White Paper

Progressive HD Video in the Multi-Screen World

Mike Knee, Snell Ltd

June 2010

Abstract

In television production facilities, standard broadcast HD formats are rapidly giving way to full 1080-line 50 or 60 Hz progressive scanning with 3 Gbit/s interfaces, and are likely to be closely followed by stereoscopic 3D at up to twice that data rate. Such installations need to make the best of the pictures they receive, from mobile phone cameras, HDTV cameras, film scanners, 3D rigs or interlaced archive sources. Meanwhile, destination formats are proliferating, with increasing requirements for re-purposing of video material for displays ranging from tiny portable devices, through conventional TV screens, to high-end home theatre and 3D displays.

This paper gives an overview of many of the issues surrounding progressive production, including picture quality, flexibility for content repurposing, compression efficiency, dealing with legacy content and equipment, deinterlacing and interlacing, transparency and reversibility, latency in live production, and stereoscopic 3D processing.

The paper illustrates, with the support of examples and numerical measurements, that full 1080p 50 or 60 production capability is hugely beneficial even in mixed-format environments.

Introduction

Interlace served us well in the 20th century. It was the basis of all the analogue broadcast formats, providing a compromise between vertical and temporal resolution within the scanning rate and bandwidth constraints of cameras, transmission channels and displays. Interlace even survived the transition to HDTV, with 1080i being a popular format worldwide. However, several factors have now conspired to make progressive scanning (at frame rates of 50 Hz or higher) the format of choice for the future; this has long been recognized, for example by the EBU (1). The relevant factors include increased processing speeds and memory capacity, reduced storage costs, increasingly sophisticated processing, special effects and re-purposing, inherently progressive cameras and displays, and greater public awareness of picture quality. Computer-based portrayal of moving images has largely bypassed or ignored the existence of interlace. But in the world of broadcasting, continuity with the past is more important, for two reasons. There are millions of hours of content captured in interlaced formats, both at standard and high definition, and there are thousands of broadcast installations whose equipment is designed for interlaced signals.

This paper provides a technical discussion of the benefits of high frame rate progressive HD production and of the issues involved in handling "legacy" (lower-resolution and interlaced) signals in a progressive environment.

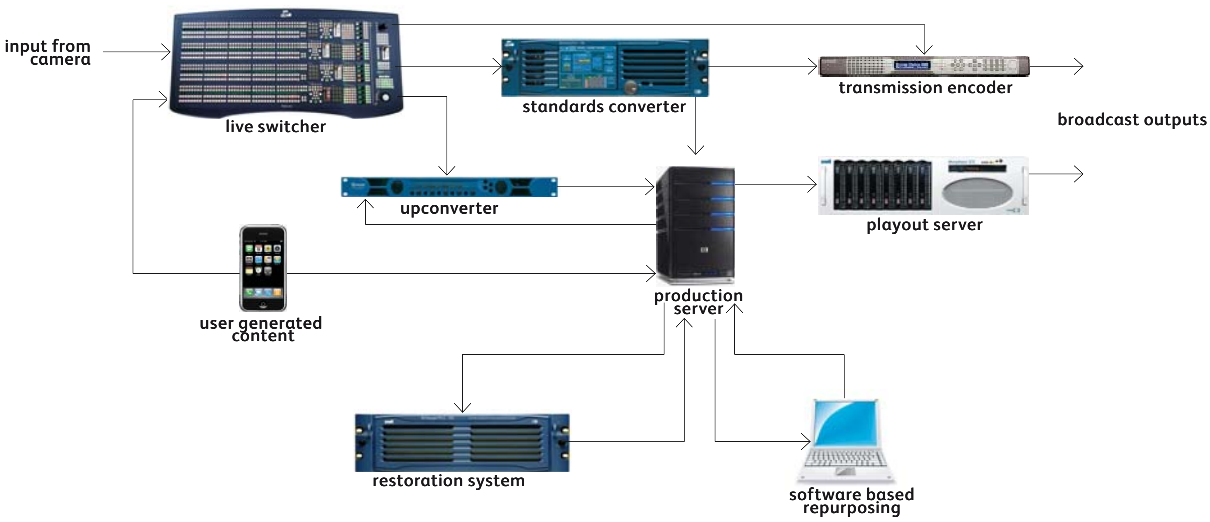

Figure 1: Elements of a television production chain

Figure 1 provides a starting point for our discussions. It attempts to summarise the workflows that might exist in both live and off-line broadcast production. In principle, almost all the signal paths shown could be in any video format, from existing SD or broadcast HD standards to high frame rate progressive HD formats in both 2D and 3D. We shall argue that a high frame rate progressive HD format should be used as widely as possible, regardless of the input and output formats being required at a particular moment or in a particular case. We shall show that conversion between formats should generally be minimized, though we also describe circumstances where unlimited, transparent upconversion and downconversion is possible and forms a practical part of the workflow. Particular attention will be paid to the handling of "legacy" formats within this structure, and to the impact of compression. The conclusions we make are backed up by simulations. There is not enough space in this paper to present all the simulation results, but a short case study will be presented as an example. This case study looks at the effect of moderate compression in different configurations of a production chain in which the original source is in a 1080p50 format. After the case study, we discuss in more general terms the results of other simulations and the implications of working with progressive signals throughout the production chain.

Case Study - Progressive Source with Compression

This example gives some simulation results based on a few possible workflows involving moderate compression of both interlaced and progressive versions of pictures whose original source format was 50 Hz 1080-line progressive. We first summarize the principles and assumptions on which the simulations were based.

Principles and Assumptions

Source Material

Most of the material we used in the simulations came either from the SVT test set (2) or from 72 Hz material treated as 50 Hz from the 2002-5 European METAVISION project (3). In the case study, one example was taken from each: CrowdRun, which is slow moving but with much detail, and Hurdles, which is less detailed but has fast, complex motion (even when slowed to 50 Hz). It is encouraging to note that the EBU is seeking to improve the range and availability of 1080p50 test material in its strategic programme on Harmonisation and Interoperability of HDTV Production Standards (HIPS) (4).

Compression

MPEG-2, Windows Media Video 9 and MPEG-AVC coding were used in the simulations, all at high bit rates with intraframe coding only. In the case study, MPEG-AVC coding was performed at 100 Mbit/s using the open-source x264 encoder.

Comparisons

Progressive outputs are compared to equivalent uncompressed progressive reference signals. Interlaced outputs could be compared to interlaced reference signals, but this would ignore the fact that modern displays are progressive. Interlaced outputs were therefore deinterlaced using a typical display deinterlacing algorithm and then compared to the progressive reference signals.

Error Measure

A combination of informal subjective assessment, luminance PSNR (5) and more sophisticated error measures such as Structural Similarity Index (6) and in-house visibility-weighted measures were used. In the case study, luminance PSNR values are reported, and this turns out to be give reliable comparisons, consistent with the other measures, in this domain of high quality and relatively few visible impairments or artefacts.

Deinterlacing

Any deinterlacing in the chain being tested (except at the display) used one of two algorithms, a high-complexity, high quality motion compensated algorithm previously described by us (7) and a simpler motion adaptive algorithm. In the case study, the simpler algorithm was used.

Interlacing

By default, interlacing is performed by discarding lines, but sometimes this is preceded by a vertical averaging filter as shown in the results.

Rescaling

This is an important operation, simulating the basic process that needs to be carried out for any reframing or re-purposing application. In the case study, rescaling was only carried out on progressive pictures and consisted of high-quality linear upconversion from 1080 to 1085 lines and cropping the result back to 1080 lines. This small change was chosen to minimize the change in picture content while ensuring that the interpolation algorithm was fully exercised.

Simulation Results

Figure 2 is a combination of a block diagram of the various signal flows used in the example, with PSNR graphs for the two source sequences. The blue graphs show results for progressive pictures and the red graphs show equivalent results for interlaced pictures (at the output of a display deinterlacer).

Here are some observations on these results

- It is always better to take a progressive output from the production process than an interlaced output. In some cases (IX and RIX) the differences are small, reflecting only the superiority of the motion adaptive deinterlacer over the simpler display deinterlacer. But in other cases, when there is any more complexity to the processing, the differences are much more marked, and would not be compensated for by any future improvements in display deinterlacing algorithms.

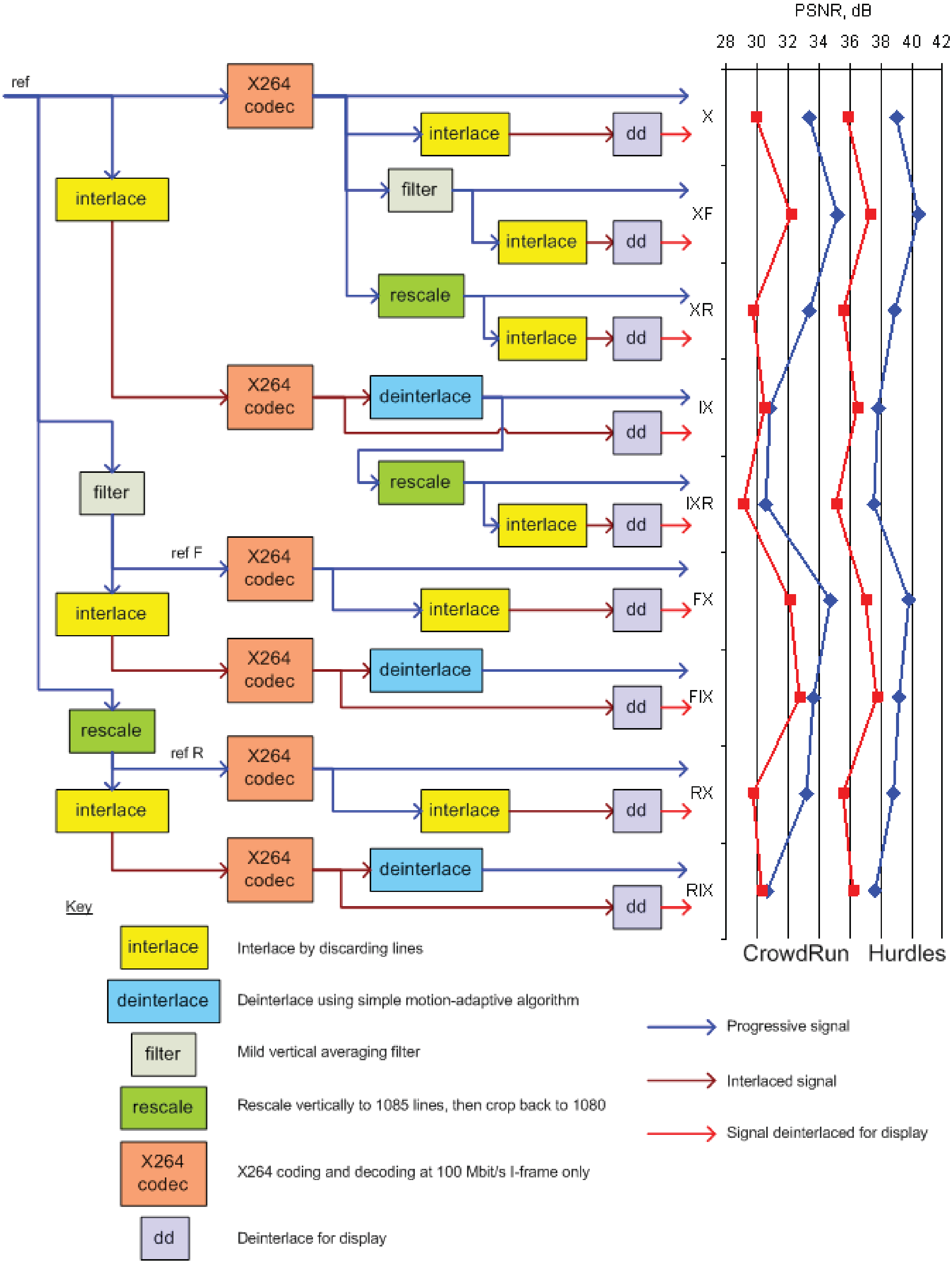

Figure 2: Case study block diagram and results

- Looking at progressive outputs, it is better to compress the progressive version of the signal than the interlaced version, even though the data rate at the codec input is twice as high. Compare points X to IX, FX to FIX and RX to RIX. The benefit is particularly marked where rescaling is involved.
- For progressive output, it does not matter whether rescaling is performed before or after compression (compare XR to RX).
- Mild vertical filtering prior to interlacing, which simulates an interlaced camera, is compression friendly (compare FX to X and FIX to IX), but it is even better to perform this filtering after the compression (compare XF to FX). This means that if it is desired, it only needs be done just prior to generating an interlaced output.

Other Simulation Results

We now summarize results taken from a large number of tests that were similar to the above but performed over a wider range of conditions and source material.

- If the original source is interlaced, it can be more efficient to compress the interlaced signal directly, but only for slow-moving material.
- With an interlaced source, if any processing such as rescaling is performed, it is better to deinterlace the input as soon as possible, regardless of picture content.
- A high-quality motion compensated deinterlacer is preferable, but the distinction between that and a simpler, motion-adaptive deinterlacer becomes less important as the compression becomes stronger.
- If an interlaced source has had some vertical bandwidth reduction, for example coming from an interlaced camera with a progressive sensor followed by line averaging, some mild vertical equalizing boost can be beneficial either before or after the compression, provided the noise level of the source is low.
- The conclusions are independent of the choice of (high bit rate intraframe) compression.

Implications of a Progressive Production Format

The main conclusion from the simulations is that it is beneficial to work with progressive signals as much as possible. We now discuss the implications of this practice on various aspects of the production workflow.

Interlaced Inputs and Outputs

Figure 3 shows that there are four possible signal flows through a production plant, depending on whether each of the source and destination is interlaced or progressive. If the main body of the plant works with progressive signals, then interlaced inputs will need to be deinterlaced, and interlaced outputs will need to be created by interlacing, as shown in the figure. Once conversion in either direction is carried out, knowledge of whether the original source was interlaced or progressive is lost unless specific steps are taken to preserve the history as metadata, in much the same way as MOLE (8) (in a compression chain) or the Video Index (9). It is possible that interlaced signals emerging from the plant will re-enter the same or another plant, bringing a potential for multiple cascading of interlacing and deinterlacing processes.

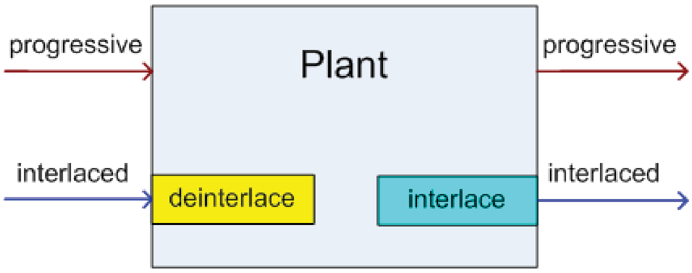

Figure 3: Input/Output Conversions

In order to ascertain whether a metadata signalling system is required, let us look at the interlacing and deinterlacing processes in turn.

Interlacing

Essentially, interlacing, for example conversion from 50 frames per second to 50 interlaced fields per second, consists of throwing away alternate lines in a vertical-temporal quincunx pattern. It may be expected that some pre-filtering should be applied to the progressive signal in order to reduce aliasing. However, since progressive signals are our reference, the efficacy of such filtering should be judged according to the quality of a subsequent deinterlaced signal. Simple deinterlacers might do better with signals that were interlaced with pre-filtering. However, a motion compensated deinterlacer such as that described in (7) can be optimized for its performance on non-pre-filtered interlaced signals. High-level knowledge of the motion and spatial detail of the scene can be exploited to remove the effects of almost all the aliasing introduced by the interlacing process. So we have found that the best way to interlace a signal, if there is any chance of subsequent processing, is to throw away lines. If the only remaining process is a low-quality deinterlacer in a display, then some pre-filtering may be advantageous, but as display deinterlacing improves, the benefit of such pre-filtering is likely to diminish.

Deinterlacing
The aim of the deinterlacer is to make from an interlaced source a picture that is as close as possible to one originating from a progressive source. Modern interlaced cameras usually consist of progressive capture followed by a vertical averaging process to create the interlaced output. Ideally, a deinterlacer should equalize the effect of this filter, adding some vertical boost, to produce an approximation to the original captured progressive signal. Such an equalizer could also be applied following any interlacer that included a similar filter. In principle, then, an interlaced signal could carry metadata indicating what vertical filtering it had undergone, so that the deinterlacer could apply appropriate equalization. The deinterlaced signal would need to preserve the metadata and add a record of the equalization process. However, the benefits of such equalization have been found to be small and may not justify the increase in housekeeping required.

Another issue with deinterlacing is that it is not strictly necessary to leave input lines untouched. In some deinterlacing algorithms, the visibility of any "mouse teeth" artefacts that have not been removed by motion compensated interpolation can be significantly reduced by a vertical filter applied to the output. However, this approach is not necessary in high-quality motion compensated deinterlacing and it loses the reversibility of deinterlacing followed by interlacing.

Recommended Approach
So here is a summary of a recommended approach to handling interlace in a progressive environment:

- deinterlace the input straight away, using a high quality motion compensated deinterlacer, and retain the progressive version of the signal through all subsequent processing and compression;
- if interlaced outputs are required which might undergo further processing, interlace by throwing away lines;
- interlacing for transmission or display may be preceded by a mild vertical filter;
- a system of equalizing deinterlaced signals that may have come from softer sources (including vertical filters) may improve quality in some circumstances but is probably not worth the additional housekeeping overhead and uncertainty.

Re-Purposing (The Multi-Screen World)
The main operation involved in any kind of re-purposing for different display platforms is some kind of rescaling of all or part of the picture. Realistically, rescaling requires progressive pictures, except for the special case of horizontal-only rescaling. And even if the required output is thought to be interlaced, it is better to keep the output progressive and only create the interlaced version at the final output.

Re-purposing may also involve complex analysis of the picture material, for example to find the region of interest in a dynamic reframing or intelligent aspect ratio conversion algorithm (10). For such analysis tasks a progressively scanned picture is essential.

Transparent Upconversion
One of the exciting recent developments designed for a workflow based on progressive signals is that of transparent upconversion, where a subsequent downconversion process can produce an output that is identical (to an arbitrary accuracy) to that of the original input. This is possible because of the redundancy inherent in the upconversion process. Because more than one output sample is generated for each input sample, there are infinitely many ways to upconvert a signal and retain the ability to downconvert transparently. By careful design and optimization of upconversion filters, it is possible to upconvert with high quality in such a way that the inverse downconversion filter also has high quality when applied to a native high-resolution signal, as described in a patent application by Weston (11).

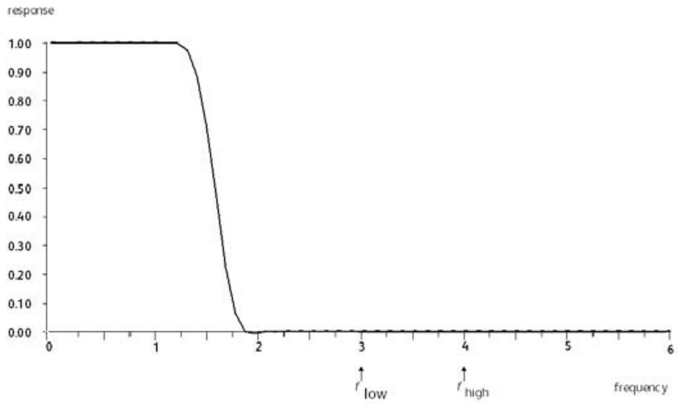

Figure 4: Transparent conversion filter response

Figure 4 shows the frequency response of both the upconversion and downconversion filters for conversion in the ratio 3:4 (for example, for conversion between 1440 and 1920 samples per line). The frequency axis is in units of 1/3 of the lower or 1/4 of the higher sampling frequency. This is a very flat frequency response with good stopband performance, yet the cascade of upconversion and downconversion processes is transparent to better than one 10-bit quantizing level.

When combined with the approach to deinterlacing and interlacing outlined above, the notion of transparency extends to conversion from any interlaced format to any progressive format with higher spatial resolution.

Latency
The top part of Figure 1 shows a live production chain. In live production, latency or overall delay may be important. This may be construed as an argument against progressive production in a simpler production chain which could be kept in interlaced format. However, even in live production there are occasions in which progressive processing is required, so it still makes sense to convert any interlaced inputs to progressive at the earliest opportunity. Even the very best deinterlacers have a latency of only two field periods, and it is better to suffer this latency once than to risk accumulating delay in multiple deinterlacing and interlacing operations.

3D Considerations

Stereoscopic 3D television unfortunately doubles the uncompressed bandwidth requirements for TV production. This means that it might be tempting to go back to an interlaced production format in order to fit both channels of a 3D signal into a 3 Gbit/s container. However, we have seen that in the compressed domain we do not save bandwidth by using interlace. One of the key considerations of 3D is that the two channels should undergo equivalent processing as far as is possible. Artefacts or impairments that might be tolerable in a 2D picture might be more objectionable if they occur in the left eye but not the right eye, for example. So it is important to minimize the number of high-level processing operations (such as deinterlacing) in which different decisions might be made in the two channels. We have demonstrated in practical, on-air experiments that good-quality standards conversion performed separately on the two channels gives perfectly acceptable 3D quality. However, that was with a standards converter that makes significant use of global information in its motion measurement. More localized algorithms, whether for standards conversion or for deinterlacing, are liable to make incompatible switching decisions in the two channels, leading to greater visibility of artefacts.

Extending the Chain

So far in this paper we have concentrated on the production part of the broadcast chain and have made the case for the use of progressive formats where possible, even when the source and destination of the signals are in interlaced formats. In this section we make the case for extending the use of progressive formats across the whole chain, as shown in Figure 5. The main argument is that both cameras and displays already work with high frame rate progressive formats and only ever need to deal with interlace because of legacy equipment and transmission paths. Consumers are becoming acquainted with 1080p50 and 60 pictures through Blu-ray DVD, games consoles and camcorders. Consumer DVI and HDMI interfaces handle these formats and there remain few obstacles to the final removal of interlace from everywhere except archive handling. The biggest remaining obstacle is, of course, the final transmission path and infrastructure. In the UK, the regulatory body Ofcom has considered the use of MPEG AVC Level 4.2, which would cover 1080p50 transmission, in the DVB-T2 multiplex. There is some evidence that, on a simplistic comparison, there is a small bit rate penalty in transmitting 1080p50 compared with 1080i25. However, the comparisons do not take into account the inherent picture quality benefits of progressive over interlaced scanning or the potential quality compromise resulting from deinterlacing in the display. Put simply, interlace is not a particularly efficient form of compression and it would be counter-productive to be stuck with it.

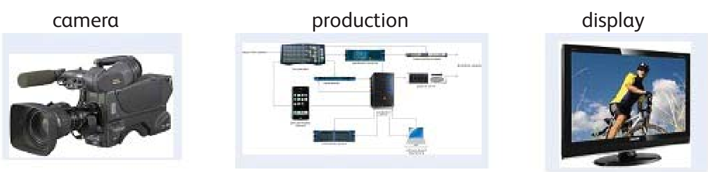

Figure 5: Extending the chain

Conclusions

We have discussed the issues surrounding the use of high frame rate progressive HDTV signals in television production, especially with regard to mixed input and output formats and to the effects of compression. We conclude that 50 or 60 Hz progressive formats should be used as widely as possible, with upconversion being performed early in the production chain and with any downconversion being performed only when required.

References

1. EBU. 2005. FUTURE HIGH DEFINITION TELEVISION SYSTEMS: The need to develop television production equipment for a progressively scanned image format of 1920 horizontal by 1080 vertical resolution at 50 and 60 Hz frame rates. EBU Technical Recommendation R115-2005.

2. Haglund, L. 2006. The SVT High Definition Multi Format Test Set.
Document available at http://media.xiph.org/svt/SVT_MultiFormat_v10.pdf

3. IST METAVISION Project, 2000-03. Search for Metavision at http://www.ist-world.net

4. Nufer, Christoph. 2010. HIPS. tech-i (Insight from EBU Technical) issue 03, March 2010, page 10.

5. Sayood, Khalid. 2005. Introduction to Data Compression. Morgan Kaufmann, 3rd edition, pages 198-9.

6. Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, 2004. Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, Apr. 2004.

7. Knee, M J. 2006. International HDTV content exchange. Proc. IBC 2006.

8. SMPTE. 2000. MPEG-2 video recoding data set. SMPTE Standard 327M-2000.

9. SMPTE. 2008. Video Index Information Coding for 525- and 625-Line Television Systems. SMPTE Recommended Practice RP 186-2008.

10. Knee, M.J and Piroddi, R. 2008. Aspect processing – the shape of things to come. Proc. IBC 2008 (re-published as “Aspect ratio problems in television today and some new solutions” in SMPTE Motion Imaging Journal 2010 vol 19 no 1 pp 35-41)

11. Weston, M. 2008. Digital image processing. International Patent no. WO 2009/138804 A2

Acknowledgements

The author would like to thank the Directors of Snell for their permission to publish this paper.