PARAMETER TRANSMISSION VIA ROI IN JPEG2000 FOR VARIABLE-COEFFICIENT INVERTIBLE DEINTERLACING

†*Takuma Ishida,* †*Jun Uchita,* †*Shogo Muramatsu,* †*Hisakazu Kikuchi, and* ‡*Tetsuro Kuge*

†Department of Electrical and Electronic Engineering, Niigata University
8050, Ikarashi 2-no-cho, Niigata, 950-2180, Japan
phone: +81-25-262-6746, fax: +81-25-264-2130, email: takumaro@telecom0.eng.niigata-u.ac.jp
‡Science & Technical Research, NHK
1-10-11 Kinuta, Setagaya-ku, Tokyo, 157-8510, Japan

ABSTRACT

A coefficient-parameter embedding method is presented for invertible deinterlacing with variable coefficients in the application to Motion-JPEG2000 (MJP2). Invertible deinterlacing, which the authors have developed before, can be used as a preprocess of frame-based video codec, such as MJP2, for interlaced videos. When the conventional field-interleaving is used instead, comb-tooth artifacts appear around edges of moving objects. On the other hand, the invertible deinterlacing technique allows us to suppress the comb-tooth artifacts and also to recover an original picture on demand. As previous works, the authors have developed a variable coefficient scheme with a motion detector, which realizes adaptability to local characteristics of given pictures. When applying this deinterlacing technique to an image codec, however, it is required to send coefficient parameters to receivers for original picture recovery. This work proposes a parameter-embedding technique and constructs a standard stream which consists both of picture data and parameters. The parameters are embedded into the first *LH* subband of the wavelet domain through the ROI (region of interest) function of JPEG2000, while maintaining the capability of comb-tooth suppression and quality recovery.

1. INTRODUCTION

Two formats of interlaced scanning and progressive scanning are in use for recording and displaying motion pictures [1–3]. The intraframe-based coding of interlaced pictures such as NTSC signals assumes field interleaving. Such a scheme can offer some advantages. Because an excellent still picture coding such as JPEG2000 (JP2) is directly applicable, its various options can be utilized for creating, editing and archiving video contents as well as in coping with various network environments and front-end terminals. Unfortunately the field interleaving causes horizontal comb-tooth artifacts around the boundaries of moving objects [4]. In the case of scalable transform-based coding such as Motion-JPEG2000 (MJP2) [5], quantization errors introduced in vertical high frequency components of the comb-tooth artifacts are annoyingly perceivable at low bitrates. To suppress the unfavorable artifacts, the intraframe-based coding system with a pre-filter was proposed [4]. It is shown to be effective in decoding videos at a certain target bit rates. Especially it is true for low bit-rate applications. Note that scalable codec systems should cover a wide range of decoding bit-rate. In high bit-rate decoding with this approach, however, the resolution of a picture is decreases due to lowpass filtering. To solve this problem, we developed an invertible deinterlacing technique with sampling density preservation as a preprocessing to scalable intraframe-based coding [6, 7]. We further developed variable-coefficient invertible deinterlacing so that the filter characteristics can be selected according to local properties of pictures [8, 9]. For the application to video codec systems, however, the variable-coefficient invertible deinterlacing has to send the coefficient parameters to the receiver for recovering the original pictures.

This work was in part supported by the Grand-in-Aid for Scientific Research No.14750283 from Society for the Promotion of Science and Culture of Japan.

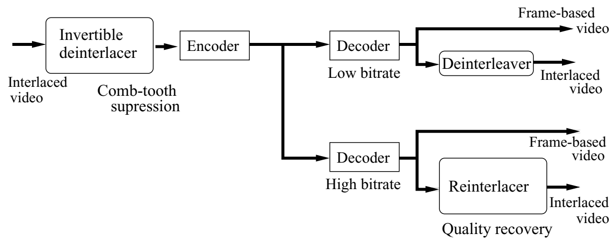


Figure 1: Intraframe-based coding system with deinterlacer.

In this work, we propose to embed the coefficient parameters to the MJP2 stream by using the ROI (Region of Interest) functionality. Since specified ROI shapes can be detected by decoder, the region where the deinterlacing is applied can be transmitted. As a result, encoders and decoders become able to share the information of parameter position and values. In our proposed method, the coefficient parameters are embedded into the LH_1 component because the comb-tooth artifacts strongly appears in this subband. It is investigated if we can keep the standard stream format of MJP2 without significant loss of the comb-tooth suppression and quality recovery capabilities even though the parameters are embedded.

2. REVIEW OF INVERTIBLE DEINTERLACING

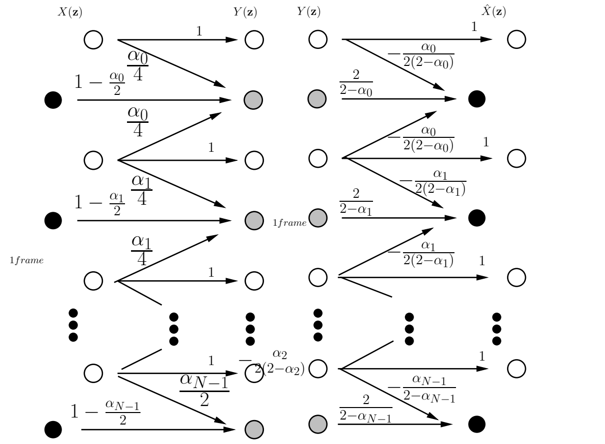
In this section, we introduce the variable-coefficient invertible deinterlacing as a preliminary.

2.1 Application Scenario [10]

Figure 1 shows an outline of our suggested codec system [10]. This system uses an invertible deinterlacer as a pre-filter. We suggest a codec system with the deinterlacing to support both of frame and field-based display. The comb-tooth artifacts can be suppressed for both of field and frame-based pictures at low bitrates, whereas the quality of field-based pictures can be maintained with reinterlacer at high bitrates. For low bit-rate decoding, a simple deinterleaver is performed to split fields from frames. Both of field and frame-based videos can be obtained from one standard code-stream on demand. Although producing a progressive scanning video given by doubling the sampling density via another deinterlacing for the reinterlaced result is beyond the scope of this work, it surely is possible.

2.2 Variable-Coefficient Filters

As found in the articles [8,9], we introduced the variable deinterlacing technique. The pair of the deinterlacing and reinterlacing filters



(a) Deinterlacing (b) Reinterlacing

Figure 2: Efficient implementation of deinterlacing with variable coefficients, where the symmetric extension method is applied.

as follows:

$$H_n(\mathbf{z}) = 1 + \left(1 - \frac{\alpha_n}{2}\right) z_{\mathrm{T}}^{-1} + \frac{\alpha_n}{4}(z_{\mathrm{V}}^{1} + z_{\mathrm{V}}^{-1}), \quad (1)$$

$$F_n(\mathbf{z}) = z_{\mathrm{T}}^{-1}\left\{\frac{2}{2-\alpha_n} + z_{\mathrm{T}}^{-1} - \frac{\alpha_n}{2(2-\alpha_n)}(z_{\mathrm{V}}^{1} + z_{\mathrm{V}}^{-1})\right\}, \quad (2)$$

where α_n is a parameter and $\mathbf{z}$ is a 3×1 vector which consists of variables in a 3-D z-domain, that is $\mathbf{z} = (z_{\mathrm{T}}, z_{\mathrm{V}}, z_{\mathrm{H}})^T$. The characteristics of these filters are controlled among temporal, vertical-temporal and vertical filters in the range of $0 \leq \alpha_n < 2$. In particular, for $\alpha_n = 1$, the coefficients of the pair are given by power of two, and become the same as ones shown in the articles [6, 7, 10, 11]. In addition, the deinterlacer becomes simple field interleaving for $\alpha_n = 0$, so that the original quality can be maintained.

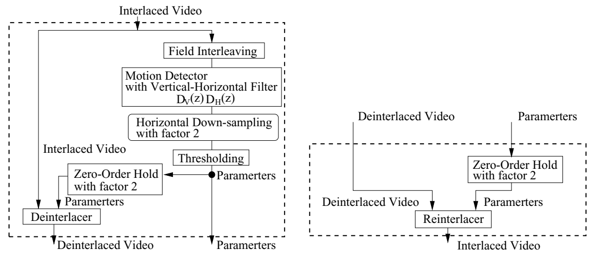
Figure 2 illustrates efficient operations with the filters in Eqs.(1) and (2). The white, black and gray circles indicate pixels on top line, bottom line of $X(\mathbf{z})$ and odd line of deinterlaced frame $Y(\mathbf{z})$, respectively. The odd line of deinterlaced frame pictures can be obtained by weighted sum of three lines with weights beside arrows. It is verified that our variable-coefficient method can be computed through the in-place implementation. Even if the value of α_n is varied for each sample on odd line, $X(\mathbf{z})$ can be recovered from $Y(\mathbf{z})$ as shown in Fig. 2. Note that the property of perfect reconstruction can be kept for this implementation independently from the choice of α_n.

2.3 Adaptive Control Method

The parameter α_{n} can have any value in the range of $0 \leq \alpha_{\mathrm{n}} < 2$. The value, however, should be transmitted to a decoder for reinterlacing, if the inverse process is desired. It is thus significant to limit the possible quantities for efficient transmission of α_{n}. In addition, the reduction of the computational complexity is another concern. To cope with these two practical requirements, we proposed to switch the value of α_{n} between 0 and 1 [8, 9].

To detect regions prone to yield comb-tooth artifacts, we suggested to apply a vertical high-pass filter prior to deinterlacing. The motion detection to predict comb-tooth artifacts was switched between 0 and 1 by a threshold value T. The invertible deinterlacing with variable coefficients avoids flickering by locally suppressing the comb-tooth artifacts, while keeping quality recovery by reinterlacing.

Simultaneous transmission of parameters decreases the bit-rate assigned to the picture data within a specified bit-rate. In a previous work, to reduce the coefficient parameters, we further proposed



(a) Deinterlacing (b) Reinterlacing

Figure 3: Process flow of parameter reduction method [12].

a parameter reduction method without significant loss of comb-tooth suppression capability. Figure 3 shows the flow graphs of the method, where a horizontal low-pass filter is added in the motion detector. As a result, the amount of parameters can be reduced and the quality of recovered pictures are improved compared with the original full parameter method at same bitrates. The details of this reduction method were shown in the article [12].

Note that it is still necessary to send the coefficient parameters to a receiver, and the simultaneous transmission of parameters is preferable to the separate transmission. In this work, we propose to embed the parameters into MJP2 stream by using the ROI functionality without significant loss of the performance.

3. PROPOSED ROI APPROACH

In this section, we propose to embed the coefficient parameters into ROI of MJP2 so that we can make all of data one standard bitstream.

3.1 Overview of ROI Maxshift Method

JPEG2000 supports ROI coding. The ROI function achieves non-uniform distribution of the image quality between a specified region and the background region. According to the ROI Maxshift method defined in JP2 part I, the background bit-planes are down-shifted below all of the ROI coefficients [5]. ROI can have any shape, which does not need to be transmitted to the decoder side. From these reasons, we propose to use the ROI shape for transmitting the position where the deinterlacing is applied.

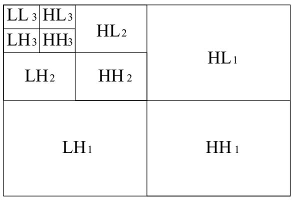
3.2 Choice of Target Subband

The ROI function can independently specify its shape in each subband. To suppress the influence on the image quality due to embedding coefficient parameters to ROI, we propose to embed the parameters to one subband domain. Note that the size of parameters is $W/2 \times H/2$ and fits to one of level-1 subband domain, where W and H denote the width and height of the original picture.

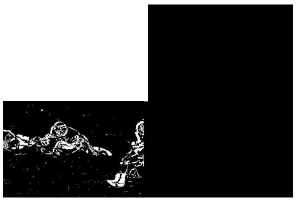
The coefficient parameters are determined by the output of a horizontal-lowpass and vertical-highpass filter, that is comb-tooth detection filter. Thus, the LH_1 subband coefficients should be treated carefully to recover the original picture. From this reason, coefficient parameters are embedded into subband LH_1 as ROI.

3.3 Progression Order

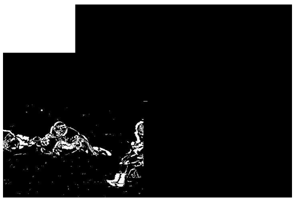
There are five different progression orders supported in JP2 [5]. The LRCP (Layer Resolution Component Position) progression is one of the main progression types. The LRCP progression arranges code-stream firstly in terms of layer and then of resolution. Since our invertible deinterlacer is meaningful for the SNR scalability, we here investigate only the LRCP progression case. When the LRCP progression is used, a problem arises. If only the ROI subband is given priority in the stream, the lower significant bits, that would be discarded if the ROI weren't used are regarded as important bits



(a) Notation of each subband.



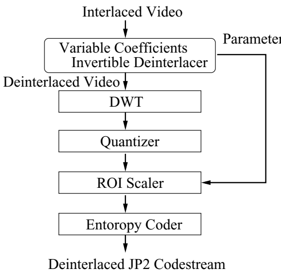
(b) Proposed mask for LL_1.



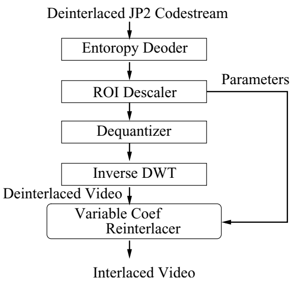
(c) Proposed mask for LL_2.

(d) Proposed mask for LL_3.

Figure 4: Notation of subband and ROI masks.



(a) encoding flow



(b) decoding flow

Figure 5: Proposed process flow.

for decoding. As a result, disagreeable pictures are yielded especially at low bitrates since other significant bits out of ROI subband coefficients are discarded.

To solve this problem, we suggest to add the entire components of the subband $LL_n (= \{LL_{n+1}, HL_{n+1}, LH_{n+1}, HH_{n+1}\})$ to ROI, where n denotes the depth of level. As examples, our proposed masks are given in Fig. 4. In addition, we suggest to replace the value of the least significant bit (LSB) in LL_n of ROI region to zero so as to preserve some bits in non ROI coefficients that would be pushed out if the replacement weren't applied. We verified that we can achieve a similar quantization through the expounded quantization supported in JP2. With regard to the choice of n we will discuss in the next section.

3.4 Processing Flow

Figure 5 (a) and (b) show our proposed process flows integrated into an encoder and a decoder of JP2. The coefficient parameters from deinterlacer are passed to the ROI Scaler, and are given to LH_1. All coefficients in LL_n are given as ROI. At a high bit-rate decoder, the information on LH_1 of the ROI mask is detected at ROI Descalers, and they are passed to the reinterlacer. Lastly, a picture is reconstructed. At a low bit-rate decoder, or a standard decoder, the reinterlacing process is skipped.

4. PERFORMANCE EVALUATION

To show the significance of our proposed approach, we demonstrate the comb-tooth suppression capability at low bitrates from view of subjective quality and show PSNRs as the quality recovery capability at high bitrates. In this evaluation, successive frame pictures of

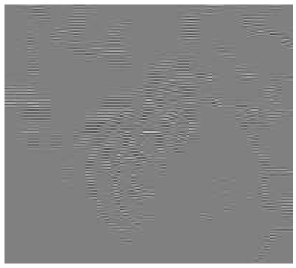
(a) Simple field interleaving

(b) Separate transmission [12]

(c) ROI transmission ($n = 1$)

(d) ROI transmission ($n = 2$)

(e) ROI transmission ($n = 3$)



(f) ROI transmission without ROI in LL

Figure 6: Decoded pictures at low bit-rate (0.1 bpp).

Football sequence (720×480 pixel, 8-bit grayscale) are used. Each frame picture is encoded at 2.0 bpp and then decoded at 2.0 and 0.1 bpp by JP2 [14].

4.1 Low Bitrate Performance

Figures 6 (a) - (f) show the decoded pictures at 0.1 bpp, where n denotes the depth of LL levels in which entire coefficients are maxshifted and the LSB in ROI are discarded. The simple field interleaving does not require any transmission of parameters.

The comb-tooth artifacts produced by the simple field interleaving are clearly perceived in Fig. 6 (a). In contrast, those artifacts are significantly suppressed by the invertible deinterlacer as shown in Figs. 6 (b) - (e). The deeper the entire maxshift operation is applied to LL_n components, the more blurry the result becomes. We can recognize that the maxshift operations to LL_1 or LL_2 are moderate at low bit-rate decoding in this experiment.

4.2 High Bitrate Performance

Figure 7 shows PSNRs of decoded pictures at high bitrates. The reinterlacing is used to recover the original quality at the decoder side. Here, the following methods are compared:

- Separate transmission without decimation [8, 9]
- Separate transmission with decimation [12]
- ROI transmission without LL maxshift

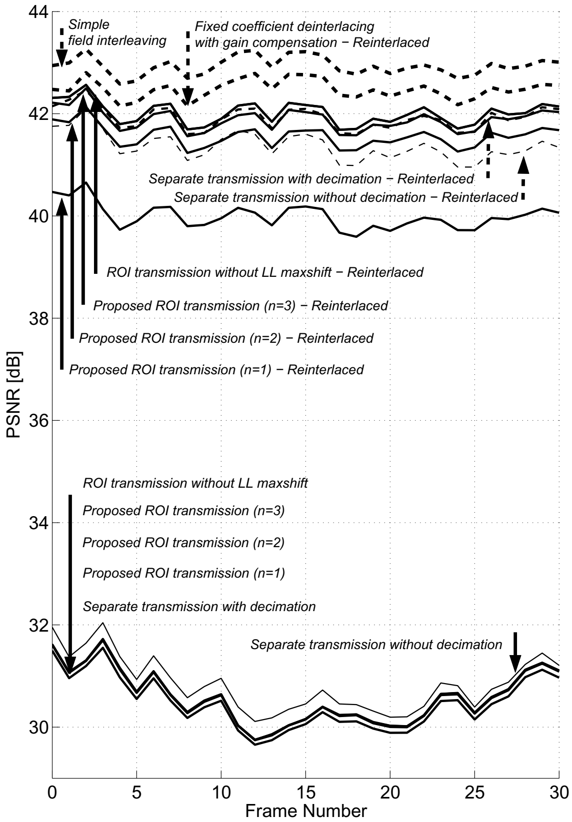


Figure 7: High bit-rate results, where our proposed quantization is not applied.

- Proposed ROI without LL quantization ($n = 1$)
- Proposed ROI without LL quantization ($n = 2$)
- Proposed ROI without LL quantization ($n = 3$)

In addition, results before reinterlacing are shown in Fig. 7. Quality recovery of reinterlacing can be verified. For reference, the following two schemes are also shown:

- Simple field interleaving
- Fixed coefficient deinterlacer with gain compensation [13]

In our proposed methods, PSNRs are improved as the depth n increases. When $n = 3$, it reaches to the result of separate transmission technique with parameter decimation. This tendency is opposite to that of the low bit-rate decoding. Actually, the result without LL maxshift outperforms the proposed methods, although pictures at low bitrates are not acceptable.

In addition, Figure 8 demonstrates the effect of our proposed quantization. Results of the LSB quantization in LL_n subband outperform those without quantization. We can find a tradeoff between the performances in low and high bit-rate decoding. The proposed method for $n = 1$ gives a good compromise in this experiment.

It should be noted that the field interleaving shows well result in the quality recovery. This technique is simple and requires no parameter transmission. The suppression capability at low bitrates is, however, inferior to the other methods. The fixed coefficient deinterlacing also performs well without transmission of parameters. This technique, however, applies the filtering process to whole of picture, thus still parts are not guarded.

5. CONCLUSIONS

We presented ROI embedding method for parameters of variable-coefficient invertible deinterlacing. By our proposal, it became possible to share coefficient parameters in one standard MJP2 codestream. To give appropriate pictures at low bitrates, we also pro-

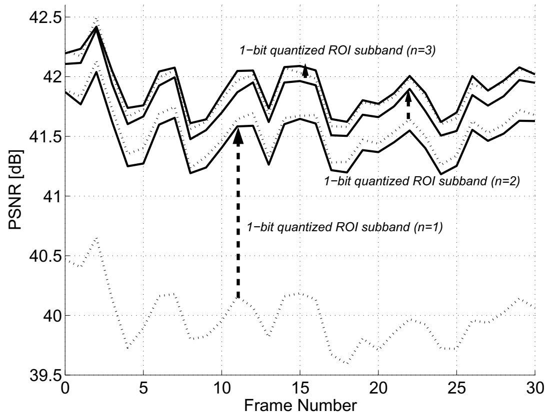


Figure 8: Effect of discarding LSB of ROI in LL

posed to set a certain LL subband region as ROI. It was shown that the depth of LL levels specified as ROI gives us a tradeoff with respect to performances between low and high bit-rate decoding. The demonstrated experiments did not apply neither of the DWT gain compensation nor any remedy to parameter loss caused by zero-value coefficients. As well, the ROI function is not necessarily the only solution to embed the parameters into standard code-streams. As future work, we will consider improving performances of parameter embedding by taking measures for these problems into account.

REFERENCES

[1] A. M. Tekalp, *Digital Video Processing*. Prentice Hall, Inc. 1995.
[2] Y. Wang, J. Ostermann and Y. Zhang,*Video Processing and Communications*. Prentice Hall, Inc. 2002.
[3] G. de Haan and E. B. Bellers, "Deinterlacing-an overview," in *Proc. IEEE*, vol. 86, pp. 1837-1857, Sept. 1998.
[4] T. Kuge, "Wavelet picture coding and its several problems of the application to the interlace HDTV and the ultra-high definition images," in *IEEE Proc. ICIP*, WA-P2.1, Sept. 2002.
[5] D. S. Taubman and M. W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practice*. Kluwer Academic Publishers, 2002.
[6] S. Muramatsu, T. Ishida and H. Kikuchi, "A Design Method of Invertible De-interlacer with Sampling Density Preservation," in *IEEE Proc. ICASSP*, vol.4, pp. 3277-3280, 2002.
[7] S. Muramatsu, T. Ishida, and H. Kikuchi, "Invertible Deinterlacing with Sampling Density Preservation: Theory and Design," in *IEEE Trans. on Signal Processing*, vol. 51, No. 9, pp. 2243-2356, Sept. 2003.
[8] T. Ishida, T. Soyama, S. Muramatsu, H. Kikuchi, T. Kuge, "A Lifting Implementation of Variable-Coefficient Invertible Deinterlacer with Embedded Motion Detector," IEICE Trans. on Fundamentals, vol. E86-A, no.8, pp. 1942-1948, Aug. 2003.
[9] T. Ishida, S. Muramatsu, H. Kikuchi, T. Kuge, "Invertible Deinterlacer with Variable Coefficients and Its Lifting Implementation," in *IEEE Proc. ICME*, No. IMSP-L5.3, Vol. III, pp. 105-108, 2003.
[10] H. Kikuchi, S. Muramatsu, T. Ishida, and T. Kuge, "Reversible conversion between interlaced and progressive scan formats and its efficient implementation," in *Proc. EUSIPCO*, No. 448 2002.
[11] T. Soyama, T. Ishida, S. Muramatsu, H. Kikuchi, "Lifting Architecture of Invertible Deinterlacing," IEICE Trans. on Fundamentals, Vol. E86-A, No.4, pp. 779-786, Apl. 2003.
[12] J. Uchita, T. Ishida, S. Muramatsu, H. Kikuchi, and T. Kuge, "Parameter Decimation for Invertible Deinterlacing with Variable Coefficients," in *Proc. ITC-CSCC2003*, pp. 925-928, 2003.
[13] T. Ishida, S. Muramatsu, J. Zhou, S. Sasaki, and H. Kikuchi, "DWT Gain Compensation of Motion-JPEG2000 for an Invertible Deinterlacer," in *IEEE Proc. ICIP*, No.TA-P1.12, 2003.
[14] Cannon, EPFL and Ericsson, "http://jj2000.epfl.ch"