increases with the possible time interval between the IP picture.

This paper presents a novel encoding method based on non-causal interpolative prediction in the time direction. We show that non-causal interpolative prediction can reduce the prediction error and that using OQS in the non-causal interpolative prediction provides higher PSNR than the conventional B-frame coding method.

2. PICTURE STRUCTURE IN ENCODING

One method used in various video formats to reduce file size is interframe prediction. For many frames of moving image, the only difference between one frame and another is the result of either the camera moving or an object in the frame moving. In reference to a video file, this means much of the information of one frame will be the same that of the next frame. The three major picture types used in the interframe prediction are I, P, and B. Figure.1 shows a simplified diagram of the picture structure.

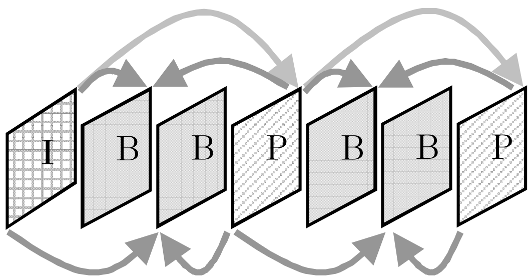

Fig. 1: Picture structure in encoding

I-pictures are the least compressible but don't require other video frames to decode. P-pictures can use data from previous I frames to decompress and are more compressible than I frames. B-pictures can use both previous and forward frames for data reference to get the highest amount of data compression.

3. INTERPOLATIVE PREDICTION CODING USING THE OPTIMAL QUANTIZING SHEME

This section describes interpolative prediction coding process and the problem that the coding error power expands in the decoding process. We also show the optimal quantizing scheme suppresses the coding error power.

3.1 Inter polative prediction coding

Interpolative prediction coding is a block-based coding method. Figure. 2 shows a n pixel $\times$ 1 line image block as an example of interpolative prediction coding.

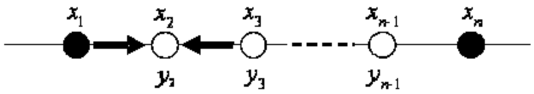

Fig. 2: Examlpe of interpolative predictive coding

The black circles in the figure are pixels already encoded by some method. The other pixels are prediced from the average values of the neighboring two pixels. The encoding process for $x_i (i = 2, 3, \cdots, n-1)$ is expressed as

$$y_i = x_i - \frac{1}{2}(x_{i-1} + x_{i+1}) \tag{1}$$

where y_i denotes the prediction error of x_i. The encoding process of all the pixels within the block can be expressed in vector-matrix from as

$$\boldsymbol{y} = \boldsymbol{C}\boldsymbol{x} \tag{2}$$

where $\boldsymbol{x}$ is the vector obtained by lexicographic ordering of the pixel value within the block, $\boldsymbol{y}$ is the prediction error vector, and $\boldsymbol{C}$ is the $n \times n$ coding matrix such that

$$\boldsymbol{C} = \begin{bmatrix} 1 & & & & & \\ -\frac{1}{2} & 1 & -\frac{1}{2} & & 0 & \\ & -\frac{1}{2} & 1 & -\frac{1}{2} & & \\ & & \ddots & \ddots & \ddots & \\ & 0 & & -\frac{1}{2} & 1 & -\frac{1}{2} \\ & & & & & 1 \end{bmatrix}.$$

The predictive errors are quantized for higher compression. Let$\boldsymbol{q}$ be the vector of the quantizing error. The coding error vector, which is obtained from the difference between the true value and the decoded value, can be expressed as

$$\boldsymbol{e} = \boldsymbol{A}\boldsymbol{q} \tag{3}$$

where $\boldsymbol{A}$ is the $n-2 \times n-2$ submatrix composed of the elements of $\boldsymbol{C}^{-1}$ corresponding to the pixels to be encoded. The mean square coding error, i.e., the coding error power, is

$$\begin{aligned} D &= E[\boldsymbol{e}^T\boldsymbol{e}] \\ &= Tr[\boldsymbol{A}^T\boldsymbol{A}]\sigma_q^2 \\ &= Tr[\boldsymbol{B}]\sigma_q^2 \ \ (\boldsymbol{B} := \boldsymbol{A}^T\boldsymbol{A}) \end{aligned} \tag{4}$$

where σ_q^2 is the quantization error power and $Tr[]$ represents trace operation.

Therefore, it is found that the coding error power is lager than the quantization error power because of $Tr[\boldsymbol{B}] > 1$. In the conventional DPCM, these powers can be equal.

3.2 Optimal Quantizing Sheme

The optimal quantizing scheme (OQS) uses proposed as a highly efficient coding method to solve the above problem [2], [3]. Figure. 3 shows the block diagram of the OQS encoder.

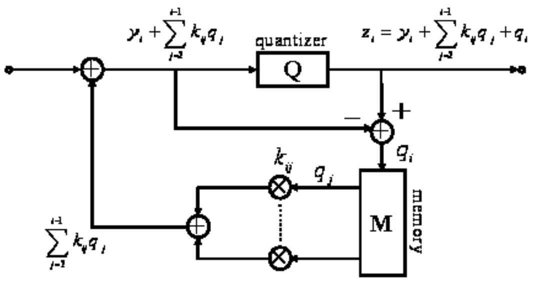

Fig. 3: Block diagram of optimal quantizing scheme

In this scheme, previously quantized errors within the block are stored in memory. The differential quantization of the i th pixel is performed after adding the stored errors q_j $(j = 1, 2, \cdots, i-1)$ multiplied by coefficient $k_{i,j}$ to the prediction error y_i. The i th output of the scheme is computed as

$$z_i = y_i + \sum_{j=2}^{i-1} k_{i,j} q_j + q_i. \tag{5}$$

Because of the structure of the OQS, the coefficient matrix $\boldsymbol{K}$ became the upper triangular matrix

$$\boldsymbol{K} = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ k_{21} & 1 & 0 & & \vdots \\ k_{31} & k_{32} & 1 & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ k_{n1} & \cdots & \cdots & k_{n,n-1} & 1 \end{bmatrix}.$$

Here, we determine the value of $\boldsymbol{K}$. The differential quantization error vector through the OQS is written by

$$\boldsymbol{q'} = \boldsymbol{K}\boldsymbol{q}, \tag{6}$$

so that the coding error vector for the OQS is computed as

$$\boldsymbol{e'} = \boldsymbol{A}\boldsymbol{q'}. \tag{7}$$

Hence, using (7), the coding error power in the OQS result in

$$\begin{aligned} G &= E[\boldsymbol{e'}^T\boldsymbol{e'}] \\ &= Tr[\boldsymbol{K}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{K}]\sigma_q^2 \\ &= Tr[\boldsymbol{K}^T\boldsymbol{B}\boldsymbol{K}]\sigma_q^2 \\ &= Tr[\boldsymbol{F}]\sigma_q^2 \end{aligned} \tag{8}$$

where $\boldsymbol{F} := \boldsymbol{K}^T\boldsymbol{B}\boldsymbol{K}$ is diagonal matrix. Here, the matrices dividing $\boldsymbol{B}$ and $\boldsymbol{K}$ such that

$$\boldsymbol{B} = \left[\begin{array}{c|c} b_{11} & \boldsymbol{b}_1^T \\ \hline \boldsymbol{b}_1 & \boldsymbol{B}_1 \end{array}\right] = \left[\begin{array}{c|cc} b_{11} & \multicolumn{2}{c}{\boldsymbol{b}_1^T} \\ \hline \boldsymbol{b}_1 & b_{22} & \boldsymbol{b}_2^T \\ \cline{2-3} & \boldsymbol{b}_2 & \boldsymbol{B}_2 \end{array}\right] = \left[\begin{array}{c} \ddots \\ \begin{array}{c|c} b_{ii} & \boldsymbol{b}_i^T \\ \hline \boldsymbol{b}_i & \boldsymbol{B}_i \end{array} \end{array}\right]$$

$$\boldsymbol{K} = \left[\begin{array}{c|c|c|c} 1 & & & \\ \hline \boldsymbol{k_1} & 1 & & \\ \cline{2-4} & \boldsymbol{k}_2 & \ddots & \\ & & 1 & \\ & & \boldsymbol{k}_i & \ddots \end{array}\right],$$

each diagonal element of $\boldsymbol{F}$ can be expressed as

$$f_i = \left(1 \quad \boldsymbol{k}_i^T \right) \boldsymbol{B}_{i-1} \begin{pmatrix} 1 \\ \boldsymbol{k}_i \end{pmatrix} S \tag{9}$$

where $\boldsymbol{B}_0 = \boldsymbol{B}$. The optimal coefficient vector $\boldsymbol{k}_{i,opt}$ and the minimum of f_i in the OQS can be determined as follows

$$\boldsymbol{k}_{i,opt} = -\boldsymbol{B}_i^{-1}\boldsymbol{b}_i \tag{10}$$

$$f_{i,min} = b_{ii} - \boldsymbol{b}_i^T \boldsymbol{B}_i^{-1}\boldsymbol{b}_i \tag{11}$$

Therefore, the minimum coding error power is given by

$$G = D - (\sum_i \boldsymbol{b}_i^T \boldsymbol{B}_i^{-1}\boldsymbol{b}_i)\sigma_q^2. \tag{12}$$

Since the second term in (12) is nonnegative, the coding error power in the OQS is smaller than that when the scheme is not used.

3.3 OQS for inter-frame prediction

We use Non-causal interpolative prediction in the time direction and reduce the prediction error of B picture. In addition, PSNR is improved by using the optimal quantizing sheme. In this paper, "one-dimensional interpolative prediction" in the temporal domain, and use the block of the same position with the frame before and after compared with a certain block of the frame for encoding. The difference is obtained at every block by non-causal interpolative prediction. Figure.4 shows the encoding expression ($\boldsymbol{y} = \boldsymbol{C}\boldsymbol{x}$) to third frame when 8 pixel $\times$ 8 lines one block as example in non-causal interpolative prediction.

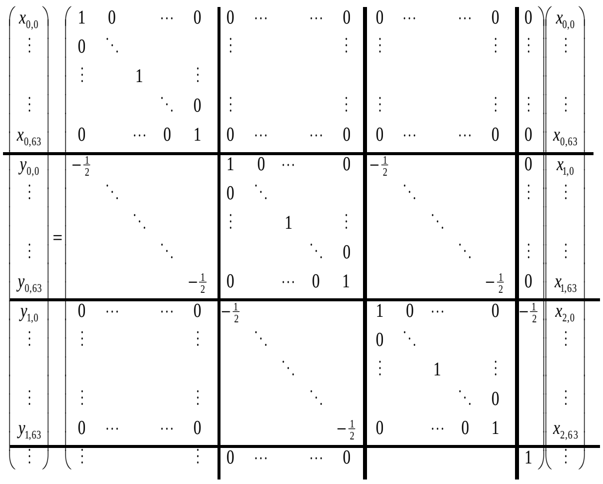

Fig. 4: Example of encoding expression

4. COMPUTER SIMULATION

4.1 Simulation condition

Computer simulations were carried out to evaluate the prformance of the proposed method. In the conventional bi-directional prediction method, each frame is predicted as a linealy weighted both ends of the objective frame for encoding. We use only luminance component of an image with 720 $\times$ 480 resolution. Block size 8 pixel $\times$ 8 lines and frames of both ends used have true values. The motion compensation was not used, the position where the block of intraframe was fixed. Moreover, the simulation result changes greatly depending on the characteristic of the quantizer. In this paper, we use a linear quantizer with constant steps. The step width is changed by quantization parameter QP.

4.2 Comparison by Entropy and Prediction error

We compared the conventional method (causal-interpolative prediction) and the propoed method (non-causal interpolative prediction) in the prediction error and the Entropy. Figure.5 and Figure.6 shows the histogram of the prediction error when the number of frame is five. It can be seen that the proposed method gathers error in the center distribution, and can suppress the prediction error compared with the conventional method.

Next, Table 1 shows the result of the entropy. The proposed method can reduce by 0.74bit compared with the conventional method in 5 frames. Moreover, the difference between the conventional method and the proposed method is 1.22bit in 10 frames. When comparing it by each method, The conventional method increased by 0.38bit, and tended to decrease by 0.09bit by the proposed method. Even if the number of frames increases, in the proposed method the entropy doesn't change. Therefore, it can be said that there is effectiveness because the prediction error is reduced and entropy is kept low when non-causal interpolative prediction is applied in the time direction.

Table 1: Entropy comparing of each number of frames

Entropy (bit)	Frame:5	Frame:10
conventional method	6.77	7.15
proposed method	6.02	5.93

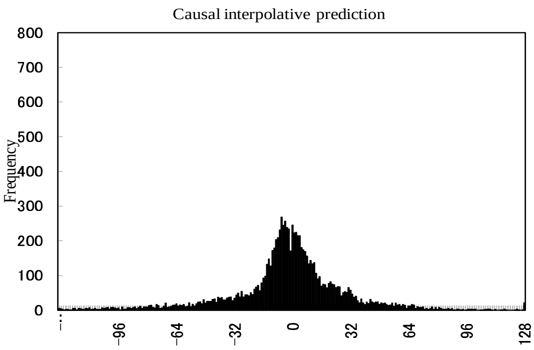

Fig. 5: Histogram of causal interpolative method

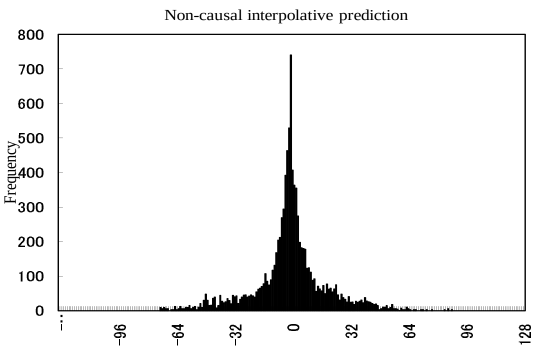

Fig. 6: Histogram of non-causal interpolative method

4.3 Comparison by PSNR and entropy of after decoding

Figure.7 shows the measured peak signal to noise ratio (PSNR) as a function of the entropy and compares causal and non-causal interpolative prediction coding. Quantization parameters(QP) has changed into 3 to 30. There is a moving object in the obtained block. Entropy decreases 0.44bit but PSNR is low about 5.6dB by non-causal interpolative prediction. It is considered that the encoding error is increased when decoding. PSNR increased by 4.76dB after applying the optimal quantizing scheme to this non-causal interpolative prediction and the entropy increased by 0.18bit. Therefore, PSNR was improved at the same entropy compared with the conventional method.

5. CONCLUSION

In this paper, we showed that prediction error reduction and PSNR inprovement can be achieved by extending non-causal interpolative predition with OQS in to the time direction for B-picture coding. As future work, we take into account within compensation in the proposed method and fast algorithm on calculation inverse of the coding matrix.

6. REFERENCES

[1] A. Machizawa, M. Tanaka, “Image Data Compression Method Based on Interpolative DPCM by Area Decomposition,” IEICE Japan, Vol. J69-D, No. 3, pp. 375–382, Mar. 1986.

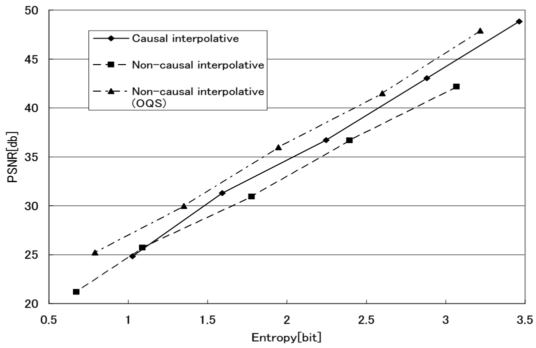

Fig. 7: Comparison of the interpolative and the non-causal interpolative and the optimal quantizing scheme

[2] T. Fujita, A. Kubota, Y.Hatori, “Adaptive Control Method by Permuting Quantization for Interpolative Prediction Coding,” IEICE Japan, Vol. J69-D, No. 3, pp. 375–382,Mar. 1986.

[3] Y. Hatori, “Optimal Quantizing Schemes in Interpolative Prediction,” IEICE Japan, Vol. J66-B, No. 5, pp. 599–606, May. 1996.