A Parameter Decimation Technique for Variable-Coefficient Invertible Deinterlacing*

Jun UCHITA, Takuma ISHIDA†, Student Members, Shogo MURAMATSU††, Hisakazu KIKUCHI††, and Tetsuro KUGE†††, Members

SUMMARY In this paper, a coefficient-parameter reduction method is proposed for invertible deinterlacing with variable coefficients. Invertible deinterlacing, which the authors have developed before, can be used as a preprocess of frame-based motion picture codec, such as Motion-JPEG2000 (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 detection filter, which realizes adaptability to local characteristics of given pictures. When applying this deinterlacing technique to an image codec, it is required to send coefficient parameters to receivers for original picture recovery. This work proposes a parameter decimation technique and shows that this reduction approach can be achieved without significant loss of comb-tooth suppression capability and improves the quality at high bit-rate decoding.

key words: invertible deinterlacing, intraframe-based coding, scalable coding, variable coefficients, Motion-JPEG2000

1. Introduction

Interlaced scanning and progressive scanning are in use for recording, transmitting and displaying motion pictures [1]–[3]. For interlaced pictures, such as NTSC signals, an intraframe-based coding system usually employs field interleave so that some still picture coding system is directly applicable. Unfortunately, this process causes horizontal comb-tooth artifacts at edges of moving objects. In the case of scalable transform-based coding such as Motion-JPEG2000 (MJP2), the comb-tooth artifacts consisting of vertical high frequency components are enhanced by their quantization in wavelet transform domain, and those result in flickering around the edges of motion objects at low bit-rate decoding. To suppress the unfavorable comb-tooth artifacts, we proposed a pre-processing technique so that the standard decoding without extra processing can serve pictures of which comb-tooth artifacts have already been suppressed [4]. Especially, it is effective for low and middle bit-rate applications.

For high bit-rate applications, however, the resolution of pictures degrades since the pre-filter blurs them. As a previous work, to solve this problem, we developed invertible deinterlacing with sampling density preservation as a preprocess of scalable intraframe-based coding [5]–[11]. With this technique, we can suppress the comb-tooth artifacts, while maintaining the resolution recovery. Here, we use the term ‘invertible’ to indicate that an inverse system analytically exists and to distinguish our technique from the deinterlacing popularly used in advanced TV receivers, where lossless encoding is not assumed. The original invertible deinterlacing was, however, not necessarily suitable for the local properties of a given picture since the coefficients of the deinterlacing filters were fixed. Later, we further proposed invertible deinterlacing with variable coefficients, where coefficients of the filter vary according to a given picture [12], [13]. Compared with the fixed-coefficient deinterlacer, PSNR is improved as a result. In order to make coefficient parameters variable, however, there arises a problem how to transmit and record coefficient parameters.

This work deals with this problem and proposes a coefficient-parameter decimation technique. Our proposed method uses horizontal low-pass filter and decimator for reducing the amount of parameters. Lossless mode of JPEG2000 without wavelet transformation is used for compression of the coefficient parameters. It is shown that the parameter reduction contributes to improve the quality of pictures at high bit-rates, since we can assign more bits to the pictures, but the parameters, than the previous work.

This paper is organized as follows: Sect. 2 outlines the invertible deinterlacer with sampling-density preservation, describes an adaptive deinterlacing with a motion-detection filter, and summarizes the performances. Section 3 proposes a procedure of parameter reduction. Section 4 evaluates the performance of our proposed method, followed by conclusions in Sect. 5.

2. Review of Invertible Deinterlacing with Variable Coefficients

As a previous work, we proposed a deinterlacing technique that preserves sampling density and possesses the invertibility [6]. In this section, let us briefly review invertible deinterlacing with variable coefficients as a preliminary. Assume that the input array $X(z)$ is given as shown in Fig. 1.
2.1 Deinterlacing with Sampling Density Preservation

Figure 2 shows a basic structure of the deinterlacer with sampling density preservation, where $\mathbf{z}$ is a $3 \times 1$ vector which consists of variables in a 3-D $z$-domain. The upsampler converts the interlaced video array $X(\mathbf{z})$ into the non-interlaced one as shown in Fig. 3. The following $H(\mathbf{z})$ is a 3-D filter, which suppresses the comb-tooth artifacts. To preserve the sampling density, the processed video arrays are finally downsampled in temporal direction so that the deinterlaced array $Y(\mathbf{z})$ is obtained. This process can be regarded as a generalization of the conventional field interleaving [5]–[13].

2.2 Reinterlacing

For an intraframe-based codec system, the deinterlaced array $Y(\mathbf{z})$ is encoded, transmitted and then decoded frame by frame. Especially, for high bit-rate decoding, the interlaced video source is expected to be reconstructed. To achieve this, we have introduced the inverse process of the deinterlacer which we refer to as a reinterlacer, where the term ‘inverse’ does not imply ‘lossless’ by itself.

Figure 4 illustrates the basic structure of the reinterlacer. In the articles [5], [6], and [7], we showed that the input array $X(\mathbf{z})$ can be perfectly reconstructed from the deinterlaced array $Y(\mathbf{z})$ if $H(\mathbf{z})$ and $F(\mathbf{z})$ satisfy some conditions. In fact, the design problem of deinterlacing filters can be reduced to a design problem of multi-dimensional two-channel maximally-decimated filter banks [7]. Thus, term ‘invertible’ can be considered as the capability of the perfect reconstruction. Some examples of deinterlacing and reinterlacing filter pairs are shown in the articles [5]–[7].

2.3 Application Scenario of the Deinterlacer

We suggest an application scenario of the deinterlacer to intraframe-based scalable coding such as MJP2 as shown in Fig. 5. The invertible deinterlacer is used as a pre-filter. The comb-tooth artifacts are avoided at low bit-rate decoding, whereas the original fine resolution is maintained by the reinterlacer at high bit-rate decoding. For lossy decoding, the invertibility of the deinterlacer is still significant because some experimental results show that the quality is recovered more than 10 dB in PSNR by the reinterlacer after decoding at 2.0 bpp as will be shown in Sect. 4.

We have verified that flickering due to the comb-tooth artifacts can be suppressed for low bit-rate decoding. The quality of still parts is, however, blurred by the pre-filtering process. Actually, a simple temporal filter is preferable for still parts.

2.4 Variable Coefficient Processing

Local adaptability can be achieved by introducing a variable
Efficient implementation of deinterlacing with variable coefficients, where the symmetric extension method is applied. The white, black and gray circles denote top line, bottom line of $X(z)$ and odd line of $Y(z)$, respectively.

The variable-coefficient filtering has an in-place implementation as shown in Fig. 6, where the black, white and gray circles indicate pixels on bottom fields, top fields and the deinterlaced frame, respectively. Note that the perfect reconstruction property is verified by this implementation independently of values of $\alpha_n$.

2.5 Adaptive Control Method by Motion Detector

The parameter $\alpha_n$ can have any value in the range of $0 \leq \alpha_n < 2$. 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 $\alpha_n$. 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 $\alpha_n$ between 0 and 1 [12], [13].

In order to detect regions prone to be comb-tooth artifacts, we suggested to apply the following vertical-direction high-pass filter prior to deinterlacing:

$$D_V(z) = \frac{1}{4} z^{-1}_V + \frac{1}{2} - \frac{1}{4} z^{-1}_V.$$  (3)

Since the coefficients of the high pass filter are identical to those in the deinterlacer with $\alpha_n = 1$ except for their signs, the intermediate results can be shared by the following deinterlacing. The motion detection to predict comb-tooth artifacts is switched between 0 and 1 by using a threshold value $T$.

Figures 7(a) and (b) show an original frame picture and the detected areas where comb-tooth artifacts are prone to appear. Threshold $T = 2^4 = 16$ is used here. The black and white regions show still and moving areas, respectively. The invertible deinterlacing with variable coefficients avoids flickering by locally suppressing the comb-tooth artifacts, while it guarantees the perfect reconstruction by reinterlac-
ing at high bit-rate decoding.

It is, however, necessary to send the coefficient parameters to a receiver. When sending the parameters along with video data, the bit-rate assigned to image decreases. This work shows that these coefficient parameters can be reduced by a horizontal decimation process without significant loss of the comb-tooth suppression capability.

3. Coefficient Parameter Reduction

This section proposes an idea of parameter decimation for a variable-coefficient invertible deinterlacer. Invertible deinterlacing with variable coefficients suppresses comb-tooth artifacts at low bit-rate decoding, and enables us to perfectly recover the fine resolution in stillness regions.

3.1 Coefficient Parameter Reduction by Horizontal Decimation

We propose a method of coefficient parameter reduction by a horizontal decimation process. Figure 8 shows a basic flow of the proposed deinterlacing.

- Field Interleaving produces frame pictures so that the special detection filter can be applied.
- Motion Detector with Vertical-Horizontal Filter $D_V(z)\ D_H(z)$ detects horizontal-low and vertical-high frequency components.
- Horizontal Down-sampling with factor 2 [16] reduces the size of parameter array to half width by remaining every two column of the array and discarding the others.
- Thresholding quantizes the decimated parameters to binary values.
- Zero-Order Hold horizontally expands the decimated parameter array to the original size.
- Deinterlacer adaptively suppresses the comb-tooth artifacts by using the detected parameters.

The input interlaced video is analyzed by a motion detection filter. Then, the coefficient parameters are horizontally downsampling and quantized into binary values by thresholding. Decimated coefficient parameters are passed into the deinterlacer after expansion with zero-order holding. Here we suggests to use the following horizontal low-pass filter:

$$D_H(z) = \frac{1}{4}z^{-1} + \frac{1}{2} + \frac{1}{4}z^{1}.$$  \hspace{1cm} (4)

The above filter is added to avoid aliasing and detect only horizontal-low and vertical-high frequency components. Therefore, the comb-tooth detection becomes severe and the protection of stillness region are gained.

3.2 Reinterlacing

Figure 9 shows a basic flow of corresponding reinterlacing.

- Zero-Order Hold is identical to that appears in the forward process in Fig. 8.
- Reinterlacer adaptively recovers the original pictures by using the expanded parameter array.

The inverse system expands the parameters with zero-order holding in the same way as the deinterlacer. This process allows us to perfectly recover the original pictures from the deinterlaced pictures.

4. Performance Evaluation

In order to show the significance of our proposed coefficient-parameter reduction, let us evaluate the performance for both of low and high bit-rate decoding, where successive frame pictures of Football (720 × 480 pixel, 8-bit grayscale), Mobile&Calendar (720 × 480 pixel, 8-bit grayscale) and NewYork2 (720 × 480 pixel, 8-bit grayscale) sequences are used. Each frame picture is encoded at 2.0 bpp and then decoded at both of 2.0 and 0.1 bpp by JPEG2000 [22].

Performances for low bit-rate decoding: Figures 10, 11 and 12 show the decoded pictures at 0.1 bpp with field

(a) Original frame picture
(b) Motion detection result ($T = 16$)

Fig. 7 Example of detected comb-tooth artifacts.
interleaving, the conventional invertible deinterlacing with variable coefficients and the proposed method, respectively. Comb-tooth artifacts around the boundaries of moving objects are significantly suppressed by the invertible deinterlacers as shown in Figs. 11 and 12. In contrast, these artifacts produced by the simple field interleaving are clearly perceived in Fig. 10. The point here is that the coefficient-parameter reduction does not affect the comb-tooth artifact suppression capability at all.

**Performances for high bit-rate decoding:** PSNR is shown for the performance evaluation in case of high bit-rate decoding at 2.0 bpp. The simulation procedure is shown in Fig. 13. We use the lossless mode of JPEG2000 without wavelet transformation for the encode and decode processes of coefficient parameters $\alpha_n$. The bit-rate of a frame picture is calculated by Eq.(5) by taking the amount of parameters into account.

$$B_i = R - \frac{S_i}{F},$$

(5)

where $B_i$ is a bit-rate assigned for the encoding and decoding of the $i$-th frame picture, $R$ is a specified overall bit-rate, $S_i$ expresses the amount of coefficient-parameters for the $i$-th frame, and $F$ expresses the number of pixels per frame. Figures 14(a) and (b) show examples of coefficient-parameter bit planes. Figures 15, 16 and 17 plot PSNR of decoded pictures. At the encoder side, we compares among the following four methods:

- Field Interleaving
- Fixed Deinterlacing
- Conventional Variable Deinterlacing
- Proposed Variable Deinterlacing

At the decoder side, the following two methods are simulated:

- Deinterleaving: the inverse process of field interleaving
- Reinterlacing: the inverse process of deinterlacing

Since the field-interleaving does not require any coefficient parameters, the overall bit-rate $R$ of 2.0 bpp is directly used for the bit-rate $B_i$. For example, when coefficient parameters are taken into account for the conventional variable-coefficient filter, bit-rate of 1.8613 bpp in average is assigned for the compression of frame pictures in Football. On the other hand, 1.9401 bpp in average is assigned for the frame pictures in the proposed method. In
Mobile&Calendar, the amount of coefficient parameters increase because the entire screen moves. It is closer to the result of the fixed deinterlacing than those of the others. In the NewYork2, the amount of coefficient parameters are small because moving region is little. The proposed parameter decimation indicates almost the same result with the field interleaving.

Tables 1, 2 and 3 summarize each result. The proposed method is able to reduce the coefficient-parameters with the picture recovery at high bit-rate decoding and to keep the comb-tooth suppression capability at low bit-rate decoding. In other words, PSNR of recovered pictures is improved since the amount of data assigned to coefficient parameters is reduced.

5. Conclusion

In this paper, we proposed a coefficient parameter reduction method for the invertible deinterlacing with variable coefficients. It is achieved by a horizontal decimation process. It was verified that the data amount of coefficient parameters can be reduced without significant loss of the comb-tooth suppression capability, and the quality of recovered pictures are improved at the same bit-rate compared with the conventional method.

Acknowledgement

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.

References

Table 1  Performances: Football (Averages of 30 frames).

<table>
<thead>
<tr>
<th></th>
<th>Comb-tooth suppression (Low bit-rate)</th>
<th>Coding efficiency (High bit-rate)</th>
<th>Parameter size (High bit-rate)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field interleaving</td>
<td>Poor</td>
<td>42.94 [dB]</td>
<td>-</td>
</tr>
<tr>
<td>Fixed coefficients [5]-[9]</td>
<td>Good</td>
<td>40.93 [dB]</td>
<td>-</td>
</tr>
<tr>
<td>Conventional method [12], [13]</td>
<td>Good</td>
<td>41.37 [dB]</td>
<td>0.1387 [bpp]</td>
</tr>
<tr>
<td>Proposed method</td>
<td>Good</td>
<td>41.92 [dB]</td>
<td>0.0596 [bpp]</td>
</tr>
</tbody>
</table>

Table 2  Performances: Mobile&Calendar (Averages of 240 frames).

<table>
<thead>
<tr>
<th></th>
<th>Comb-tooth suppression (Low bit-rate)</th>
<th>Coding efficiency (High bit-rate)</th>
<th>Parameter size (High bit-rate)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field interleaving</td>
<td>Poor</td>
<td>34.99 [dB]</td>
<td>-</td>
</tr>
<tr>
<td>Fixed coefficients [5]-[9]</td>
<td>Good</td>
<td>33.95 [dB]</td>
<td>-</td>
</tr>
<tr>
<td>Proposed method</td>
<td>Good</td>
<td>33.86 [dB]</td>
<td>0.1121 [bpp]</td>
</tr>
</tbody>
</table>

Table 3  Performances: NewYork2 (Averages of 205 frames).

<table>
<thead>
<tr>
<th></th>
<th>Comb-tooth suppression (Low bit-rate)</th>
<th>Coding efficiency (High bit-rate)</th>
<th>Parameter size (High bit-rate)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field interleaving</td>
<td>Poor</td>
<td>48.18 [dB]</td>
<td>-</td>
</tr>
<tr>
<td>Fixed coefficients [5]-[9]</td>
<td>Good</td>
<td>45.90 [dB]</td>
<td>-</td>
</tr>
<tr>
<td>Conventional method [12], [13]</td>
<td>Good</td>
<td>46.34 [dB]</td>
<td>0.1267 [bpp]</td>
</tr>
<tr>
<td>Proposed method</td>
<td>Good</td>
<td>47.79 [dB]</td>
<td>0.0268 [bpp]</td>
</tr>
</tbody>
</table>


Shogo Muramatsu received B.E., M.E., and D.E. degrees in electrical engineering from Tokyo Metropolitan University in 1993, 1995, and 1998, respectively. From 1997 to 1999, he worked at Tokyo Metropolitan University. In 1999, he joined Niigata University, where he is currently an associate professor at Department of Electrical and Electric Engineering, Faculty of Engineering. During a year from 2003 to 2004, he is a visiting scholar at University of Florence, Italy. His research interests are in digital signal processing, multirate systems, image processing and VLSI architecture. Dr. Muramatsu is a member of IEEE (Institute of Electrical and Electronics Engineers, Inc.), IPSJ (Information Processing Society of Japan) and ITE (Institute of Image Information and Television Engineers).

Hisakazu Kikuchi received B.E. and M.E. degrees from Niigata University, Niigata, Japan, in 1974 and 1976, respectively, and Dr. Eng. degree in electrical and electronic engineering from Tokyo Institute of Technology, Tokyo, Japan in 1988. From 1976 to 1979 he worked at Information Processing Systems Laboratory, Fujitsu, Ltd., Tokyo. Since 1979 he has been with Niigata University, where he is a Professor at Department of Electrical Engineering. During a year of 1992 to 1993, he was a visiting scientist at the California, Los Angeles. His research interests include digital signal processing, image/video processing, and wavelets as well as spread spectrum communication systems. Dr. Kikuchi is a member of IEEE (Institute of Electrical and Electronics Engineers, Inc.), ITE (Institute of Image Information and Television Engineers of Japan), and Japan Society for Industrial and Applied Mathematics, Research Institute of Signal Processing, and SPIE.

Tetsuro Kuge received M.S. degree in electrical engineering from Doshisha University in 1978. Since 1984 he was engaged in research of signal processing and picture coding of HDTV-VTR and stereoscopic HDTV at NHK Science Technical Research Laboratories. Since 1995 he was engaged in development of election reporting system at News Department and in development of information processing systems for program production at Multimedia Department. Since 2000 he has been researching JPEG2000 related picture coding for HDTV program production at Science Technical Research Laboratories. He is a member of ITE Japan, and SC29/WG1 Japan committee.