Proceedings, 4th International Workshop on HDTV and Beyond, Torino, Italy, September 1991.

Compatible Coding of Digital Interlaced HDTV Using Prediction of the Even Fields from the Odd Fields

Ti-Hao Chiang, Robert Mokry, Feng-Ming Wang, Alexandros Eleftheriadis,
Yan-Bin Yu and Dimitris Anastassiou

Image and Advanced Television Laboratory
Department of Electrical Engineering
Columbia University
New York, NY 10027, USA

Abstract

A high-compression technique is presented for coding digital interlaced HDTV. A key feature of the method consists of coding the even fields using information from the neighboring odd fields, which have been coded first with an MPEG-type technique. This approach, first proposed in [1,2], was originally designed to aid development of MPEG-2 coding techniques of CCIR-601 video, compatible with the MPEG-1 video coding standard [3]. The same approach can also facilitate the design of HDTV/Standard TV compatible coding.

In this paper, we outline two coding schemes that are based on this idea of field separation. One of them aims at optimizing the coding performance for HDTV and is noncompatible, allowing occasional joint coding of the two fields. The other is compatible with Standard TV and also considers the problem of horizontal resolution mismatch; it aims at optimizing coding of HDTV, while reducing error propagation in standard TV coding by use of predictive leak in the DCT domain.

These methods are appropriate for digital interlaced HDTV coding, including possible terrestrial and satellite transmission applications.

The even fields are coded with a digitally assisted multi-mode hybrid interpolative/recursive technique, using both predictive (hence recursive) and interpolative modes of operation. Of particular importance is the use of a combined hybrid recursive/-interpolative mode.

The technique provides for convenient multiresolution channel coding for graceful degradation in the presence of channel errors.

1. Introduction

Digital HDTV coding is closely related to the second phase of the Standard TV coding standardization effort of the International Standardization Organization, known as MPEG-2. For example, coding of CCIR-601 video at 4 Mbits/s could be used for

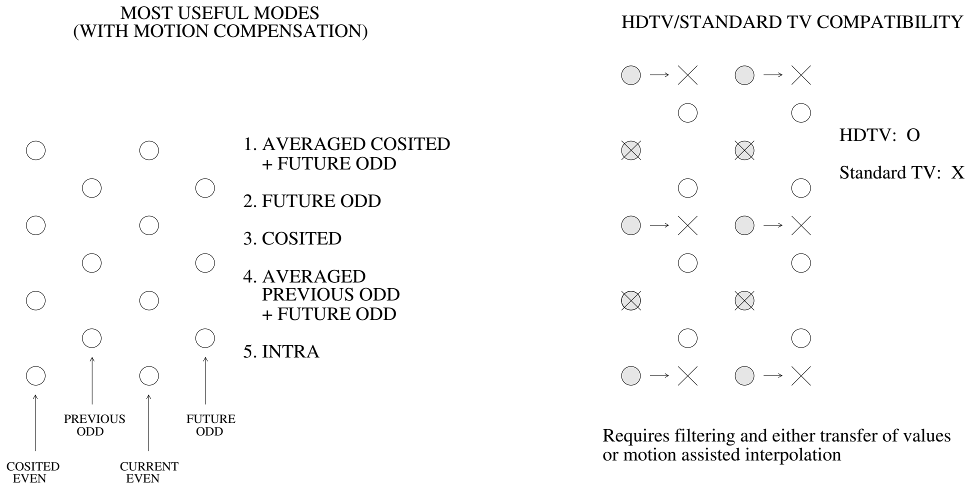

Figure 1: Sampling lattices and motion compensation modes

HDTV coding at about 20 Mbits/s, with the same SNR. This is the rate appropriate for digital terrestrial broadcasting of Advanced Television. On the other hand, coding of CCIR-601 video at 9 Mbits/s corresponds to a rate of about 45 Mbits/s for HDTV, and that could be appropriate for, e.g., DBS digital HDTV broadcasting. Therefore, it is relevant to examine MPEG-type approaches for HDTV coding. The first phase of MPEG, now nearly completed, is a hybrid predictive/DCT approach of a pyramidal nature in the temporal domain, containing various hierarchical levels (I,P and B) of the video signal. For interlaced video, MPEG-1 deals only with the odd fields. It is, therefore, natural to consider adding another level (the odd fields), in the MPEG hierarchy for MPEG-2.

Downsampling the video sequence by skipping the even fields results in a progressive sequence of only the odd fields, which can be coded first (see Fig. 1 for the pixel locations in the vertical-temporal plane). The simulation models of MPEG-1 [4] suggest coding of only the odd fields for high compression. At higher bit rates, there is room for spending bits to code the even fields as well, in a pyramidal fashion [5]. In this paper, we are using concepts similar to those used in [6] to extend coding to both fields.

2. Compatibility/robustness vs. coding performance

Multiresolution video coding (typically via subband or pyramid decomposition) provides a coded bit stream for a full-resolution video signal, so that subsets of that bit stream can be used for reconstructing the same signal at lower resolutions. If these lower resolutions correspond to standard video formats, then the coding scheme is compatible with those formats. Furthermore, this technique will also be robust, i.e. resilient to channel errors, provided that transmission is done in a prioritized multi-channel manner, so that high priority bits correspond to lower resolutions, while low priority bits are used to reconstruct higher resolutions. In this way, in the presence of channel errors, low priority channels may be lost, but the quality of the reconstructed signal degrades gracefully by shifting to lower resolutions, instead of having total communication disruption.

There is a price in coding performance to be paid for multiresolution coding, because of the extra imposed constraints. For example, consider a two- level compatible bit allocation for a high-resolution signal and a low- resolution signal. The required price in coding performance can be allocated in each of the two channels, depending on the application. For example, if good performance at low-resolution is needed, one may optimize coding for it, and then encode the remaining enhancement signal, thus degrading performance at high resolution. If good performance at high-resolution is needed, then coding is optimized as much as possible at that level, trying to allow some imperfect reconstruction of the low-resolution part from some subset of the bit stream. In some cases, it may be appropriate to aid reconstruction of the low- resolution signal by sending an extra low-rate bit stream.

3. Interpolation of the even fields from the odd fields

Decomposition between odd and even fields can be the first stage of multiresolution coding. A missing even field can be interpolated by the past and the future neighboring odd fields. We found that digitally assisted interpolation is required for good quality rendition of the even fields, and for high compression coding of the difference (enhancement) signal. The main novel feature of the proposed scheme is the joint use of prediction from either the odd or the even fields, which significantly improves performance when compared with coding based on only the even (recursive coding), or only the odd (interpolative coding) fields. The following modes of operation (requiring transmission of extra marker bits including up to two motion vectors) are used: an averaged (combined) mode between the previous co-sited even field and the future odd field; an averaged (combined) mode between the past and future odd fields; a single mode using the co-sited field; and a single mode using the future odd field (see Fig. 1).

Using a future picture for prediction is very important, because it improves prediction whenever there is uncovered background. The savings for considering uncovered background far outweigh the cost for having to skip some pictures for doing that [7].

Other modes, including intraframe, are also possible. Any error propagation is stopped whenever the encoder selects use of purely interpolative modes, without otherwise notifying the decoder, resulting in a finite-memory system.

4. Multiresolution joint source/channel coding

In the presence of channel errors, it is desirable to have graceful degradation, rather then total communication disruption whenever the bits start being misinterpreted. As mentioned above, multiresolution layered coding is suitable for this purpose, in addition to error correction, concealment and reinitialization. A particularly interesting case is terrestrial digital HDTV broadcasting, using multiresolution channel coding. Coding the even fields from the odd fields offers a convenient approach in this respect. Higher protection should be assigned for the coded bits of the odd fields than those of the even fields. When the even fields cannot be coded due to channel errors, they can be interpolated, resulting in only slight degradation.

Since the odd fields are coded with an MPEG-like algorithm, it is possible to determine a number of resolutions, resulting in a hierarchical pyramid scheme, in which a number of parallel channels is used, each with a different error protection. There are many such possible hierarchical levels, including the "intra" odd fields, the "predicted" odd fields, the "interpolated" odd fields, the horizontal resolution enhancement, and, finally, the even fields. It is a design problem depending on the particular transmission circumstances to select the optimum hierarchical levels.

In the United States, the problem of channel coding for digital terrestrial broadcasting within a bandwidth of 6 MHz is of particular importance, due to the requirements of the Federal Communications Commission. In that case, spectrum efficiency is essential: if a digital system is designed to have reliable reception in the fringe (remote) areas, then the same system will be necessarily overdesigned for the areas of close proximity to the transmitter (i.e. the inner city, where other services like mobile telephony are competing for bandwidth as well). As a result, resources are wasted where they are needed most, as was explained in [8]. One solution to that problem is the above mentioned technique of multiresolution combined source/channel coding with graceful degradation. This, coupled with multiresolution transmission (see, e.g. [9]), allows scalable prioritized reception, in which, in addition to usable reception in remote areas, also provides for manufacturing of inexpensive TV receivers capable of only processing the high priority bits.

5. Compatible (HDTV/Standard TV) approach

The sampling lattice of interlaced Standard TV is not a subset of the corresponding lattice of interlaced HDTV, as shown in Fig. 1, even if HDTV has exactly twice the number of lines of Standard TV. This is a problem that can be solved in various ways. Furthermore, in order to have good compression of Standard TV, the problem of horizontal resolution mismatch must be addressed as well. Both these issues are discussed below. Finally, the problem of different aspect ratios is solved by independent coding of the side panels at HDTV resolution.

An advantage of first independently coding the odd fields of interlaced HDTV is that they can be used for resampling to interlaced Standard TV. Let us assume that the original signal has an interlaced HDTV resolution of K/L/2, in which K can be 1125, 1250, 1050, etc., and L can be 60, 50, 59.94, etc. Dropping the even fields will yield a (K/2)/(L/2)/1 signal. Converting that signal, with high quality, into, say, a useful (K/2)/L/2 is not straightforward, although the total number of samples is identical. A simple transfer of values (see Fig. 1) will create motion jerkiness. For uniformly moving objects, temporal and spatial resolution can be exchanged: slight spatial filtering will greatly reduce this problem without significant blurring. This approach may be acceptable for some applications, including multiresolution channel coding. More accurate results are obtained if assisted by side motion vector information.

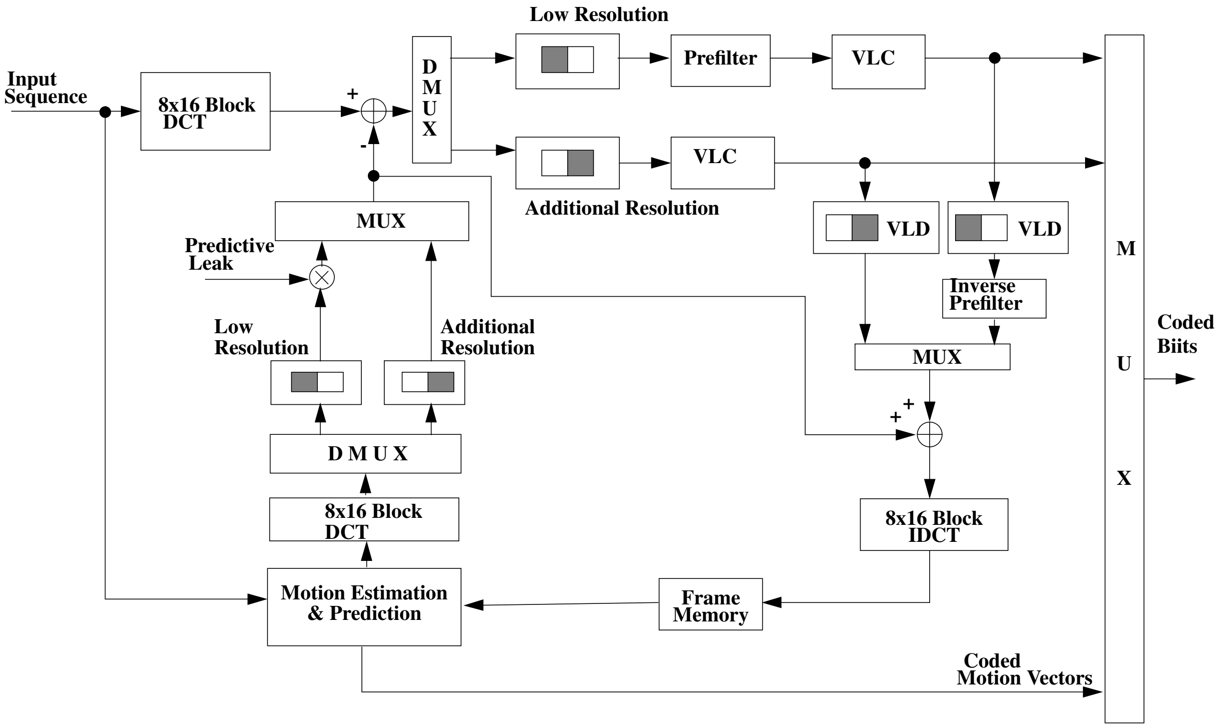

Figure 2: Two-level compatible encoder

Although such motion vector information is directly transmitted in the proposed technique for prediction of the even fields, care must be taken to avoid error propagation when using recursive modes. Therefore, it is best to independently send to the Standard TV decoder a choice among the forward, backward, and combined modes, incorporating the odd fields only. This will be an extra bit stream of low rate.

For higher compression of the low resolution signal, it is appropriate to perform horizontal downsampling before coding the (K/2)/L/1 signal, in which case extra bits must be sent for the enhancement in the horizontal dimension, when coding the HDTV signal. This can be done in various ways, including subband, pyramid, and DCT-splitting based techniques, and will result in some performance deterioration.

We follow the approach proposed in [10] for MPEG-2/MPEG-1 compatible coding, in which the odd fields of the high-resolution signal are compressed using 16x8 DCT coding, and an 8x8 DCT matrix is extracted for reconstruction of the low- resolution signal. Zig-zag scanning of DCT coefficients for each 16x8 block must first cover the low-frequency 8x8 block, followed by the high-frequency 8x8 block. In our case, we have more freedom, as the low-resolution coding loop does not need to be exactly identical to that of MPEG-1. We attempt to optimize performance of the high-resolution (HDTV) loop, and perform motion estimation in that loop. However, if the low-resolution loop uses the motion vectors of the high-resolution loop (properly divided by 2), and the 8x8 part of the 16x8 DCT for the prediction error, then errors will begin accumulating in the low resolution loop, and the quality of Standard TV will soon become unacceptable.

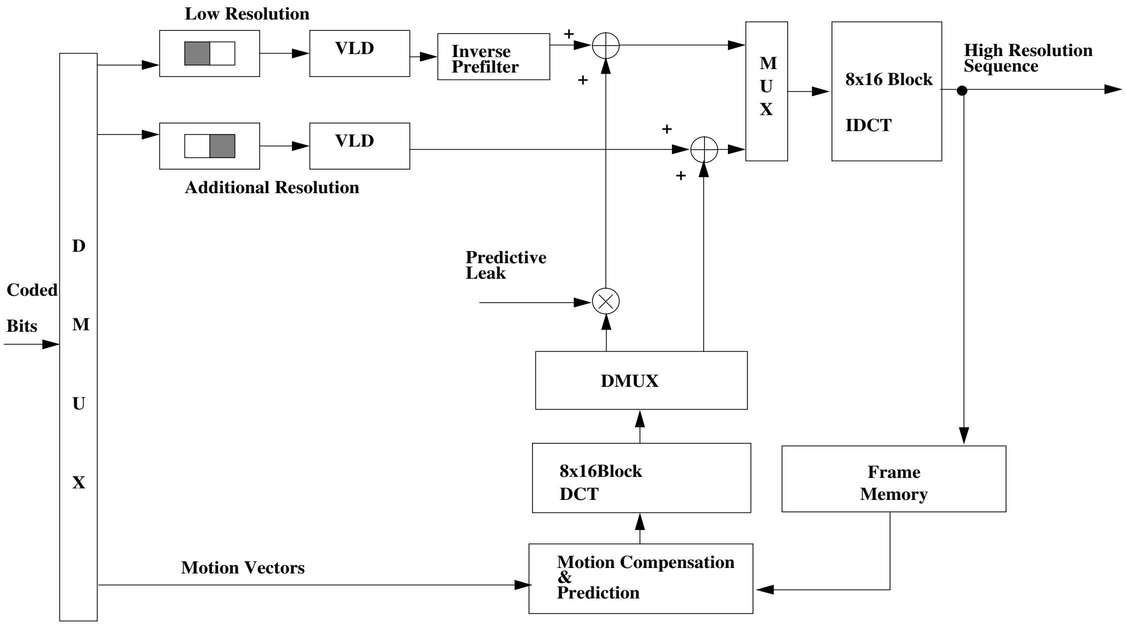

Figure 3: High resolution decoder

In order to reduce the propagation of such errors, we perform predictive leak in the low-resolution loop. In the spatial domain, this can be done by multiplying the predicted low resolution signal by a coefficient less than but close to one. We do this adaptively in the DCT domain, and selectively for only the low horizontal frequency coefficients, except for the DC coefficient, which is always sent error-free. In this way, the high frequency coefficients are not affected. Each coefficient is multiplied by a different scaling factor in order to optimize performance. This technique reduces the cost to be paid in the performance in the high resolution loop. Our results have shown that overall coding performance using DCT leak is slightly better than that of space-domain leak. Prefiltering in the DCT domain is also useful to keep balanced proportionality for each DCT coefficient. Then, the decoder can perform the inverse of the prefiltering process.

A block diagram of the encoder is shown in Fig. 2. Block diagrams of the decoders for high and low resolution are also shown in Figures 3 and 4.

6. Noncompatible approach

If high coding performance is essential, joint coding of even and odd fields can be allowed, but scalability, multiresolution channel coding, and possible compatibility with Standard TV are sacrificed, unless the algorithm is switched to a mode forcing separate field coding.

As explained above, pyramid coding has many advantages, but the performance at the highest resolution is reduced, as was shown, e.g., in the JPEG standard. For example, given an interlaced freeze-frame of a moving object, it makes sense to compress the first

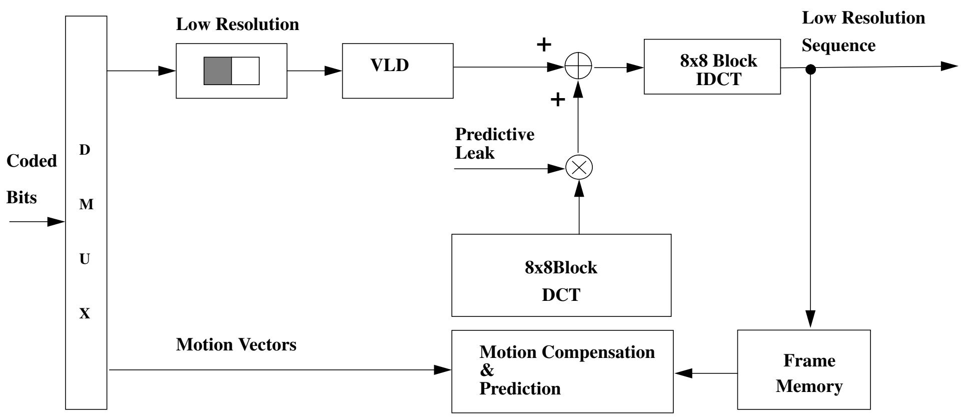

Figure 4: Low resolution decoder

field, and then the second field based on the reconstructed values of the first field (see, e.g., [11]). However, if the frame contains no moving objects, the coding performance of the final result will be compromised by this two-stage approach. It is then preferable to code the fields jointly, because they are properly aligned. Similarly, in motion compensated predictive video coding, if an object has uniform motion, the error image contains properly aligned fields, and it may be preferable to compress them jointly. This is not the case, however, when the object is accelerating, or at the critical areas around the edges of moving objects. Accordingly, the decision on whether to code the fields separately or jointly must be taken on a macroblock-by-macroblock basis. This approach is also used in [12, p. 19].

If background is uncovered, accurate prediction can only be made using future fields. Uncovered background may not occur often, and the SNR will not be affected by local inaccurate coding, but artifacts can be quite visible. Therefore, if fields are coded separately, the future (odd) field is coded first, so that the previous even field can then be coded using the results of the future odd field, with backwards motion estimation, as described in the previous section.

This implies that the whole odd field must be coded at a first stage. The macroblocks for which the even field was not coded jointly with the odd field are tagged. Coding of the even field is completed at a second stage, using the adjacent future odd field, in addition to past even and odd fields.

Conclusions

Compatibility, robustness for multiresolution channel coding, and scalability can be achieved by coding the odd fields, followed by even field coding, using the already coded future odd field to accommodate uncovered background.

We have found, however, slightly better coding performance of the final signal, by occasionally (but not always) allowing the fields to be coded jointly. The optimum approach will highly depend on the weights of relative importance assigned on the various issues. Subjective evaluation experiments were done at Columbia University using a RAM-based HDTV simulation system connected to an HDTV monitor.

References

1 F.M. Wang and D. Anastassiou, "Compatible even field coding from odd fields using the co-sited field in single and averaged modes", ISO-IEC JTC1/SC2/WG11 MPEG 90/308, December 1990.

2 F.M. Wang and D. Anastassiou, "High-quality coding of the even fields based on the odd fields of interlaced video sequences", IEEE Transactions on Circuits and Systems, Vol. 38, No. 1, January 1991, pp. 140-142.

3 MPEG Committee Draft, Part 2 Video, (ISO: Coding of Moving Pictures and Associated Audio for Digital Storage Media up to about 1.5 Mbits/s), August 22, 1991

4 MPEG Video Simulation Model Three, International Organization for Standardization, Coded Representation of Picture and Audio Information, 1990.

5 P.J. Burt, E.H. Adelson, "The Laplacian Pyramid as a Compact Image Code", IEEE Transactions on Communications, Vol. COM-31, No. 4, April 1983.

6 D. Anastassiou, "Generalized Pyramidal Coding using Nonlinear Interpolation for HDTV Coding", Proceedings, 1990 Picture Coding Symposium.

7 D. Anastassiou, T.H. Chiang, A. Eleftheriadis, R. Mokry, F.M. Wang and Y.B. Yu, "Even field coding using adjacent future odd field, in addition to past even and odd fields", ISO-IEC JTC1/SC2/WG11 MPEG 91/014, May 1991.

8 W. Schreiber, "A Critique of Purely Digital HDTV" IEEE Spectrum, April 1991, p. 72.

9 K.M. Uz, K. Ramchandran and M. Vetterli, "Multiresolution Source and Channel Coding for Digital Broadcast of HDTV", this issue.

10 "Improvements of the Multi-Purpose Coding Scheme", PTT, The Netherlands, and "MPEG1 Compatible Scheme based on DCT Splitting", LEP, France, ISO-IEC JTC1/SC2/WG11 MPEG91, May 1991.

11 D. Anastassiou, W. B. Pennebaker and J. L. Mitchell, "Gray-Scale Image Coding for Freeze-Frame Videoconferencing", IEEE Trans. Commun., vol. COM-34, pp. 382-394, Apr. 1986.

12 "DIGICIPHER HDTV System Description", submitted to the FCC by the General Instrument Corp., (Videocipher Division) on behalf of the American Television Alliance, July 26, 1991.