Presented at the SMPTE 141st Technical Conference in NY, November 19-22, 1999

Motion Compensated De-Interlacing:
The Key to the Digital Video Transition

by
Jed Deame
Director of Product Development
Teranex, Inc.
http://www.Teranex.com

Abstract

As the video industry transitions from analog to digital, more and more video processing equipment will also need to transition from analog to digital. The current analog television standards, NTSC, PAL, etc., are based on interlaced formats. As these standards transition to digital, the demand for progressive material will increase, causing a directly proportional increase in the demand for video processing products with high quality de-interlacing.

There are many ways to perform the de-interlace process, with varying levels of quality and corresponding compute requirements. This paper will emphasize the importance of proper de-interlacing, examine some of the currently used techniques for de-interlacing video, delineate a new de-interlacing technique, and discuss applications where this new de-interlacing technique can improve the end-to-end image quality of any DTV system.

Introduction

In the traditional analog television facility, typically there are many different videotape formats in use. Common formats range from archaic 2” Quadruplex to the Workhorse 1” Type C and 3/4 U-matic, to the new compressed formats such as DVCAM and DVC Pro. Although all of these analog tape formats are inherently incompatible, there is a simple method of transcoding from one format to another: a 75-ohm coax cable. Even though each of the above tape formats resolve varying number of pixels per line, the common interchange format, NTSC analog, is supplied at the output of each VTR, and accepted at the input of each VTR. This luxury is no longer available in the realm of Digital Television.

The new DTV formats, including HD-D5, HDCAM and D9-HD each have a common SMPTE 292M digital I/O port, but there is very limited plug compatibility between decks. This is due to the differing number of pixels/line, lines/frame, and progressive vs. interlace scanning. To dub from one DTV format to another requires a piece of equipment called the DTV Format Converter. The DTV Format Converter is the “black box” that converts from any DTV format to any other DTV format, to enable the transfer of data between the above DTV VTRs. The most challenging part of DTV Format Conversion is de-interlacing. De-interlacing is required any time an interlaced source is format converted, regardless of whether the destination is progressive or interlaced. De-interlacing is required in interlace to interlace conversions in order to provide more temporally aligned samples for the resampling process.

Survey of De-Interlacing Techniques

Interlaced video signals consist of two video fields, each containing the odd or even lines of the image. During the image capture process, the camera outputs the odd lines at one instant in time, and then 33 milliseconds later, outputs the even lines. This creates a temporal shift between the odd and even lines of the image, which must be addressed in frame based processing systems. This problem is illustrated in Figure 1(a) and 1(b). The de-interlace process attempts to overcome this problem by creating a clean frame from the two fields. There are two basic classes of de-interlacing algorithms: non-motion compensated and motion compensated. Both classes are described in the following sections.

Figure 1(a): Odd Field of Interlaced Image

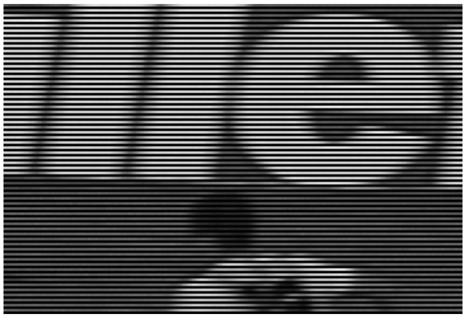

Figure 1(b): Even Field of Interlaced Image

Non-Motion Compensated De-Interlacing

There are two categories of non-motion compensated de-interlacing algorithms – linear and non-linear. Both categories contain spatial (or intra-field), temporal (or inter-field), and spatio-temporal algorithms. For brevity, only the most popular methods are examined.

Linear Techniques

The two most basic linear conversion techniques are called “Bob” and “Weave”. “Weave” is the simpler of the two methods. It is a linear filter that implements pure temporal interpolation. In other words, the two input fields are overlaid or “woven” together to generate a progressive frame; essentially a temporal all-pass. See Figure 2(a). While this technique results in no degradation of static images, moving edges exhibit significant serrations, which is an unacceptable artifact in a broadcast or professional television environment.

“Bob”, or spatial field interpolation, is the most basic linear filter used in the television industry for de-interlacing. In this method, every other line (one field) of the input image is discarded, reducing the image size from 720x486 to 720x243. The half resolution image is then interpolated back to 720x486 by averaging adjacent lines to fill in the voids. The advantage of this process is that it exhibits no motion artifacts and has minimal compute requirements. The disadvantage is that the input vertical resolution is halved before the

Figure 2(a): Fields “Weaved” together with a Linear Temporal Filter

Figure 2(b): Fields “Bob’d” together with a Linear Spatial Filter

image is interpolated, thus reducing the detail in the progressive image. The results of this process are shown in Figure 2(b).

A combination of the linear spatial and linear temporal methods is the linear vertical-temporal (VT) filter. The VT de-interlacer is the best performing of the linear filters. This method gradually reduces the vertical detail as the temporal frequencies increase. The contribution from the neighboring field is limited to the high vertical frequencies, such that motion artifacts can be minimized. Note that the vertical detail from the previous field is being combined with the temporally shifted current field, so some motion blur may occur. The results of this process are shown in Figure 3.

Non-Linear Techniques

Linear interpolators work quite well in the absence of motion, but television consists of moving images, so more sophisticated methods are required. The field-weave method works well for scenes with no motion, and the field interpolation method is a reasonable choice if there is high motion. Non-linear techniques, such as motion adaptive de-interlacing, attempt to switch between methods optimized for low and high motion. In motion adaptive de-interlacing, the amount of inter-field motion is measured and used to decide whether to use the entire input frame (if no inter-field motion detected), or discard one field (if significant motion detected). Advanced motion adaptive techniques can vary the percentage of the previous field’s data that is retained as the inter-field motion increases. The challenge in implementing an effective motion adaptive algorithm is managing the trade-off between double images, due to the inclusion of too much uncorrelated previous field information, and lost vertical resolution, when not enough previous field information is used. Another challenge in varying the temporal aperture is avoiding “resolution pumping” as objects start or stop moving. The processing requirements for motion adaptive de-interlacing are higher than that of field interpolation, at the gain of somewhat higher output image quality. The results of this process are shown in Figure 4.

Figure 3: Vertical-Temporal De-interlacing

Figure 4: Motion Adaptive De-interlacing

Motion Compensated De-Interlacing

The most advanced method of de-interlacing is motion compensation. This technique, which at one time was thought to be too complex to implement in hardware, is currently used in advanced SDTV standards converters, and is beginning to appear in high-end DTV format converters. Motion compensated de-interlacing measures the inter-field motion, and then aligns data between the two video fields, maximizing the vertical resolution of the image.

Motion compensation consists of shifting the pixels in the two temporally displaced fields to a common point in time. Determining the amount of shift for each pixel is referred to as motion estimation, in which motion vectors are identified and tracked from one field to another. This is typically implemented as a block-matching process. Typical block sizes range from 4x4 to 8x8. The implication of using blocks is that only these relatively large pixel blocks can be moved in time to the correct spatial position; with the further restriction of requiring every pixel in the block to move the same displacement.

The Teranex PixelComp™ Algorithm

Teranex has developed an enhanced version of motion compensated de-interlacing called PixelComp™. The PixelComp™ motion estimator uses a full exhaustive search on 1x1 block sizes, generating a motion vector for every pixel in the image. This allows motion of any kind to be detected at pixel granularity. The pixel motion vectors are then used to precisely position the objects in the image to the correct spatial position, maintaining the full vertical resolution of the imagery even when significant motion is detected. PixelComp™ also includes advanced techniques for filtering motion vectors to insure their accuracy. The results of the PixelComp™ de-interlace process are shown in Figure 5.

Figure 5: Teranex PixelComp De-interlacing

Hardware Implementations

Linear de-interlacers are implemented readily in commodity Field Programmable Gate Arrays (FPGAs). Motion adaptive de-interlacers require higher density FPGAs, but are still readily implemented in a compact 19” rack mount chassis. The complexity of motion compensated implementations is significantly higher, historically requiring an additional chassis containing many Application Specific Integrated Circuits (ASICs). The extra hardware is required to perform the complex motion estimation and compensation process. Previous implementations of real-time block matching hardware take many shortcuts; such as hierarchical search methods and large block sizes in order to make the design realizable. The PixelComp™ algorithm makes no such compromises. Consequently, this exacting technique is far too compute intensive to implement in FPGAs or ASICs. Supercomputer-level performance is needed for such a task.

The Teranex Video Computer Platform

Teranex has developed and perfected a parallel processing supercomputer technology that is capable of extremely high compute power with minimal component real estate. This SIMD (Single Instruction Multiple Data) parallel processing technology has been acquired by Teranex and packaged in a compact 6RU chassis. The Video Computer platform has a throughput of over one trillion operations per second and is capable of processing standard definition and/or high definition video with exacting quality in real-time.

The Video Computer platform is the basis for all current and future Teranex products. The programmable nature of the processing engine enables the system to be configured by simply installing a disk into the built in CD-ROM drive. This reconfigurability enables additional features or enhanced algorithms to be installed in the product with a simple CD-ROM update.

De-Interlacing Applications

An example of a video processing product requiring high quality de-interlacing is SDTV to HDTV up-conversion from 480i to 720p. Format up-conversion consists of two steps: de-interlacing, and resampling. De-interlacing, the most complex of the two steps, is necessary because the odd and even lines of an interlaced image are temporally shifted, and will cause a misalignment if they are merged unaltered. This misalignment must be corrected to create visually pleasing progressive images. Once a correctly aligned 480p image is created, it may then be resampled to the desired new image format.

De-interlacing is also important in compression systems. Temporal codecs, such as MPEG-2, gain compression efficiency by coding only the differences between frames, in the form of motion vectors and error signals. In an interlaced system, coding inefficiencies due to poor motion vectors may result if there is vertical motion present at a multiple of the field rate. That is, the MPEG-2 encoder will first try to predict the motion between the odd fields of consecutive frames. However, field-rate vertical motion will be uncorrelated between the odd frames since the detail will have moved into the even field. To compensate for field-rate vertical motion, the MPEG-2 adaptive field-frame encoder will combine the odd and even field, without proper de-interlacing, and attempt to predict motion between the frames. Since the frames are typically created by “weaving” the two input fields together, the inter-field motion in the combined frame will generate artificial vertical detail. This in turn makes motion prediction even more difficult, and significantly decreases the coding efficiency of any intra-coded blocks. Therefore, high quality de-interlacing before compression will increase coding efficiency significantly. This increase in coding efficiency results in smaller streams at the same quality level, or more importantly for broadcasters, higher quality levels at the same bitrate.

In addition to format converters and compression systems, high quality de-interlacing will be required in any digital processing system that operates on standard definition interlaced video. Operating on images in the progressive domain internally enables the full vertical resolution of the input to be maintained throughout the processing chain.

Conclusions

There are many techniques for de-interlacing video data, requiring varying levels of computational power, and providing varying levels of image quality. In general, the non-linear techniques outperform the linear techniques, with the possible exception of vertical-temporal de-interlacing. The VT de-interlacer exhibits different motion artifacts from the non-linear motion adaptive de-interlacer, and the relative image quality comparison is a subjective choice. Motion adaptive de-interlacing often suffers from motion detector switching artifacts and vertical detail being interpreted as motion. Vertical-temporal de-interlacing often suffers from blur in moving scenes since the vertical detail from the previous field is being combined with the temporally shifted current field.

The motion compensated de-interlacers, such as PixelComp™, clearly exceed both the linear and non-linear de-interlacers in image quality, but require significant compute resources to implement in real time. Fortunately, the Teranex Video Computer platform is capable of implementing these algorithms, as well as the non-motion compensated algorithms in real time, enabling the analog video signals of yesterday and today to seamlessly integrate into the digital television systems of tomorrow.

References

Deame, Jed, "Parallel Processing Solves the DTV Format Conversion Problem," *Proceedings of the 33rd SMPTE Advanced Motion Imaging Conference*, SMPTE, White Plains, NY, February 25, 1999, pp. 66-74.

Deame, Jed, "Teranex Fills In The GAPPs," *Videography*, June 1999, pp. 100-108.

Deame, Jed, "Teranex's PixelComp DTV Format Converter," *Broadcast Engineering*, September 1999, p. 138, 162.

de Haan, Gerard and Bellers, Erwin, "Deinterlacing – An Overview", *Proceedings of the IEEE*, Vol. 86, No. 9, Sept 1998, pp. 1839-1857.

de Haan, Gerard and Bellers, Erwin, "Deinterlacing of Video Data", *IEEE Transactions on Consumer Electronics*, Vol. 43, No. 3, Aug 1997, pp. 819-825.

de Haan, Gerard, "Motion Compensated De-Interlacing, Noise Reduction, and Picture-Rate Conversion", *IEEE Transactions on Consumer Electronics*, Vol. 45, No. 3, Aug 1999, pp. 617-624.