Using enhancement data to deinterlace 1080i HDTV

Using Enhancement Data to Deinterlace 1080i HDTV

Andy L. Lin[a], Jae S. Lim[b]
[a]Stanford University, 450 Serra Mall Stanford, CA 94305
[b]Research Lab of Electronics, Massachusetts Institute of Technology, 77 Massachusetts Avenue., Cambridge, MA USA 02139

ABSTRACT

When interlaced scan (IS) is used for television transmission, the received video must be deinterlaced to be displayed on progressive scan (PS) displays. To achieve good performance, the deinterlacing operation is typically computationally expensive. We propose a receiver compatible approach which performs a deinterlacing operation inexpensively, with good performance. At the transmitter, the system analyzes the video and transmits an additional low bit-rate stream. Existing receivers ignore this information. New receivers utilize this stream and perform a deinterlacing operation inexpensively with good performance. Results indicate that this approach can improve the digital television standard in a receiver compatible manner.

Keywords: Deinterlace, receiver compatible, 1080i, 1080/60/IS, HDTV, motion adaptive deinterlacing, enhancement data

1. INTRODUCTION

The interlaced scan (IS) transmission format has historic roots. In order for modern progressive scan (PS) devices to display IS format video, the IS format video must be properly deinterlaced. Over the years, there have been many deinterlacing algorithms. Since these algorithms are typically performed at the receiver, they are either low in quality and computationally inexpensive, or reasonable in quality and computationally expensive.

We propose the following receiver compatible approach. On the transmitter side, the system analyzes the source video and transmits an additional low bit-rate enhancement stream, which contains useful information for deinterlacing the IS video stream. New receivers decode both streams. With help from the enhancement stream, these new receivers are able to deinterlace IS video with reasonable quality, while being computationally inexpensive. Old receivers ignore the enhancement stream and decode *only* the IS video.

We implemented a specific receiver compatible system for the 1080/60/IS transmission format and compared its performance with that of traditional deinterlacing algorithms. Results indicate that the receiver-compatible system exhibits better performance than traditional deinterlacing algorithms, while being computationally inexpensive. The proposed receiver compatible system is an example of improving 1080/60/IS transmitted television quality in a receiver compatible manner.

2. BACKGROUND

2.1 HDTV Background

In 1987, the Federal Communications Commission (FCC) initiated an advanced television standardization process. The process led to the U.S. Digital Television Standard in 1996. The standard allows multiple video formats to be transmitted through the airwaves. The two popular high definition formats are 720p (60 frames/sec; 720/60/PS), and 1080i (60 fields/sec; 1080/60/IS). The progressive video format does not require deinterlacing, but the 1080/60/IS resolution format must be deinterlaced in order to be displayed on high definition television receivers.[1, 2]

2.2 Interlaced Scan

Video can be displayed in different scan modes. In progressive scan (PS) video, all pixels from each frame are displayed. In interlaced video (IS), every other line of video is displayed for each field. Thus, only half the pixels of an original

Applications of Digital Image Processing XXXIII, edited by Andrew G. Tescher,
Proc. of SPIE Vol. 7798, 77980B · © 2010 SPIE · CCC code: 0277-786X/10/$18
doi: 10.1117/12.859876

frame are displayed. For odd fields, only the pixels on odd-numbered lines of a frame are displayed. Similarly, for even fields, only the pixels on even-numbered lines are displayed. Interlaced scan often leads to line crawl and twitter video artifacts. These artifacts are worsened on modern PS displays especially if input video is not deinterlaced.

2.3 Traditional Deinterlacing Algorithms

Many deinterlacing algorithms can be classified into intra-field and inter-field algorithms. Intra-field algorithms only extrapolate data from spatial neighbors. Linear interpolation is a simple intra-field deinterlacing algorithm. Missing pixels are approximated by averaging pixels that are directly above and below the pixel of interest. Edge dependent deinterlacing algorithms are a class of intra-field deinterlacing algorithms that improve upon linear interpolation by averaging along dominant edge directions.[3] The Martinez-Lim deinterlacing algorithm is an edge dependent deinterlacing algorithm that offers high image quality while maintaining computational efficiency.[10]

Inter-field deinterlacing algorithms interpolate values for missing pixels by considering data from other fields. Whereas inter-field deinterlacing algorithms may use pixel values within the current frame, strictly inter-field deinterlacing algorithms interpolate missing pixels by considering data *only* from other fields. The simplest strictly inter-field deinterlacing algorithm is forward-field repetition (FFR), which repeats the previous field forwards in time to fill missing lines. Although FFR can reproduce theoretically perfect results for stationary frames, this algorithm produces jagged mice-teeth or field-tearing artifacts in moving frames.

Motion adaptive deinterlacing algorithms, a more advanced class of inter-field deinterlacing algorithms, switch between an intra-field algorithm for some areas of video and a strictly inter-field algorithm for other areas of video. Due to problems with strictly inter-field algorithms for moving areas of video, adaptive deinterlacing algorithms often use intra-field deinterlacing algorithms for moving areas and strictly inter-field algorithms for stationary areas.[4]

There are a series of inter-field deinterlacing methods which are more computationally expensive than motion adaptive methods. Motion compensated deinterlacing algorithms take advantage of inter-field correlation by using block-matching techniques to compute motion vectors of video blocks. These motion vectors model the changes between frames by estimating the motion for each block. Many different motion compensated deinterlacing algorithms exist, but all are computationally expensive. Moreover, motion compensated deinterlaced video often have block artifacts near the boundaries of moving objects. [5, 9, 11]

3. PROPOSED SYSTEM

3.1 Receiver-compatible Enhancement

It is difficult to modify a television standard after it has been well-established and many television receivers are already in use. One method to improve television quality without establishing a new standard is by employing a receiver compatible enhancement system. Under a receiver compatible system, the broadcaster can transmit the enhancement information over a separate bit-stream. There are two streams: the conventional video stream and the enhancement stream. Unaware of this enhancement stream, existing HDTV sets ignore the enhancement stream and decode the conventional video stream. Newer HDTV sets not only decode the conventional video stream, but also decode the enhancement stream and display superior HDTV video. See Figures 1, 2, and 3 for a comparison between the receiver compatible and the traditional transmission arrangements.[6, 7, 14]

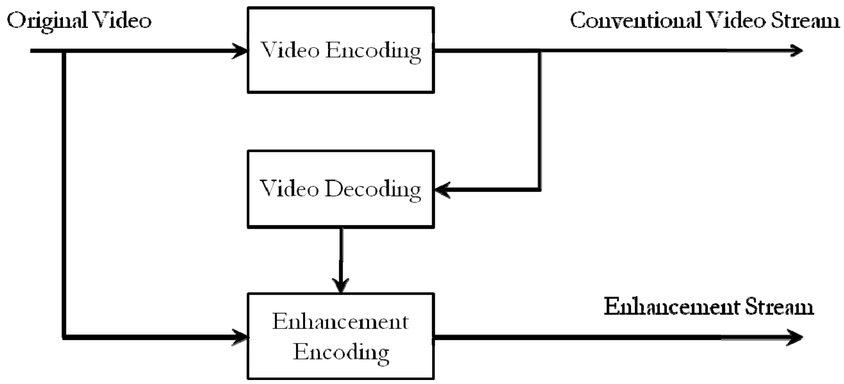

Figure 1: Receiver compatible transmitter block diagram.

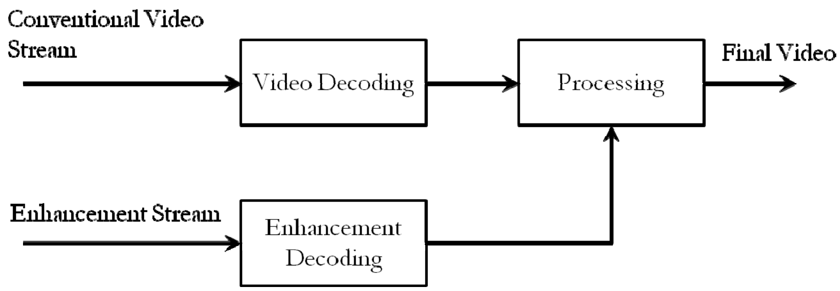

Figure 2: Receiver compatible receiver block diagram.

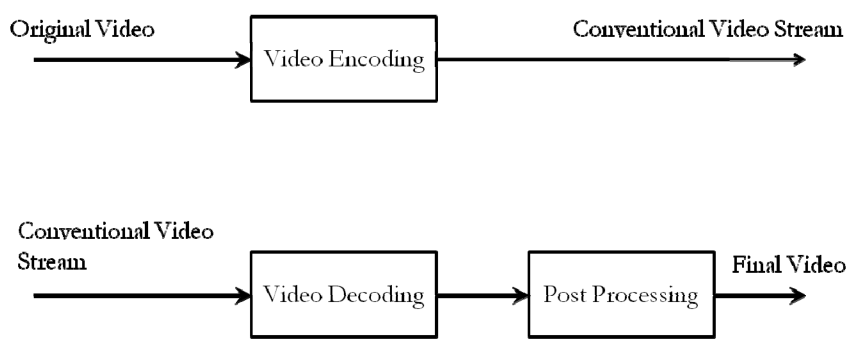

Figure 3: Traditional transmitter and receiver block diagram.

3.2 Receiver Compatible Deinterlacing Using Enhancement Data

There are several ways to employ the receiver compatible approach to improve deinterlacing quality without significantly increasing the computations and cost of receivers. One approach is to send residual video through the enhancement stream. The residual video is the difference between the interlaced, encoded video and the original video. As described in Wan[13], sending residual video is only beneficial if there is sufficient bandwidth in the system.

A more bandwidth-efficient approach is to send intra-field/inter-field switching information, which specifies whether to use intra-field or inter-field methods for different areas of the video. The intra-field/inter-field switching information is calculated and determined at the transmitter. This approach shifts the bulk of the computation from the receiver to the transmitter and is more efficient overall since there are far more receivers than transmitters. Although the final video quality from sending inter-field/intra-field switching information cannot surpass that of sending residual video, it requires a much smaller bit-rate for the enhancement stream.

At the transmitter, a mean square error (MSE)/peak-signal-to-noise ratio (PSNR) comparator can be used for every block. These MSE/PSNR comparator systems have been previously investigated by Sunshine[12] and Wan[14]. In order to optimize the MSE/PSNR of the entire frame, the algorithm optimizes each block individually. The MSE/PSNR for inter-field deinterlacing and intra-field deinterlacing is calculated for each stream. The MSE/PSNR comparator chooses the block (inter-field or intra-field) with the greatest PSNR, then encodes the result in the enhancement stream. See Figures 4 and 5 for a block diagram of the MSE/PSNR comparator.

Although MSE/PSNR is a well-accepted metric for video quality, many studies conclude that MSE/PSNR is occasionally inconsistent with subjective testing results.[15] In the case of deinterlacing, an inter-field block with miceteeth artifacts does not always have a greater MSE than an intra-field block without these artifacts. This issue led us to add a miceteeth detection and correction algorithm to our receiver compatible system.

Our mice-teeth detection algorithm is implemented as follows. The algorithm processes both the original video and inter-field stream with a horizontal-edge detecting filter. There are many possible choices for these filters, but we have found the filter in Figure 6 to work quite well. The processed original video is then subtracted from the processed inter-field stream pixel-by pixel and the square of the difference is computed. If this result is greater than a set threshold, then the pixel of interest qualifies as containing miceteeth. Additional morphological processing can help remove noise in this miceteeth detection process. If an inter-field block contains significant amount of mice-teeth, then that inter-field block is reassigned as an intra-field block. See Figure 7 for a block diagram of miceteeth detection and correction. Since most of the processing is performed at the transmitter, the receiver remains simple. A receiver compatible receiver runs a simple intra-field deinterlacing algorithm and a simple inter-field deinterlacing algorithm in parallel. The receiver can adaptively switch between these two streams in a block-wise fashion with help from the enhancement stream. See Figure 8 for a block diagram of the transmitter and Figure 9 for a block diagram of the receiver.

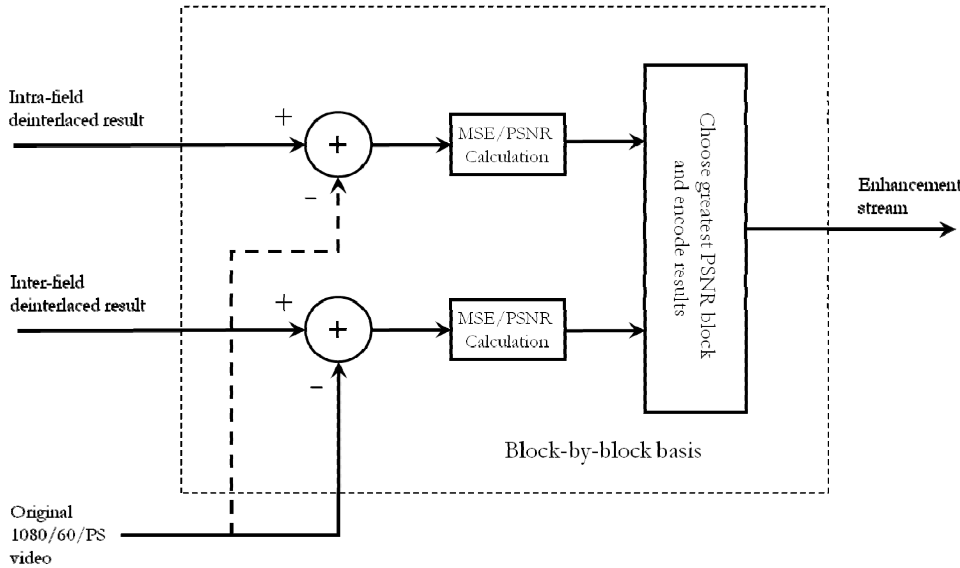

Figure 4: MSE/PSNR comparator block diagram.

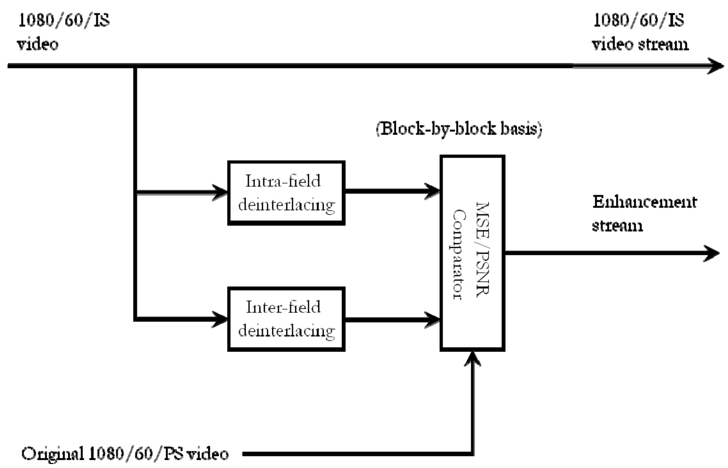

Figure 5: Block diagram for MSE/PSNR comparator receiver compatible transmitter.

1	1	1
-1	-1	-1

Figure 6: Horizontal-edge detecting window.

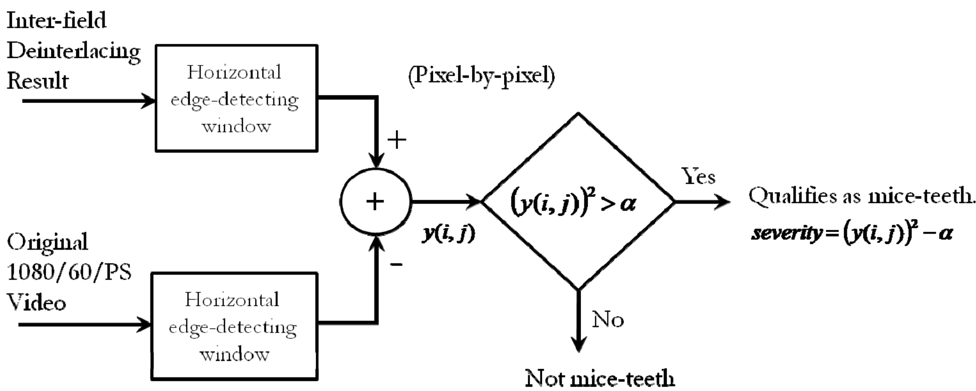

Figure 7: Mice-teeth detection block diagram.

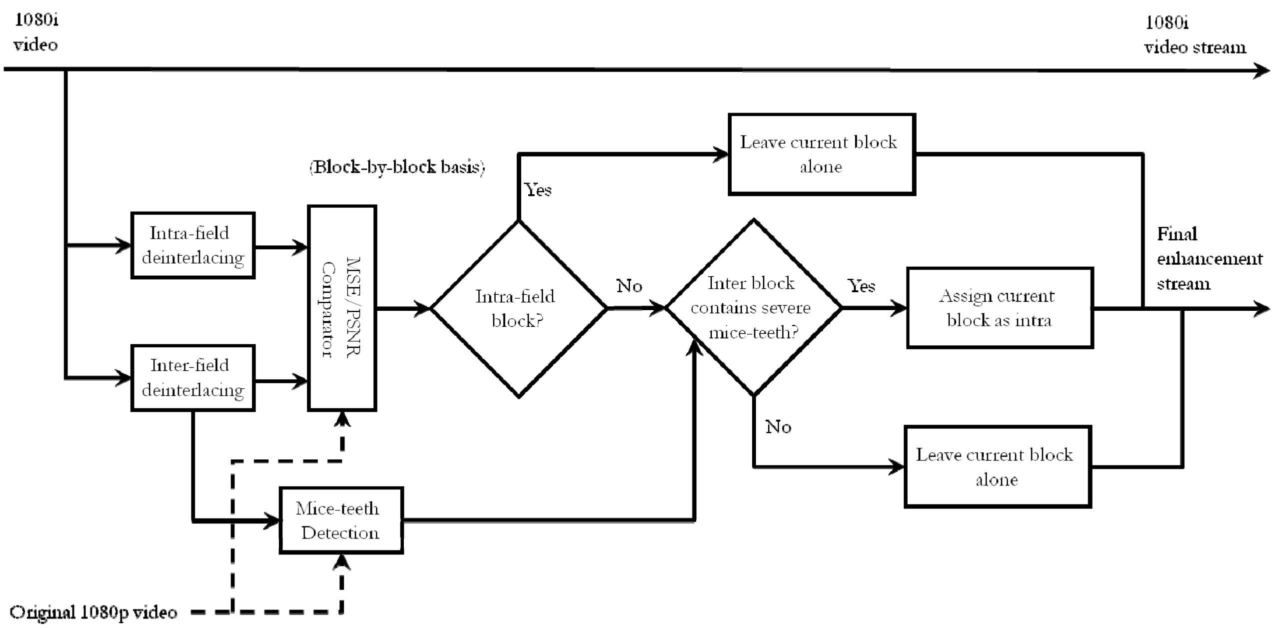

Figure 8: Transmitter block diagram.

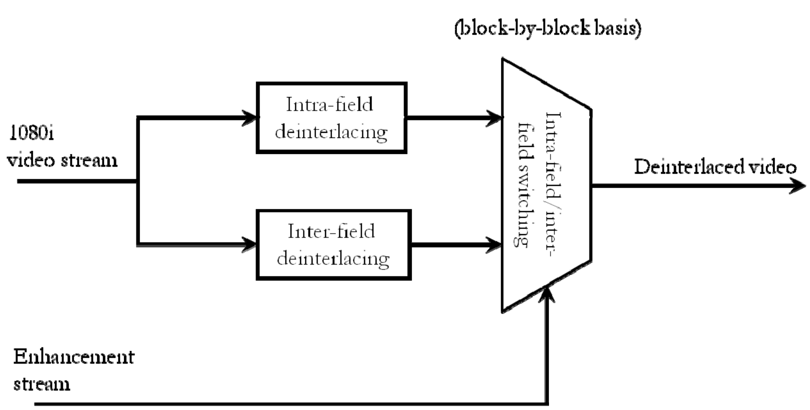

Figure 9: Receiver system block diagram.

3.3 Experimental Setup

We simulated a receiver compatible system which uses Martinez-Lim deinterlacing for the intra-field method, and forward field repetition (FFR) for the inter-field method. Different algorithms can be chosen for the intra-field and inter-field methods, but the concepts are the same. To minimize the bit-rate of the enhancement stream while still providing enough enhancement data, we used 32 x 32 blocks, which yield a .12 Mbit/sec bit-rate for 1080/60/IS video. Note that the .12 Mbit/sec bit-rate is less than 1 percent of the total 19.4 Mbit/sec bandwidth of a television channel. This bit-rate could be further reduced if entropy coding is considered. See Lin[8] for bit-rate calculations and more details. We compared this new system with the Martinez-Lim algorithm and a simple 4-field motion-detection algorithm, examples of traditional deinterlacing systems found in HDTV receivers. The 4-field motion-detection algorithm is similar to the one described by Heng[4].

This 4-field motion detection algorithm is based on a 3-field motion detection algorithm, which is described as follows. Motion at a missing pixel location is estimated by subtracting the 2 temporally neighboring, known pixels. Within a single block, if the sum of the absolute values of these subtractions exceeds a set threshold β, then the block is determined to contain motion. Block-wise processing is used to provide an algorithm which is closer in computations when compared to the receiver compatible receiver. Let I represent the received 1080/60/IS (1080i) video and let p represent the absolute values of the subtractions at each pixel at time t. Let m represent an image that describes the inter-field/intra-field switching on a block-by-block basis for a frame at time t. The subscript for m signifies the number of fields involved in the motion detection scheme. The blocks are identified by b_1, and b_2. N_1 and N_2 represent the set of all pixel values in the block specified by b_1 and b_2. This motion detection procedure is summarized by the following equations:

$$p(n_1, n_2, t) = |I(n_1, n_2, t+1) - I(n_1, n_2, t-1)| \tag{1}$$

$$m_3(b_1, b_2, t) = \begin{cases} 1, & \sum_{n_2 \in N_2} \sum_{n_1 \in N_1} p(n_1, n_2, t) > \beta \\ 0, & otherwise \end{cases} \tag{2}$$

The 4-field motion detection algorithm is based on the 3-field motion detection algorithm. For each block, if the 3-field algorithm detects motion for either the current frame or the previous frame, then the block is determined to contain motion, as shown in the following equation:

$$m_4(b_1, b_2, t) = m_3(b_1, b_2, t) \vee m_3(b_1, b_2, t-1) \tag{3}$$

With respect to computational requirements of a receiver, the receiver compatible receiver is comparable to the Martinez-Lim deinterlacing algorithm and is simpler than the 4-field motion detection algorithm. While the motion detection algorithm must perform computations including subtractions and many comparisons to perform intra-field/inter-field switching, the receiver compatible receiver only decodes the transmitter-calculated enhancement stream which contains intra-field/inter-field switching information.

4. RESULTS

The receiver compatible deinterlacing approach outperforms both the Martinez-Lim and the motion detection algorithms in terms of PSNR and visual quality. See Figure 10 for summarized results and Figure 11 for an example frame. Note that the video sequences *Speed Bag*, *Pedestrians*, and *Rush Field* all feature a relatively stationary camera, while the other sequences were recorded with either a panning or zooming camera.

For the stationary sequences, the receiver compatible algorithm consistently outperforms both traditional algorithms by a wide margin. The receiver compatible algorithm outperforms the Martinez-Lim algorithm on an average of about 2 decibels (dB) in PSNR, a significant difference. The receiver compatible algorithm is also visibly superior and shows much more detail in stationary areas of the scene. The stationary areas from the Martinez-Lim algorithm display severe twitter for near-horizontal lines. On average, the receiver compatible algorithm outperforms the motion detection algorithm by about 1 dB. The motion detection algorithm does not introduce twitter artifacts or a loss of vertical detail in stationary areas. However, the motion detection algorithm introduces visible mice-teeth, especially near boundaries of moving objects. The receiver compatible algorithm does not exhibit this deficiency.

For the sequences that feature a panning or a zooming camera, the receiver compatible algorithm shows less improvement over the traditional deinterlacing algorithms. This result is expected because the receiver compatible algorithm can only switch between an intra-field and an inter-field algorithm for different parts of the frame. Since moving images feature little non-translational correlation between frames, both the receiver compatible and the motion detection algorithms assign most of the frame as intra-field blocks.

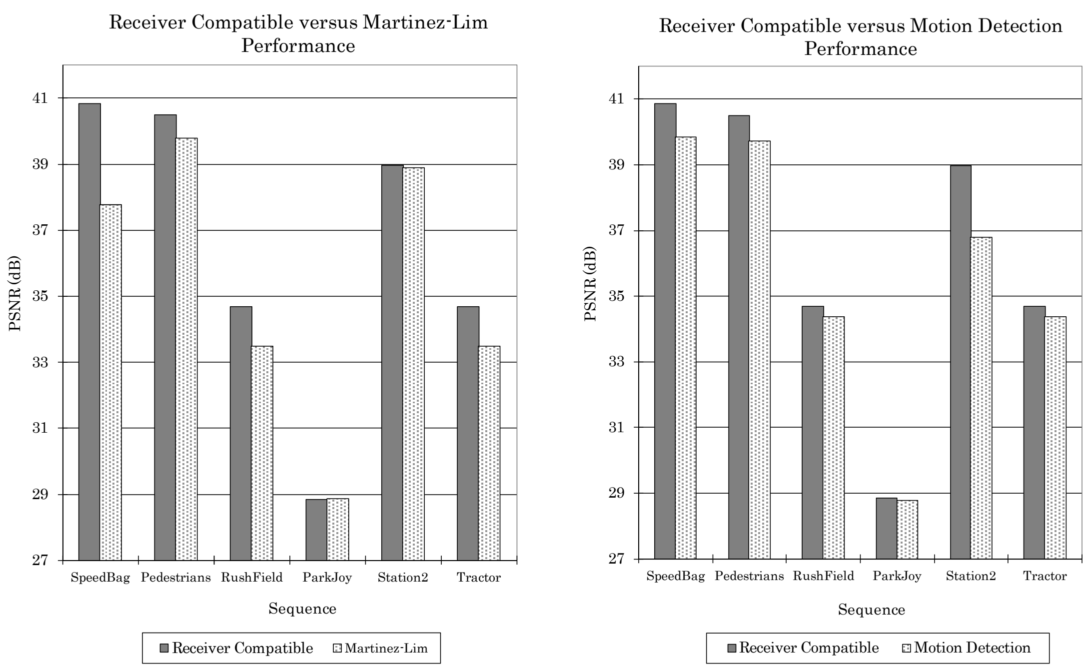

Figure 10: The receiver compatible system significantly outperforms traditional deinterlacing algorithms.

(a) Original frame.

(b) Martinez-Lim deinterlacing results. Note the loss of details in the text.

Figure 11: Comparison of traditional deinterlacing techniques against the receiver compatible system.

(c) Motion detection results. The text is now sharp. However, there are a few mice-teeth artifacts present (right side of the boxer's shirt, and under his chin).

(d) Receiver compatible results. The vertical details of stationary areas are retained and there are no mice-teeth artifacts.

Figure 11 (continued)

5. CONCLUSION

Our proposed receiver compatible system sends a small bit-rate enhancement stream in parallel with the existing IS stream to allow for simpler deinterlacing of 1080/60/IS (1080i) HDTV at the receiver. Existing receivers ignore this new stream while new receivers use the enhancement stream to improve HDTV quality. The receiver compatible system shifts computation from the receivers to the transmitter, resulting in improved video quality with no increase in computations for receivers. Our proposed algorithm uses a MSE/PSNR comparator with mice-teeth correction. This new receiver compatible system outperforms traditional deinterlacing algorithms in terms of MSE/PSNR as well as visual quality.

REFERENCES

[1] ATSC digital television standard, ATSC: A/53D, Advanced Television Systems Committee, Washington, D.C. (2005).

[2] Challapali, K., Lebegue, X., Lim, J.S., Paik, W.H., and Snopko, P.A., "The Grand Alliance system for US HDTV," Proceedings of the IEEE, 83(2), 158-174 (1995).

[3] De Haan, G.; Bellers, E.B., "Deinterlacing-an overview," Proceedings of the IEEE, 86(9), 1839-1857 (1998).

[4] Heng, B., "Application of Deinterlacing for the Improvement of Surveillance Video," M.S. thesis, Massachusetts Institute of Technology, Cambridge, MA US (2001).

[5] Kwon, O.; Kwanghoon Sohn; Chulhee Lee, "Deinterlacing using directional interpolation and motion compensation," IEEE Transactions on Consumer Electronics, 49(1), 198-203 (2003).

[6] Lim, J.S., "A migration path to a better digital television system", SMPTE Journal, 103(1), 2-6 (1994).

[7] Lim, J. and Sunshine, L., "HDTV transmission formats and migration path," International Journal of Imaging Systems and Technology, 5(4), 286-291 (1994).

[8] Lin, A.L. "Using Enhancement Data for Deinterlacing 1080i High Definition Television," M. Eng. thesis, Massachusetts Institute of Technology, Cambridge, MA US (2009).

[9] Lin, C., Sheu, M., Chiang, H., Liaw and C., Lin, J., "Motion adaptive de-interlacing with local scene changes detection," Proc. of the Second International Conference on Innovative Computing, Information and Control, pp. 142 (2007).

[10] Martinez, D.M. and Lim, J.S., "Spatial interpolation of interlaced television pictures," Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 2045–2048 (1989).

[11] Patti, A.J., Sezan, M.I. and Tekalp, A.M., "Robust methods for high-quality stills from interlaced video in the presence of dominant motion," IEEE Transactions on Circuits and Systems for Video Technology, 7(2), 328-342 (1997).

[12] Sunshine, L., "HDTV transmission format conversion and migration path," Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA US (1997).

[13] Wan, W., "Adaptive format conversion information as enhancement data for scalable video coding," Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA US (2002).

[14] Wan, W. and Lim, J.S. "Adaptive format conversion for scalable video coding," in Proc. of SPIE 4472, 390-401 (2001).

[15] Winkler, S. and Mohandas, P., "The evolution of video quality measurement: From psnr to hybrid metrics," IEEE Transactions on Broadcasting, 54 (3), 660-668 (2008).