# INVERTIBLE DEINTERLACING WITH VARIABLE COEFFICIENTS AND ITS LIFTING IMPLEMENTATION

*Takuma ISHIDA*†, *Shogo MURAMATSU*†, *Hisakazu KIKUCHI*† *and Tetsuro KUGE*††

† Dept. of Electrical and Electronic Eng., Niigata University
Niigata 950-2181, Japan
†† Science & Technical Research Laboratories, NHK
Tokyo 157-8510, Japan

## ABSTRACT

Invertible deinterlacing with variable coefficients is proposed to suppress comb-tooth artifacts caused by field interleaving of interlaced scanning video. A vertical highpass filter is applied to detect moving artifacts around boundaries of moving objects. The coefficients of a deinterlacing filter is varied depending on the motion intensity so that the deinterlacing filter may be matched to the local characteristics of moving pictures. Note that the deinterlacing filter is motion-adaptive and is time/translation-varying, while the deinterlacing is still kept to be invertible. The deinterlacing filter performance and its contribution to intraframe-based video coding are evaluated. In addition, since the processing of motion detection and a part of deinterlacing filtering can be shared, their efficient implementation is derived in the form of lifting popular in wavelets.

## 1. INTRODUCTION

Two formats of interlaced scanning and progressive scanning are in use for recording and displaying motion pictures [1, 2]. The intraframe-based coding of interlaced pictures such as NTSC signals assumes field interleaving, and such a scheme can offer some advantages. It is because an excellent still picture coding such as JPEG2000 is directly applicable and various options can be developed at creating, editing and archiving video contents as well as in their transcoding to cope with various network environments and front-end terminals. Unfortunately the field interleaving causes horizontal comb-tooth artifacts around the boundaries of moving objects [3]. In the case of scalable transform-based coding such as Motion-JPEG2000 (MJP2), quantization errors introduced by discarding vertical high frequency components in the comb-tooth artifacts are perceivable on low bit-rate decoding. To suppress the unfavorable artifacts, an intraframe-based coding system with a pre-filter was proposed [3]. Especially, it is effective for low bit rate applications.

In high bit-rate applications and some applications typical to transcoding, the resolution of a picture is not so much satisfactory because of lowpass filtering. The invertible deinterlacing with sampling density preservation had been developed to solve this problem as a preprocessing to scalable intraframe-based coding [4, 5].

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.

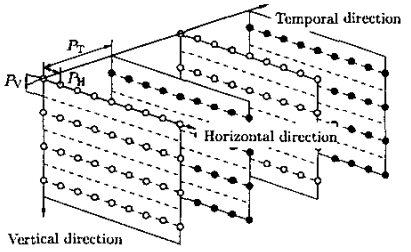



**Fig. 1**. Interlaced scanning with the sampling lattice $\mathcal{L}(\mathbf{V})$. The white and black circles are sample points on top and bottom fields, respectively.

This work describes further improvements offered by an additional treatment of local characteristics of a given picture, where 'local' implies locations both in time and space. It is the invertible deinterlacing with variable coefficients, and its system is a sequence of motion detection by a highpass filter followed by adaptive deinterlacing. A simple but efficient lifting implementation of the deinterlacing is also presented by sharing the identical operations between motion detection and deinterlacing. It is demonstrated that a deinterlacing pre-filter is effective and significant improvement in coding performance is also gained.

## 2. REVIEW OF INVERTIBLE DEINTERLACING

As a previous work, we proposed a deinterlacing technique that preserves sampling density and is invertible [4]. This technique is briefly reviewed as a preliminary. Assume that an input array $X(\mathbf{z})$ is given by a sampling lattice $\mathcal{L}(\mathbf{V})$ shown in Fig.1, where $\mathbf{V} = \begin{pmatrix} P_T & P_T & 0 \\ -P_V & P_V & 0 \\ 0 & 0 & P_H \end{pmatrix}$. $P_T$, $P_V$ and $P_H$ are the temporal period between successive fields, the vertical and horizontal sampling periods in a frame, respectively. $\mathcal{L}(\mathbf{V})$ is the set of sample points given by $\mathbf{Vn}$ for all $3 \times 1$ integer vectors $\mathbf{n}$, where $\mathbf{V}$ is a $3 \times 3$ non-singular matrix [6, 7].

### 2.1. Deinterlacing with Sampling Density Preservation

Figure 2 shows a basic structure of a deinterlacer with sampling density preservation, where circles denoted by $\uparrow \mathbf{Q}$ and $\downarrow \mathbf{R}$ are the upsampler with a factor $\mathbf{Q}$ and downsampler with a factor $\mathbf{R}$, respectively [1, 6, 8, 9]. The upsampler converts the interlaced





This paper was originally published in the Proceedings of the 2003 IEEE International Conference on Acoustics, Speech, & Signal Processing, April 6-10, 2003, Hong Kong (cancelled). Reprinted with permission.

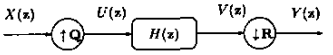



**Fig. 2**. Basic structure of a deinterlacer with sampling density preservation.

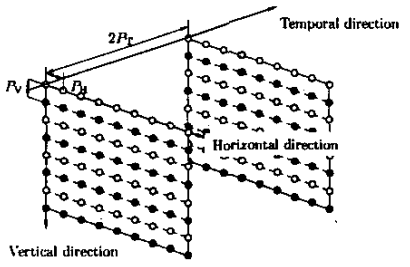



**Fig. 3**. Sampling lattice $\mathcal{L}(\mathbf{V}')$ of deinterlaced array with temporal decimation.

video array $X(\mathbf{z})$ into a non-interlaced video array. The sampling factor $\mathbf{Q}$ is $\begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ for the sampling matrix $\mathbf{V}$. $H(\mathbf{z})$ is a 3-D filter. It removes imaging components caused by upsampling and avoids aliasing to be caused by downsampling. The ratio has to be two to preserve the same density as the original. To keep the lattice orthogonality, one choice of $\mathbf{R}$ is given by $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. In this case, $V(\mathbf{z})$ is temporally downsampled. Figure 3 shows the sampling lattice $\mathcal{L}(\mathbf{V}')$ of the deinterlaced array $Y(\mathbf{z})$, where $\mathbf{V}' = \mathbf{V}\mathbf{Q}^{-1}\mathbf{R} = \begin{pmatrix} 2P_T & 0 & 0 \\ 0 & P_V & 0 \\ 0 & 0 & P_H \end{pmatrix}$.

### 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 in perfect. To do this, we introduced its inverse converter i.e. *reinterlacer*. Figure 4 illustrates the basic structure of a reinterlacer. The interlaced array $X(\mathbf{z})$ is perfectly reconstructed from the deinterlaced array $Y(\mathbf{z})$, if $H(\mathbf{z})$ and $F(\mathbf{z})$ satisfy a set of conditions. [4]

### 2.3. Intraframe-based Coding System with a Deinterlacer

An application scenario of intraframe-based scalable coding such as MJP2 [5] can be suggested as shown in Fig. 5, where an invertible deinterlacer is used as a pre-filter. The comb-tooth artifacts are avoided on low bit-rate decoding, whereas fine resolution is maintained by a reinterlacer on high bit-rate decoding. A pair of

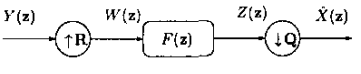



**Fig. 4**. Basic structure of a reinterlacer.

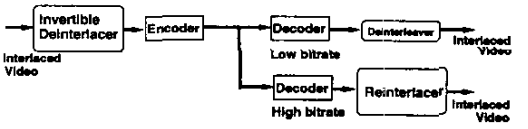



**Fig. 5**. Intraframe-based coding system with a deinterlacer.

filters for deinterlacing and reinterlacing can be as follows [4].

$$H(\mathbf{z}) = 1 + \frac{1}{2}z_T^{-1} + \frac{1}{4}(z_V^1 + z_V^{-1}). \quad (1)$$

$$F(\mathbf{z}) = z_T^{-1}\left\{2 + z_T^{-1} - \frac{1}{2}(z_V^1 + z_V^{-1})\right\}. \quad (2)$$

These filters have the following properties: i) normalized amplitude to keep brightness, ii) regularity to avoid the checkerboard effect [10, 11], and iii) vertical symmetry to afford the symmetric border extension [10,12]. The deinterlacing filter $H(\mathbf{z})$ works well for avoiding comb-tooth artifacts in low bit-rate decoding. However, the detailed quality at fine still areas is degraded by filtering. For those still parts, a simple temporal filter is preferable rather than Eq.(1).

## 3. VARIABLE-COEFFICIENT PROCESSING

We are going to vary filter coefficients depending on the local characteristics of given pictures. Problems are how to choose the coefficients and how to maintain the invertibility.

### 3.1. Variable-Coefficient Filters

We propose the deinterlacing and reinterlacing filters as follows.

$$H_n(\mathbf{z}) = 1 + (1 - \frac{\alpha_n}{2})z_T^{-1} + \frac{\alpha_n}{4}(z_V^1 + z_V^{-1}), \quad (3)$$

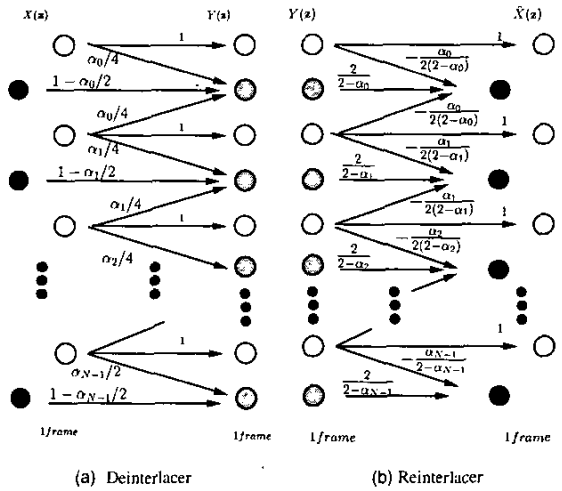
$$F_n(\mathbf{z}) = z_T^{-1}\left\{\frac{2}{2-\alpha_n} + z_T^{-1} - \frac{\alpha_n}{2(2-\alpha_n)}(z_V^1 + z_V^{-1})\right\}, \quad (4)$$

where $\alpha_n$ is a parameter in the range of $0 \leq \alpha_n < 2$. Different filtering modes are selectable among temporal, vertical-temporal and vertical filters. In particular, if $\alpha_n = 1$, the transfer function of the filter is identical to that given by Eq.(1). Also, if $\alpha_n = 0$, the deinterlacer reduces to a simple field-interleaver so that fine resolution is maintained.

The variable-coefficient filtering can have an in-place implementation as shown in Fig.6, where black, white and gray circles indicate pixels on bottom fields, top fields and a deinterlaced frame, respectively. Note that the perfect reconstruction property is incorporated in this implementation and it is independent from the values of $\alpha_n$.

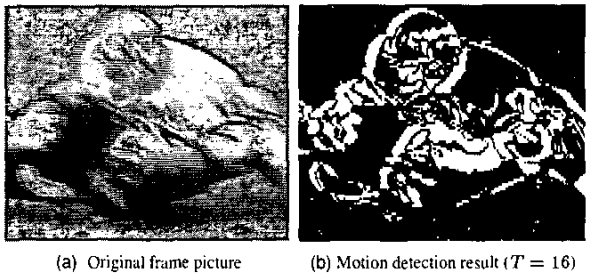
### 3.2. Adaptive Control Method

The parameter $\alpha_n$ can have any value in the range of $0 \leq \alpha_n < 2$. Its value should be provided to a decoder for reinterlacing, if high bit-rate decoding is desired. It is thus significant to limit its actual values for efficient transmission of $\alpha_n$. Also, the reduction in computational complexity is another concern. To cope with these



(a) Deinterlacer (b) Reinterlacer

**Fig. 6**. In-place implementation of deinterlacing with variable coefficients. The symmetric extension has been applied.



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

**Fig. 7**. Example of motion detection.

two realistic requirements, the value of $\alpha_n$ is switched between 0 and 1.

In order to detect comb-tooth artifacts to be possible to appear, a vertical-orientation highpass filter $D(\mathbf{z}) = \frac{1}{2}z_T^{-1} - \frac{1}{4}(z_V^1 + z_V^{-1})$ is applied prior to deinterlacing. Since the coefficients of the highpass filter are identical to those in Eq.(1) except for their signs, the intermediate computations can be shared by this highpass filtering and deinterlacing to follow. The motion detection to predict comb-tooth artifacts is performed as follows.

1 Compute an output frame $I_{x,y}$ of the highpass filter
2 If $|I_{x,2n+1}| \leq$ threshold $T$, set $\alpha_n = 0$
3 Otherwise, set $\alpha_n = 1$

When $T = 0$, this system results in fixed-coefficient deinterlacing. When $T \geq 128$ for 8-bit grayscale images, it reduces to just simple interleaving. Figure 7 (a) and (b) show an original frame picture and the detected areas where comb-tooth artifacts are prone to appear, where $T$=$2^4$=16 is used. The black and white regions show still and moving areas, respectively.

### 3.3. Lifting Implementation

As has been previously explained, a simplified variable-coefficient deinterlacing system can have an efficient implementation by computation sharing. In particular, the lifting structure can be developed for combining a sequence of computations in highpass filtering and deinterlacing.

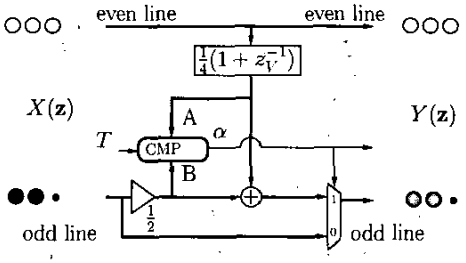



**Fig. 8**. Lifting implementation of deinterlacing embedded with motion detection.

Figure 8 illustrates our proposed implementation, where CMP indicates a comparator that calculates the selection signal $\alpha$ that actually controls $\alpha_n$, as follows.

$$\alpha = \begin{cases} 0, & |A - B| \leq T \\ 1, & |A - B| > T. \end{cases} \quad (5)$$

## 4. PERFORMANCE EVALUATION

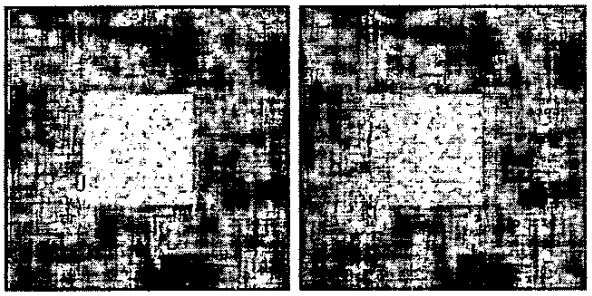
***Performance as Pre-filter*:** At first, the variable-coefficient deinterlacing is compared with the conventional field interleaving and fixed-coefficient deinterlacing in terms of pre-filtering performance. Figure 9 (a) shows an artificial test frame picture. The background is given by a separable first-order AR process ($\rho = 0.95$, 256×256 pixels), and a square is moving to the right in the velocity of 16 pixels per frame. The procedure of performance evaluation is as follows.

1 Make an interlaced image $I_s$ by decimating the test frame sequence of Fig.9 (a) by **Q**
2 Deinterlace $I_s$ for several threshold values
3 Calculate PSNR

Figure 9 (b) shows the deinterlaced frame picture. Figure 10 shows the PSNR dependency against $T$. Two end points at $T = 0$ and 128 correspond to the plain field interleaving and fixed-coefficient deinterlacing, respectively. It is observed that a single peak exists around $T = 12$ in the variable-coefficient deinterlacing.

***Performance for Low Bit-rate Decoding*:** A JPEG2000 encoder [13] is applied to every frame picture of *football* (720×480 pixels per frame, 8-bit grayscale) and the coding bit-rate is 0.1 bpp for low bit-rate applications. Figure 11 (a) and (b) show the decoded pictures with field interleaving and the proposed method, respectively. Comb-tooth artifacts around the boundaries of moving objects have been significantly suppressed in (b) of the figure. It is hard to perceive the flickering artifacts as a moving picture. In contrast, the comb-tooth artifacts produced by simple interleaving are strong as found in (a) and are perceived in flickering artifacts as a moving picture. In fact it is very annoying for low bit-rate encoding/decoding.

***Performance for High Bit-rate Decoding*:** The *football* sequence is again used for an experiment. Every deinterlaced frame picture has been encoded at 2.0bpp and then decoded at 2.0bpp by



(a) Original frame picture. (b) Deinterlaced picture ($T = 16$).

**Fig. 9**. Test picture produced by a separable first-order AR process ($\rho = 0.95$).

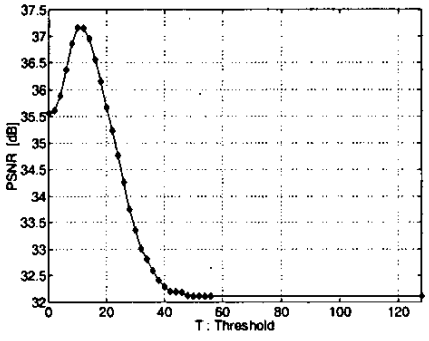



**Fig. 10**. PSNR dependency against $T$.

JPEG2000. Since the experiment assumes high bit-rate applications, reinterlacing has been applied to every decoded picture. Figure 12 is the PSNR plots of decoded pictures.

The variable-coefficient deinterlacing is involved with several threshold values among $0 \leq T \leq 128$, and its PSNR performance is surely superior to that in fixed-coefficient deinterlacing with $T = 0$. It is observed that PSNR is improved as the threshold value increases. It is worth to note that deinterlacing capability to suppress comb-tooth artifacts will deteriorate as $T$ grows high. Since $T = 128$ implies the system reduces to the field interleaving, no effect for artifact suppression is expectable in such a system. The plain field interleaving corresponding to $T = 128$ shows the highest level in PSNR, but the subjective quality offered by this system is worse and unfavorable; this statement must be

(a) Plain interleaving ($T = 128$) (b) The proposed method ($T = 16$)

**Fig. 11**. Low bit-rate decoding (0.1bpp)

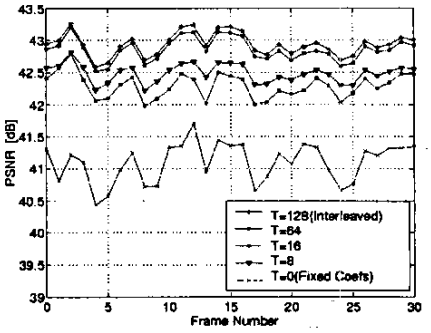



**Fig. 12**. Coding performance at 2.0bpp.

simply justified by the problem itself why deinterlacing is desirable. In contrast, the proposed method makes it possible to offer a trade-off among deinterlacing capability, picture resolution, and coding performance.

## 5. CONCLUSION

In this work, invertible deinterlacing with variable coefficients has been proposed. The coefficients of a deinterlacing filter is varied to fit the local characteristics of moving pictures. A simplified switchable-coefficient version and an efficient hardware implementation scheme having in-place computation have been also derived. Satisfactory performances have been demonstrated, and a desirable trade-off among comb-tooth artifact reduction, picture resolution and coding/decoding performance is possible.

## 6. REFERENCES


[1] A. Murat Tekalp, *Digital Video Processing*, Prentice Hall, Inc., 1995.

[2] G. de Haan and E. B. Bellers, "Deinterlacing-an overview," *Proc. IEEE*, vol. 86, pp. 1837–1857, Sept. 1998.

[3] T. Kuge, "Wavelet picture coding and its several problems of the application to the interlace HDTV and the ultra-high definition images," *IEEE Proc. of ICIP*, WA-P2.1, Sept. 2002.

[4] S. Muramatsu, T.Ishida, and H. Kikuchi, "A design method of invertible deinterlacer with sampling density preservation," *IEEE Proc. of ICASSP*, vol. 4, pp. 3277–3280, May 2002.

[5] H. Kikuchi, S. Muramatsu, T. Ishida, and T. Kuge, "Reversible conversion between interlaced and progressive scan formats and its efficient implementation," *Proc. of EUSIPCO*, no. 448, Toulouse, 2002.

[6] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice Hall, 1993.

[7] S. Muramatsu and H. Kiya, "Multidimensional parallel processing methods for rational sampling lattice alteration," *Proc. IEEE ISCAS*, vol. 1, no. 3, pp. 756–759, Apr. 1995.

[8] T. Chen and P.P. Vaidyanathan, "Recent developments in multidimensional multirate system," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 3, no. 2, pp. 116–137, Apr. 1993.

[9] Y. Wang, J. Ostermann, and Y. Zhang, *Video Processing and Communications*, Prentice-Hall, 2002.

[10] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, 1996.

[11] Y. Harada, S. Muramatsu, and H. Kiya, "Multidimensional multirate filter and filter bank without checkerboard effect," *Proc. of EUSIPCO*, pp. 1881–1884, Sept. 1998.

[12] H. Kiya, K. Nishikawa, and M. Iwahashi, "A development of symmetric extension method for subband image coding," *IEEE Trans. Image Processing*, vol. 3, no. 1, pp. 78–81, Jan. 1994.

[13] Cannon, EPFL and Ericsson, "http://jj2000.epfl.ch"