A Novel Encoding Method for P-frames

Cui Wang · Akira Kubota · Yoshinori Hatori

Abstract In this paper, the feasibility and efficiency of non-causal prediction for P-frames is examined, and based on the findings, a new P-frame coding scheme is proposed. Motion-compensated inter-frame prediction, which has been used widely in low-bit-rate television coding, is an efficient method for reducing temporal redundancy in a sequence of video signals. To this end, the proposed scheme combines motion compensation with non-causal prediction based on an interpolative, but not Markov, representation. Nevertheless, energy dispersion occurs in the scheme as a result of the interpolative prediction transform matrix being non-orthogonal. To solve this problem, we have introduced a new conditional pel replenishment method. On the other hand, we have applied rotation scanning, which is also applied for feedback quantization, as a quantizer. Simulation results show that the proposed coding scheme achieves an approximate 44 dB when entropy is less than 1 bit/pixel.

Keywords Novel coding scheme · Inter-frame coding · Interpolative representation

1 Introduction

Motion-Compensated (MC) image coding, which takes advantage of frame-to-frame redundancy to achieve a high data

C. Wang (✉) · Y. Hatori
Interdisciplinary Graduate School of Science and Engineering,
Tokyo Institute of Technology, Tokyo, Japan
e-mail: tracywang0528@hotmail.com

Y. Hatori
e-mail: hatori@ip.titech.ac.jp

A. Kubota
Faculty of Science and Engineering, Chuo University, Tokyo, Japan
e-mail: kubota@elect.chuo-u.ac.jp

compression rate, is one of the most popular inter-frame coding techniques [1, 2]. For the H.26x family of video coding standards, motion estimation (ME)/MC coding tool combined with an orthogonal transform (OT), such as a discrete cosine transform (DCT), has been introduced. This tool now plays an important role in the field of inter-frame coding. According to the conditional pixel replenishment method and quantization control in the DCT coefficient domain, the H.26x standards gain considerable coding efficiency and transmission bandwidth reductions by applying this method.

Conversely, we have developed a hybrid I-frame en-coding method based on non-causal interpolative prediction and differential feedback quantization that utilizes intra-frame spatial correlation [3]. To verify the efficiency of our hybrid coding method, we have coded four test sequences using the method. As a result, a considerable coding efficiency has been achieved by applying the developed method [3].

In this paper, a new configuration for P-frame coding is presented. In designing this hybrid coding scheme, we show that orthogonal transforms need not be considered as constraints.

2 Proposed Scheme

The proposed coding scheme is shown in Fig. 1. In this model, MC predictive coding is performed first, and then, the residual signal is encoded by an interpolative prediction (IP) method based on an 8×8 block [3]. We term this hybrid coding method the "MC+IP synthesis configuration."

2.1 Motion-Compensated Predictive Coding

MC predictive coding in the proposed method is identical to that used for inter-frame MC prediction of P-frames in the H.264 video coding standard. Here, the number of reference frames is set equal to one.
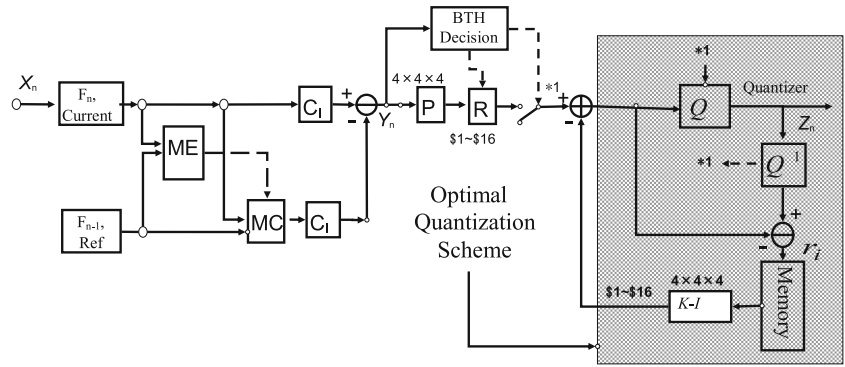
Figure 1 Proposed coding scheme.

2.2 Interpolative Prediction

The residual signal after motion compensation is coded by an IP method, which is also known as non-causal prediction, based on 8×8 blocks. Given a block whose boundary conditions were known, coding based on an interpolative model could be realized [4].

Given a 8×8 block, and the pixels in this block, $x_1 \sim x_{64}$ are rearranged in a conventional order as a 64×1 vector X_n. Then multiply X_n by the predictive matrix C to get the prediction errors. The encoding matrix C used in this paper is described in Eqs. (2)~(5). This predictive matrix here is similar to the one presented in Ref. 3, except for the elements that correspond to the four corner pixels of a block in the intra-block processing.

A configuration of MC predictive coding with interpolative prediction is shown in Fig. 2.

In our configuration, the predictive error Y_n can be expressed as

$$Y_n = C \times X_n - C \times f\left(X'_{n-1}\right) = C \times \left(X_n - f\left(X'_{n-1}\right)\right) \quad (1)$$

where X_n is the input 8×8 block signal in 64×1 vector form after last order scanning; X'_{n-1} is the reference vector in the last reconstruct frame at the exact same position; $f(\ )$ indicates the MC function, and therefore, $f(X'_{n-1})$ is the vector of X'_{n-1} after
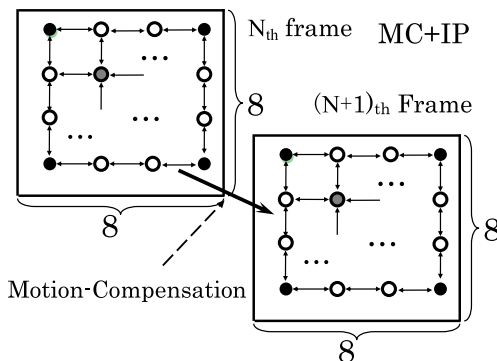

Figure 2 MC predictive coding with interpolative prediction.

MC processing; and C is the 64×64 predictive matrix, which can be represented as follows:

$$C = \begin{bmatrix} A_1 & & & & O \\ A_3 & A_2 & A_3 & & \\ & \ddots & \ddots & \ddots & \\ & & A_3 & A_2 & A_3 \\ O & & & & A_1 \end{bmatrix} \quad (2)$$

$$A_1 = \begin{bmatrix} 1 & & & & \\ -\frac{1}{2} & 1 & -\frac{1}{2} & O & \\ & \ddots & \ddots & \ddots & \\ & O & -\frac{1}{2} & 1 & -\frac{1}{2} \\ & & & & 1 \end{bmatrix} \quad (3)$$

$$A_2 = \begin{bmatrix} 1 & & & & \\ -\frac{1}{4} & 1 & -\frac{1}{4} & O & \\ & \ddots & \ddots & \ddots & \\ & O & -\frac{1}{4} & 1 & -\frac{1}{4} \\ & & & & 1 \end{bmatrix} \quad (4)$$

$$A_3 = \begin{bmatrix} -\frac{1}{2} & & & & \\ & -\frac{1}{4} & & O & \\ & & \ddots & & \\ & O & & -\frac{1}{4} & \\ & & & & -\frac{1}{4} \end{bmatrix} \quad (5)$$

Here, A_1, A_2, and A_3 are 8×8 matrices, shown as Eqs. (3)~(5). About more details of matrix C see Appendix.

2.3 Optimal Quantization Scheme

The difference signals output by the interpolative process, which correspond to IP errors of MC residuals, are sequentially input into the feedback quantizer [5]. Accordingly,

coding errors resulting from power expansion in the inter-block processing, as a result of using a non-orthogonal system, can be solved [5, 6].

2.4 Conditional Pel Replenishment

As already mentioned, because an OT is not employed in our method, energy is not concentrated in one location, but is distributed throughout an entire block. Consequently, determining whether pixels should be quantized is an issue. For this reason, we have introduced conditional pel replenishment into the scheme.

As shown in Fig. 3, the data of a previously decoded frame F'_{n-1}, used for motion compensation, can also be used as a reference for conditional replenishment pixel control before the current pixel data are quantized. As a result of pel replenishment, transmission bandwidth is constrained.

Specifically, the decoded data of the previous frame are also processed by IP based on 8×8 blocks to obtain a set of reference values; whereupon these values are compared with a predefined pixel threshold (PTH). Because the reference values are obtained from the decoded data, this conditional pixel replenishment can be achieved without additional overhead information.

On the other hand, differences of IP outputs between current and previous frames (reference values) are then compared with the 4×4 sub-block threshold (BTH) (obtained by preliminary experiments). Thus, we can determine whether the pixels should be quantized in the 4×4 sub-block of the 8×8 block under consideration in the current frame. At this stage, since the reference values are obtained from the decoded data, motion vector information (which the decoder has already acquired), and current 4×4 sub-block data (which have not been transferred to the decoder yet), it is necessary to add the information for each 4×4 sub-block to verify whether it should be quantized. As a result, this conditional pixel replenishment can be achieved with 1 bit (ON/OFF) additional overhead information for every 16 pixels (4×4 sub-block).

A threshold replenishment algorithm for adaptive vector quantization was first proposed by Fowler [7] in 1998, and has
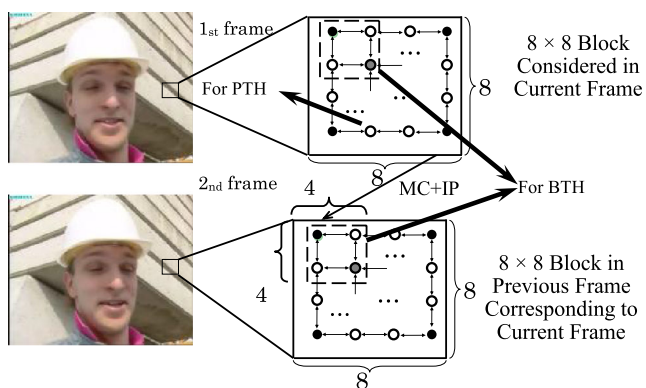

Figure 3 Conditional pel replenishment.

subsequently been used in various coding technologies [8–11]. The proposed conditional pixel replenishment method without codebook used in our method is based on that algorithm, because quantization is not performed on vectors, but on each individual pixel in an 8×8 block. Furthermore, the reference values are the output of IP, not the distortion measure between a code-vector and a quantization input vector [7]. Therefore, the distortion measure for each block is not computed, meaning that the computational time of the proposed method is less than that of Ref. 6.

However, the threshold values here are predefined and must be modified according to each frame of a video sequence. Improving the threshold selection process is considered to be an area for future research.

2.5 Rotation Scanning

In this paper, in order to achieve replenishment of pixels in the spatial domain, we have proposed a new approach for improving image quality. An input 8×8 block signal is reordered to adopt this sub-block system.

In our feedback quantization system, the power of a coding error can be expressed as

$$E = \sum \left(f_i \cdot \sigma_{q_i}^2 \right) \tag{6}$$

Here, f_i is the feedback coefficient for one 8×8 block, and σ_{qi}^2 is the power of the quantized error. When a 4×4 sub-block conditional pixel replenishment is performed, non-significant sub-blocks will not be quantized; as a result, the power of the quantized error in (6) is changed to the power of the prediction error. Generally, the power of the quantized error is less than that of the prediction error. Therefore, if we could reduce the value of f_i at the non-significant sub-block positions, we could suppress the increase in coding error powers in (6). According to Ref. 4, the value of f_i is defined by three matrices: the predictive matrix C_I, the scanning order matrix P, and the transform matrix D. Because OT is not used in our system, matrix D has been determined; predictive matrix C_I defines the prediction error, which also has been determined.

When the scanning order is 4×4×4, as shown in Fig. 4, a resultant 8×8 block of f_i values is shown in Table 1.

Each sub-block's $\sum f_i$ is as follows:

$$
\begin{aligned}
\#1 \quad & \sum (f_i) = 19.4527 \\
\#2 \quad & \sum (f_i) = 26.2304 \\
\#3 \quad & \sum (f_i) = 26.8412 \\
\#4 \quad & \sum (f_i) = 40.1022
\end{aligned}
\tag{7}
$$

Therefore, after deciding whether the current 4×4 sub-block is to be quantized (e.g., the sub-blocks of #4 need not be quantized because $\sum (f_i)$ of the sub-blocks of #4 is the largest value among all the sub-blocks), according to the
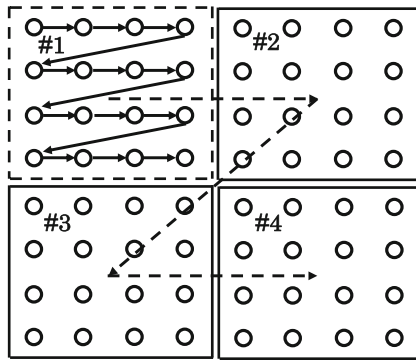
Figure 4 4×4×4 scanning order.

definition of distortion in a feedback quantization system (as shown in (6)), it is necessary to reorder the input signal to ensure that non-significant sub-blocks can be quantized as early as possible. Table 2 depicts this processing.

In Table 2, we show how, depending on the insignificant/ significant (expressed as 0 and 1 in various positions of a 4×4 sub-block) sub-block case, rotation processing should be assigned. L indicates a left rotation operation (a counterclockwise rotation); whereas R indicates a right rotation (a clockwise rotation); R^2 indicates two right rotations, and "none" indicates no rotation. However, there are two special cases in this table, which are marked with gray backgrounds: 0110→ 0111 and 1001→0111. In these cases, after a rotation operation, the last sub-block, #4, has been set to 1 mandatorily on the basis of preliminary experimental results.

The effects of improving coding efficiency by these rotation operations are shown in Table 3.

2.6 Features of the Proposed Scheme

The proposed inter-frame coding scheme has three characteristic features:

- OT is not used.
- Conditional pel replenishment is performed without additional overhead information.
- A new hybrid coding framework, MC+IP combined with feedback quantization, is employed.

Table 1 8×8 block of f_i values.

1.000	0.800	1.429	1.867	2.269	2.461	2.627	2.913
0.800	0.878	1.020	1.012	1.072	1.051	1.050	0.844
1.441	1.017	1.277	1.292	1.398	1.352	1.323	1.548
1.913	1.025	1.328	1.355	1.437	1.440	1.382	2.059
2.745	1.023	1.344	1.442	1.578	1.489	1.405	2.432
2.550	1.022	1.337	1.374	1.594	1.504	1.414	2.711
2.759	1.023	1.334	1.367	1.504	1.498	1.426	2.920
3.532	0.846	1.554	2.059	2.481	2.768	5.220	8.163

Table 2 Rotation processing for all cases.

Before rotation				After rotation		
$\begin{bmatrix}0&0\\0&0\\0&1\\1&1\end{bmatrix}$	$\begin{bmatrix}0&0\\0&1\\1&1\\1&1\end{bmatrix}$	$\begin{bmatrix}0&0\\1&1\\0&1\\1&0\end{bmatrix}$	None	$\begin{bmatrix}0&0\\0&0\\0&1\\1&1\end{bmatrix}$	$\begin{bmatrix}0&0\\0&1\\1&1\\1&1\end{bmatrix}$	$\begin{bmatrix}0&0\\1&1\\0&1\\1&1\end{bmatrix}$
$\begin{bmatrix}0&0\\1&0\\1&0\\1&0\end{bmatrix}$	$\begin{bmatrix}1&0\\0&1\\1&0\\1&1\end{bmatrix}$		L	$\begin{bmatrix}0&0\\0&1\\0&0\\1&1\end{bmatrix}$	$\begin{bmatrix}0&1\\1&1\\0&1\\1&1\end{bmatrix}$	
$\begin{bmatrix}0&1\\0&0\end{bmatrix}$	$\begin{bmatrix}0&1\\0&1\end{bmatrix}$	$\begin{bmatrix}1&1\\0&1\end{bmatrix}$	R	$\begin{bmatrix}0&0\\0&1\end{bmatrix}$	$\begin{bmatrix}0&0\\1&1\end{bmatrix}$	$\begin{bmatrix}0&1\\1&1\end{bmatrix}$
$\begin{bmatrix}1&0\\0&0\end{bmatrix}$	$\begin{bmatrix}1&1\\0&0\end{bmatrix}$	$\begin{bmatrix}1&1\\1&0\end{bmatrix}$	R^2	$\begin{bmatrix}0&0\\0&1\end{bmatrix}$	$\begin{bmatrix}0&0\\1&1\end{bmatrix}$	$\begin{bmatrix}0&1\\1&1\end{bmatrix}$

OT is not used in our coding scheme; instead, an IP method has been adopted. In fact, OT could not utilize the relations between pixels, but could merely transform the signal from the spatial to frequency domain [12–23]. In contrast, an IP method can compress the signal by eliminating the correlation between pixels within a frame [24–28]. Generally, the MC prediction error is independent of time; however, a spatial correlation still exists. Accordingly, we have replaced OT by an IP method, because in Ref. 3, we have shown that IP can be achieved as a "transform coding."

Conversely, the non-orthogonality of the IP transform matrix means that the power expansion problem, in which coding errors expand when decoded, exists in the proposed method. As a result, feedback quantization is necessary.

3 Simulation

We now present simulation results acquired by using the proposed P-frame coding scheme, and show predictive error comparison between our method and the traditional causal prediction, in this case, H.264 baseline [29]. The experiments here were designed to make two major points. First, we seek to show the coding efficiency of MC+IP without DCT. Second, we want to emphasize the importance of using the Conditional Pel Replenishment and Rotation technology in MC+IP configuration. Therefore, we give the experiment results without Conditional Pel Replenishment and Rotation (symbol as proposed (NO BTH)) and the results without Rotation only (symbol as proposed (NO ROT)) respectively.

To eliminate the influence of the I-frame, since its decoded image is used as the first reference frame when performing motion compensation, the first frame of the test sequences for the two methods have been coded by the H.264 I-frame baseline under the same parameter values.

Table 3 Coding efficiency improvement by rotation operation.

Type	Frequency of appearance (%)	PSNR before rotation (dB)	PSNR after rotation (dB)	Entropy of quantized values (bits/pixel)	Entropy of overhead (bits/block)
$1: 0.0.0.0	33.71 %	43.178	43.178	0	0.529
$2: 0.0.0.1	5.177 %	42.089	42.089	1.461	0.221
$3: 0.0.1.0	5.619 %	39.637	41.936	1.503	0.233
$4: 0.0.1.1	3.725 %	42.723	42.723	1.582	0.177
$5: 0.1.0.0	4.735 %	39.259	43.065	1.494	0.208
$6: 0.1.0.1	3.535 %	42.983	42.874	1.541	0.170
$7: 0.1.1.0	2.146 %	37.738	43.842	1.407	0.119
$8: 0.1.1.1	3.788 %	42.685	42.685	1.627	0.179
$9: 1.0.0.0	4.987 %	38.212	42.159	1.471	0.216
$10: 1.0.0.1	1.199 %	39.817	44.686	1.327	0.077
$11: 1.0.1.0	3.220 %	38.816	43.126	1.517	0.160
$12: 1.0.1.1	3.283 %	40.879	43.771	1.583	0.162
$13: 1.1.0.0	3.283 %	37.402	43.296	1.586	0.162
$14: 1.1.0.1	4.735 %	41.478	44.205	1.586	0.208
$15: 1.1.1.0	3.914 %	38.221	43.638	1.605	0.183
$16: 1.1.1.1.	12.94 %	44.555	44.555	1.698	0.382
Full-screen average	100.000 %	41.190	43.183	0.626	3.385

The first twenty frames of foreman.cif CIF (352×288) and first ten frames of three CIF (bus, flower and highway) test video sequences from the YUV Video Sequences website (http://trace.eas.asu.edu/yuv/) were used. MC coding parameters (both methods are the same at this point) have been set as follows:

- Search range: 32 pixels
- Total number of reference frames: 1
- ME scheme: fast full search
- PSliceSearch8×8: 1 (used, all other types are not used)
- DisableIntraInInter: 1 (Disable Intra mode for inter slice)
- Rate-distortion-optimized mode decision: used

Besides these parameters, the threshold values for the proposed scheme have been adapted to the input frames.

3.1 Simulation of Prediction Errors

A simulation of the prediction errors results is shown first. By this comparison, we can see the distribution of errors obtained by the proposed scheme, and whether spatial correlation exists in the signal after motion compensation. Table 4 lists several
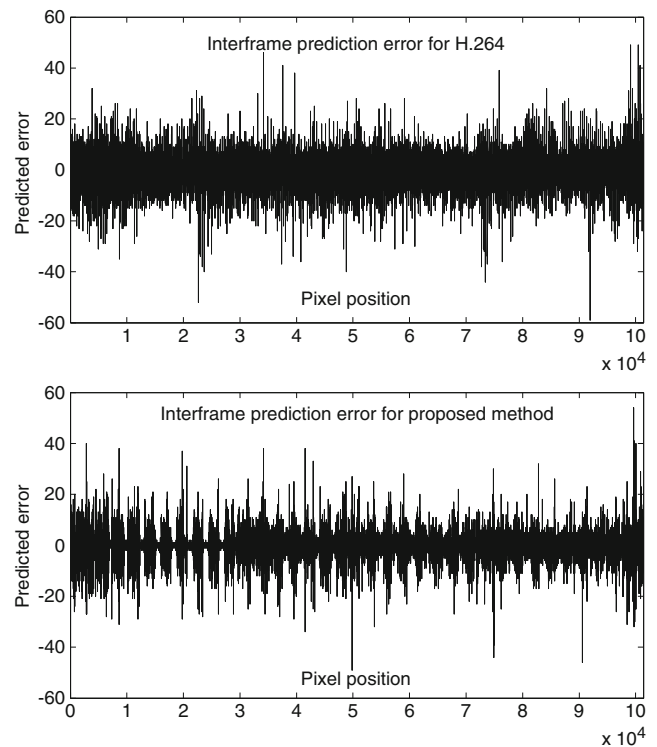
Figure 5 Simulation results of inter-frame prediction errors both proposed method and traditional causal prediction (strictly speaking, here the prediction errors of proposed is the output of interpolative prediction, not just the output of MC).

Table 4 Statistic values of residual signals.

Test image	Characteristic value					
	Number of 0's		Average error		Entropy	
	PM	Old	PM	Old	PM	Old
Foreman	35325	16862	9.66	15.48	3.249	3.829

Figure 6 Coding efficiency for *foreman.cif*, ρ_h=0.9726 ρ_v=0.9583.
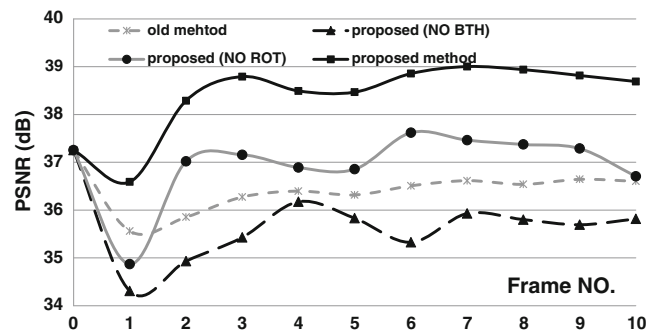
Figure 8 Coding efficiency for *flower.cif*, ρ_h=0.9087 ρ_v=0.7893.

statistical values for both methods that reflect the distribution of their respective residual signals.

In this table, "PM" stands for "proposed method"; "Old" stands for traditional causal prediction; entropy is calculated based on Shannon theory, and average error is the average value of two signal powers. The number of 0's indicates how many pixels have been accurately predicted.

Figure 5 shows the prediction errors for the first P-frame at each pixel position when using the first frame of *foreman.cif* as a test image. Here, the x-axis represents pixel position, and the y-axis represents the value of prediction error.

The distribution of the prediction errors for the proposed method is clearly more concentrated around zero than that for old methods, and the residual signal power is approximately 37.6 % lower under the proposed method. Therefore, we believe that our scheme provides improved coding efficiency if an appropriate quantization method is employed.

3.2 Comparison of Coding Efficiency

Next, we show the coding results for the proposed method. As stated before, the threshold values in our method must be changed for each frame, and these thresholds are obtained by a preliminary experiment. Since PTH and BTH are typically within a certain range, the results presented here are also limited. For this reason, the conditional pel replenishment used here has the potential for improvement.

Figures 6, 7, 8 and 9 show comparisons between the methods' coding efficiency when entropy of old method and proposed are almost the same for the respective test sequences of *foreman, bus*, *flower* and *highway*. ρ_h and ρ_v indicate the correlation of the test images in the horizontal and vertical directions, respectively.

In these plots, the horizontal axis expresses the frame number of the coded frames (0–20 frames for foreman.cif and 0–10 frames for other test sequences) and the vertical axes express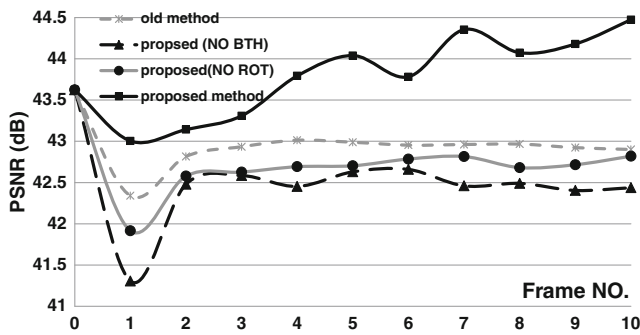 the entropy (bit/pixel; upper plot) and peak signal-to-noise ratio (PSNR; dB; lower plot) of each test image. Although the number of bits required by the proposed method (its entropy) is approximately equal to that of old method, here means H.264 baseline, PSNR for the proposed method is consistently higher (the average improvement is about 0.3–2 dB).

Table 5 shows the coding results when the proposed method is applied to the *bus* CIF video sequence as an example. Here, Q(S) is the average entropy for all pixels in the frame; and H(S) represents the overhead (motion vector) entropy.

3.3 Simulation of Motion Vectors

Table 6 lists the entropy of motion vectors in each frame for the two methods. Moreover, Fig. 10 shows the decoded frame obtained by the two methods for four test sequences, where the lines denote the motion vectors. The proposed method

Figure 7 Coding efficiency for *bus.cif*, ρ_h=0.8989 ρ_v=0.8414.
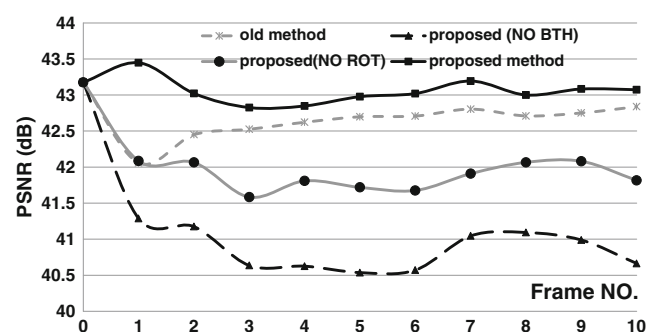
Figure 9 Coding efficiency for ρ_h=0.9964 ρ_v=0.8710.

Table 5 Coding results for proposed method.

P-frame number	THB	$H(S)$ (bits/pixel)	$Q(S)$ (bits/pixel)	PSNR(dB)
1	50	0.131	1.098	36.980
2	45	0.079	0.859	37.841
3	45	0.076	0.843	37.854
4	40	0.076	0.964	38.529
5	40	0.074	0.797	37.106
6	45	0.073	0.770	37.437

clearly contains less motion-vector information than causal method. Furthermore, the corresponding decoded frames without the motion vectors are shown in Fig. 11.

3.4 Effect of Adaptive Coding Between Intra/inter Modes

Figure 12 shows the simulation results of adaptive coding between the inter/intra coding modes. Insertion of the PinP configuration (Picture-in-Picture configuration, as shown in Fig. 13) is set to start at the first frame, and completes at the eighth frame. In Fig. 12, the PSNR of "inter mode only" indicates the value of the PSNR when all frames from the second sequence are forced to be codes in inter frame mode and the quantity of transmission bits is almost the same as that of the proposed adaptive scheme.

If the coding mode is fixed to be inter coding, as shown here, improvement of coding efficiency begins at the second P-frame, and after coding 2 frames, efficiency eventually approaches that of the adaptive system. The arrow in this figure represents how much improvement can be achieved by the proposed adaptive scheme.

On the other hand, because the adaptive intra/inter mode selection method with overhead information has been applied in the old methods, when using the PinP test sequence shown in Fig. 13, a significant decrease in PSNR does not occur.

Table 6 Entropy of motion vectors.

(Bus)			Highway	
Frame No.	H.264	PM	H.264	PM
1→2	0.131	0.048	0.079	0.080
2→3	0.080	0.047	0.077	0.082
3→4	0.074	0.046	0.074	0.083
4→5	0.091	0.075	0.074	0.090
5→6	0.072	0.072	0.081	0.087
6→7	0.070	0.074	0.080	0.092
7→8	0.077	0.076	0.085	0.085
8→9	0.074	0.135	0.075	0.087
9→10	0.078	0.098	0.079	0.084
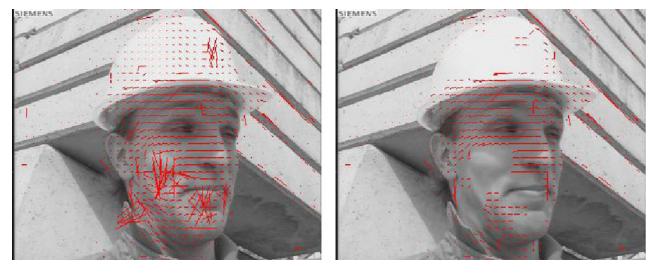
foreman.cif (4th frame)

bus.cif (1st frame)
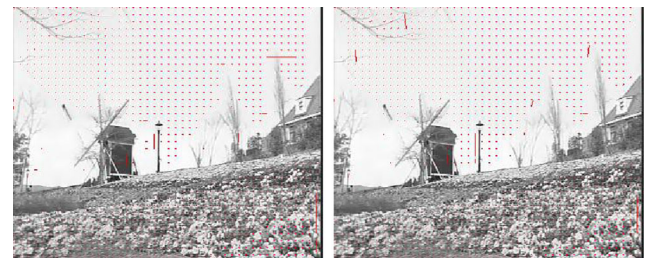
flower.cif (5th frame)

highway.cif (5th frame)

Figure 10 Comparison of motion vectors for the decoded frame obtained by traditional coding scheme (*left*) and the proposed scheme (*right*).

In addition, in the traditional coding scheme, such as H.264, because intra-frame coding mode can be selected at any time throughout the entire coding period, the overhead for every macro-block is approximately 0.01 bits/pixel more than in the proposed method. However, the disadvantage of the proposed method is that adaptive mode selection is not sufficiently flexible. In the future, the authors intend to improve upon this point.

4 Conclusion

In this paper, we have proposed a new inter-frame coding scheme in which the OT (e.g., a DCT) used in conventional
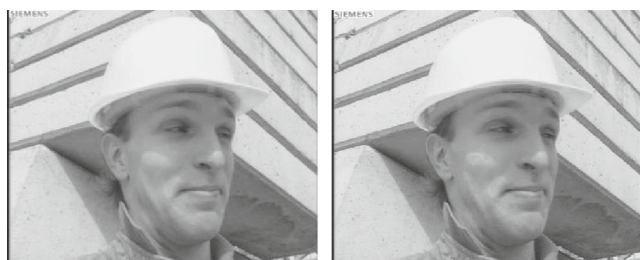
foreman.cif (4th frame)

bus.cif (1st frame)

flower.cif (4th frame)

highway.cif (4th frame)

Figure 11 Comparison of decoded frame obtained by traditional coding scheme (*left*) and the proposed scheme (*right*).
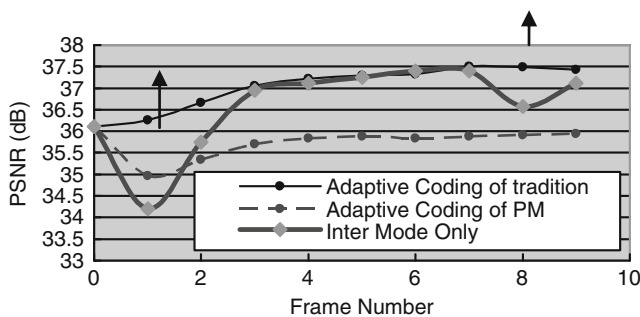
Figure 12 Coding efficiency for adaptive mode selection.

I frame   1stP frame

7th frame   8th P frame

Figure 13 PinP configuration.

hybrid coding schemes is replaced by a non-causal IP method. Application of this IP method can potentially reduce the amount of signal power a priori. Since IP methods utilize spatial correlations between pixels, we believe them to be more effective than OT (which merely perform a domain transform) for residual data that have been compressed after MC inter-frame prediction. However, combining our method with OT is also of great research interest, as shown above in Section 2 of this paper.

Accordingly, we have also introduced conditional pel replenishment to our scheme. Moreover, no additional overhead information is added by employing our method. Our model thus has the three characteristic features shown in Section 2.6. Simulation results of the proposed method in Section 3 showed that over four test sequences, the proposed scheme achieved a very considerable coding efficiency.

As a topic for future research, the conditional pel replenishment method utilized in our scheme could be improved, and this should be addressed first. Other areas that could also be explored are whether the proposed scheme can maintain high coding efficiency if the test sequence becomes large.

In conclusion, we have introduced a different approach to P-frame hybrid coding that utilizes spatial correlation for the MC residual signal. Since our hybrid video coding method has achieved high coding efficiency without employing OT, we have shown the feasibility of non-orthogonal transforms for effective coding.

Appendix

$$
\mathbf{C}_I^{inter} =
\begin{pmatrix}
1_{1st} & & & & & & & & & & & & & & \\
-\frac{1}{2} & 1_{2nd} & -\frac{1}{2} & & & & & & & & & & & & \\
 & \ddots & \ddots & \ddots & & & & & & & & & & & \\
 & & -\frac{1}{2} & 1_{7th} & -\frac{1}{2} & & & & & & & & & & \\
 & & & 1_{8th} & & & & & & & & & & & \\
-\frac{1}{2} & & & & 1_{9th} & & & & & -\frac{1}{2}^{17th} & & & & & \\
 & -\frac{1}{4}^{2nd} & & & -\frac{1}{4} & 1_{10th} & -\frac{1}{4} & & & -\frac{1}{4}^{18th} & & & & & \\
 & & -\frac{1}{4}^{7th} & & & -\frac{1}{4} & 1_{15th} & -\frac{1}{4} & & & -\frac{1}{4}^{23rd} & & & 0 & \\
 & & & -\frac{1}{2}^{8th} & & & 1_{16th} & & & & -\frac{1}{2}^{24th} & & & & \\
 & & & & & & & \cdot_{17\sim24} & & & & & & \\
 & & & & & & & & \cdot_{25\sim32} & & & & & \\
 & & & & & & & & & \cdot_{33\sim40} & & & & \\
 & & & & & & & & & & \cdot_{41\sim48} & & & \\
 & & & & & & & & & & & \cdot_{49\sim56} & & \\
 & & & & & & & & & & 1_{57th} & & & \\
 & & & & & & & & & & -\frac{1}{2} & 1_{58th} & -\frac{1}{2} & \\
 & & & & & & & & & & & \ddots & \ddots & \ddots \\
 & & & & & & & & & & & -\frac{1}{2}^{62nd}_{63th} & 1_{63th} & -\frac{1}{2}^{64th}_{63th} \\
 & & & & & & & & & & & & & 1_{64th}
\end{pmatrix}
\tag{A.1}
$$

References

1. Netravali, A. N., & Rabbins, J. D. (1979). Motion compensated television coding: part I. *Bell System Technical Journal, 58*(3), 631–670.

2. Jayant, N. S., & Noll, P. (1984). *Digital coding of waveforms*. Englewood Cliffs: Prentice Hall.

3. Wang, C., Kubota, A., & Hatori, Y. (2010). A new hybrid parallel intra coding method based on interpolative prediction. *28th Picture Coding Symposium (PCS) (December 7–10), Nagoya, Japan*, pp. 406–409.

4. Jain, A. K. (1975). Image coding via a nearest neighbors image model. *IEEE Transactions on Communications, COM-23*(3), 318–331, (Mar).

5. Hatori, Y. (1983). Optimal quantizing scheme in interpolative prediction. *The Journal of the Institute of Electronics, Information and Communication Engineers, J66-B* No.5, Tokyo.

6. Wang, C., & Hatori, Y. A parallel hybrid video coding method based on noncausal prediction with multimode. ISVC'11 Proceedings of the 7th international conference on Advances in visual computing – Volume Part II.

7. Fowler, J. E. (1998). Generalized threshold replenishment: an adaptive vector quantization algorithm for the coding of nonstationary sources. *IEEE Transactions on Image Processing, 7*, 1410–1424.

8. de Lima Filho, E. B., da Silva, E. A. B., de Carvalho, M. B., & Pinage, F. S. (2008). Universal image compression using multiscale recurrent patterns with adaptive probability model. *IEEE Transactions on Image Processing, 17*(4), 512–527.

9. Miaou, S.-G., & Yen, H.-L. (2001). Multichannel ECG compression using multichannel adaptive vector quantization. *IEEE Transactions on Biomedical Engineering, 48*(10), 1203–1207.

10. Fowler, J. E. (2000). Adaptive vector quantization for efficient zerotree-based coding of video with nonstationary statistics. *IEEE Transactions on Circuits and Systems for Video Technology, 10*(8), 1478–1488.

11. Shen, G., Zeng, B., & Liou, M.-L. (2003). Adaptive vector quantization with codebook updating based on locality and history. *IEEE Transactions on Image Processing, 12*(3), 283–295.

12. Wallace, G. K. (1991). The JPEG still picture compression standard. *Communications of the ACM, 34*(4), 31–44.

13. Le Gall, D. (1991). MPEG: a video compression standard for multimedia applications. *Communications of the ACM, 34*(4), 47–58.

14. Liou, M. (1991). Overview of the px64 kbps video coding standard. *Communications of the ACM, 34*(4), 60–63.

15. Ahmed, N., Natarajan, T., & Rao, K. R. (1974). Discrete cosine transform. *IEEE Transactions on Communications, COM-23*, 90–93.

16. Antonini, M., Barlaud, M., Mathieu, P., & Daubechies, I. (1992). Image coding using wavelet trans-form. *IEEE Transactions on Image Processing, 1*(2), 205–221.

17. Li, Z.-N., & Drew, M. S. (2004). *Fundamentals of multimedia*. New Delhi: Pearson Education.

18. Mori, S., Kubota, A., & Hatori, Y. (2010). Examination of hybrid coding method by interpolative prediction and DCT quantization. IEVC 2010, 2C-3, Nice, France, (Mar).

19. Shannon, C. E. (1948). a mathematical theory of communication. *Bell System Technical Journal, 27*, 623–656.

20. Jack, K. (2001). *Video demystified*. Mumbai: Penram International Publishing Pvt. Ltd.

21. Pratt, W. K. (1970).Karhumen-Loeve transform coding of images. Presented at the 1970 I.E. Int. Symp. Information Theory, June.

22. Pratt, W. K., Kane, J., & Andrews, H. C. (1969). Hadamard transform image coding. *Proceedings of the IEEE, 57*, 58–68.

23. Akihiko, M., & Tanaka, M. (1986). Image data compression method based on interpolative DPCM by area decomposition. *IEICE Transactions, J69-D*(3), 375–382.

24. Akihiko, M. (1992). An analysis and a reduction method of coding error in interpolative DPCM. *IEICE Transactions, J75-D-II*(9), 1565–1572.

25. Fujita, T., Kubota, A., & Hatori, Y. (2009). Adaptive quantization ordering in feedback quantization for interpolative prediction coding. *ITE Journal, 63*(11), 1652–1658.
26. Sehgal, A., Jagmohan, A., & Ahuja, N. (2004). Wyner-Ziv coding of video: an error-resilient compression framework. *IEEE Transactions on Multimedia, 6*(2), 249–258.
27. Pyun, J.-Y. (2005). Adaptive video redundancy coding for scene and channel adaptation over error-prone network. *IEEE Transactions on Consumer Electronics, 51*(3), 967–974.
28. van der Schaar-Mitrea, M., & de With, P. H. N. (2000). Hybrid compression of video with graphics in DTV communication systems. *IEEE Transactions on Consumer Electronics, 46*(4), 1007–1017.
29. H.264/14496-10 AVC reference software manual, Jan, 2009

Akira Kubota got his BE degree from Oita University in 1997 and ME from Tokyo University in 1999. In 2002, he received the doctor degree from the same university. He worked as an assistant professor in Department of Information Processing, Tokyo Institute of Technology from 2006 to 2009. Now, he is an Associate Professor of Chuo University. His research interests is about Multi-Focus images and free iris and focus image.
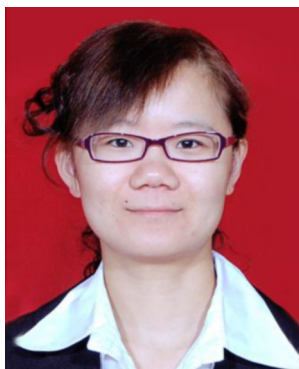
Cui Wang received her BE and ME degrees in Electronic and Information Engineering from Taiyuan University of Technology in 2005 and 2008 respectively. In 2013, she received the Doctor degree from Information Processing Department of Tokyo Institute of Technology and worked as a PHD in TIT in the same year. Her research interests include video compression, coding method for image & video signal, and image processing.

Yoshinori Hatori received his BE degree in engineering from Tokyo University in 1971 and joined KDDI Corporation in the same year. From 2003, he worked as a professor in Information Processing Department of Tokyo Institute of Technology. He interested in the research of high-efficiency coding method of image and video signal.