On Coding Efficiency and Scan-Rate Conversion

Erwin B. Bellers, and Gerard de Haan,
Philips Research Laboratories, Eindhoven, The Netherlands,
phone +31 40 2744285, email: erwin.bellers@philips.com

Abstract: This paper examines the coding efficiency of MPEG-2 coding for interlaced and progressive video, and also compares de-interlacing and picture rate up-conversion before and after coding. We found receiver side de-interlacing and picture rate up-conversion (i.e. after coding) to give better image quality at a given data rate. In contrast with some other publications, we found interlaced video coding to be better than progressive video for many relevant sequences, even when comparing the results on a progressive display.

Keywords: MPEG-2, coding, de-interlacing, picture-rate up-conversion.

1 Introduction

The human visual system (HVS) is less sensitive to flickering details than to large-area flicker [1]. Since the very beginning, television broadcast standards, apply interlacing to profit from this fact. Interlace, however, complicates many image processing tasks [2]. Particularly, it complicates scanning-format conversions. These conversions were necessary in the past mainly for international programme exchange, but with the advent of high-definition television, digital video broadcast, videophone, the Internet, and video on PCs, many scanning formats have been added to the broadcast formats, and the need for conversion between formats is increasing [3].

The increasing need for consumer video format conversion together with the new options for digital video broadcasting have restarted the discussion on interlace in broadcast standards. Although in earlier publications [4, 5, 6] a comparison of the effectiveness of MPEG coding on interlaced and progressive video has been reported, there is reason to believe that some very relevant aspects are missing in the research published so far. Particularly, the effect on the blockiness of the decoded images has not been investigated, and perhaps even more important, no subjective evaluation was presented for average bit rates. Moreover, a high quality de-interlacer is missing, although commercially available [7]. Finally, we found that most papers focus on near stationary, but highly detailed image sequences and largely neglect the coding efficiency for less detailed sequences with stronger motion.

We believe that the common opinion of interlace versus progressive video is biased towards progressive, while a profound comparison between both formats is missing. In this paper we, therefore, restudied the MPEG-2 coding efficiency for interlaced and progressive video both objectively and subjectively. Moreover, we compared de-interlacing at the transmitter and at the receiver side, as this up-conversion to a higher level is not an uncommon practice for broadcasts using the ATSC (Advanced Television Systems Committee) digital television standard.

Similarly, we expected a change in coding efficiency as a result of picture-rate up-conversion. This is particular relevant for motion portrayal improvement of movie material [8], as the picture rate of film differs from that of the display. Again we compared the overall quality difference for picture-rate upconversion at the transmitter and at the receiver side.

The evaluation includes objective comparisons by means of the Peak Signal-to-Noise Ratio (PSNR) and a Block Impairment Metric (BIM) [9]. The subjective assessment includes both expert and non-expert viewers.

Section 2 describes the experimental setup, while section 3 briefly addresses the algorithms used. Section 4 shows the evaluation results and finally, section 5 presents our conclusions.

2 The experiments

We studied three cases:
1. A comparison between a progressive-coding chain versus an interlaced-coding chain (figure 1).

2. A comparison between transmitter side and receiver-side de-interlacing (figure 2).

3. A comparison between picture-rate upconversion at the transmitter and receiver side (figure 3).

We shall refer to the ‘progressive coding chain’ if the complete chain including the video source and the display are progressive. The chain that contains a progressive video source and display, but an interlaced codec is referred to as the ‘interlaced coding chain’. In both chains the source and the display are progressively scanned. The progressive chain has the evident advantage that interlace artifacts will be totally absent, whereas the interlaced coding chain has the advantage of a reduced data rate to the encoder. The question to be answered by the experiments, is which of these advantages is largest in terms of picture quality.

In the second experiment, we investigated the coding efficiency of receiver and transmitter side de-interlacing. This comparison seems relevant, as a significant percentage of the video material is only available in an interlaced format. The advantage of transmitter side de-interlacing is that coding artifacts cannot affect the de-interlacer. The drawback is the increased data rate to the encoder, which may lead to a decreased image quality at a given data rate.

A breakthrough in motion estimation in the beginning of the nineties, enabled sophisticated motion-compensated picture rate up-conversion techniques integrated on a single consumer priced IC [8, 10]. These ICs were partly designed to improve the poor motion portrayal of film on TV, which is due to the limited 24, 25 or 30 pictures per second of film originated video. As such motion portrayal improvement can be realised either at the transmitter, or at the receiver side, we compared the resulting picture quality for both situations, assuming MPEG-2 coding in our third and last experiment.

3 The algorithms and settings

The picture-rate up-conversion and de-interlacing algorithms are those presented in [10], which are available on a commercial IC. The interlacing is just a simple sub-sampling, while the MPEG-2 codec that we used in our experiments is the Berkeley MPEG-2 codec [11]. This codec is referenced by and available to many, and complies with the test model 5 (TM5) of the MPEG-2 standard.

In line with earlier publications [12, 4, 6, 13], we chose an equal number of I and P pictures per Group Of Pictures (GOP) for progressive and interlaced coding. Moreover, the temporal instances of the I and

\(^1\)To be entirely correct, we further developed the de-interlacing method described in [7].
P pictures are chosen to be identical for progressive and interlaced coding. With a GOP representing 0.5 seconds of video, we used this for most experiments, again in agreement with settings found in the literature [12, 4, 6], \( N_{GOP} = 12 \) and \( M_P = 3 \) for the interlaced coding chain, and \( N_{GOP} = 24 \) and \( M_P = 6 \) for the progressive-coding chain (see also figure 4).

The upconversion of 24 Hz movie material to 60 Hz, used in the third experiment, required a different setting to maintain identical temporal positions for the I and P pictures. We used \( N_{GOP} = 12 \) and \( M_P = 2 \) for the receiver-side up-conversion and \( N_{GOP} = 30 \) and \( M_P = 5 \) for the transmitter-side up-conversion.

The progressive coding chain applies frame prediction, i.e. predictions are formed from reference frames [14]. For the interlaced chain we experimentally found that field based prediction resulted in (slightly) better quality. This setting was, therefore, used in the evaluation.

The main profile main level does not support 50 Hz or 60 Hz progressive coding. Therefore, we applied the main profile main level to both the 50 Hz interlaced and the 24/25 Hz progressive coding experiments, while the main profile high level was used to the 50/60 Hz progressive coding experiments.

In agreement with the experiments found elsewhere in the literature [12, 6, 4], we applied the bit rates 2Mb/s, 4Mb/s and 6Mb/s (for the luminance video data only).

4 The quality criteria

Although much criticized, it seems that the common Mean Squared Error (MSE) is still the most generally accepted measure, in coding papers usually presented as a Peak Signal-to-Noise ratio (PSNR). We use the following definitions:

\[
MSE(n) = \frac{1}{N} \sum_{\vec{x}} (F_{\text{org}}(\vec{x}, n) - F_{\text{out}}(\vec{x}, n))^2
\]

and:

\[
PSNR(n) = 10 \log \left( \frac{255^2}{MSE(n)} \right)
\]

respectively, where \( F_{\text{org}}(\vec{x}, n) \) is the original luminance source sequence, \( N \) the number of pixels, and \( F_{\text{out}}(\vec{x}, n) \) is the output signal at the receiver side. We shall use PSNR for the average over all frames \( n \) of the sequence.

However, the main defect of block-coding algorithms, is the resulting blockiness of the decoded images. This artifact is not well reflected in the resulting MSE and PSNR figures. To quantify the blocking artifacts of the decoded images, we apply the Blockiness Impairment Metric (BIM) designed for this purpose by Wu and Yuen [9], which is reported to give an objective figure well representing the subjective impression of the blockiness of the image.

Horizontal block edge impairment, \( BI_h(n) \) is measured as:

\[
BI_h(n) = \left[ \sum_{\vec{x} \in HE} (W(\vec{x}, n)(F(\vec{x}, n) - F(\vec{x} - \vec{u}_x, n)))^2 \right]^{\frac{1}{2}}
\]

where \( \vec{u}_x = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \) and where \( HE \) is the set of locations where we expect horizontal block edge impairments (assuming block sizes of 8 by 8 pixels), i.e.:

\[
HE = \{ \vec{x} = (x, y)^T | x \mod 8 = 0 \}
\]

with \( T \) for Transpose and \( W(\vec{x}, n) \) a weighting coefficient to take into account spatial masking and luminance level visibility dependence effects. To this end, \( W(\vec{x}, n) \) depends on a local average luminance level and a local variance [15, 9].

The vertical block edge impairment, \( BI_v(n) \), is defined similarly. The average of \( BI_h(n) \) and \( BI_v(n) \) is denoted as \( BIM(n) \), that we shall average over the total sequence to obtain \( BIM \) for that sequence.

5 The test set

It is not unlikely, for either of the two test criteria, that the outcome will depend on the selected picture material. Evidently, sequences with abundant vertical detail and little motion will negatively bias the resulting performance of the chain that relies on de-interlacing. Nevertheless, this choice is often seen in
publications on this topic, even though one can argue whether these sequences represent the average broadcast material very well.

We tried to prevent such a bias in our evaluation, by adding a second set of sequences with less abundant vertical detail. As a possible advantage of a decreased data rate to the encoder will not become evident if the encoder is not stressed at all, we chose the scenes in the second set to have clear motion. Snapshots of each image sequence that we selected are shown in figures 5(a,b,c) (‘first test set’) and 5(d,e,f) (‘second test set’).

6 Evaluation results

We shall present the evaluation results of the three experiments in separate subsections.

6.1 Progressive versus interlaced

Figure 6 shows the difference in PSNR and BIM score of the progressive and interlaced coding chain. A positive difference indicates a preference for the progressive, and a negative number for the interlaced chain. We can observe that the progressive chain is advantageous in terms of PSNR for three out of six of the evaluated conditions, whereas the interlaced chain is advantageous for all bit rates in terms of the BIM figure. The difference in the BIM, however, is small for the interlaced chain using the first test set. This had to be expected as the blockiness metric takes the reduced visibility of blockiness in detailed image parts into account.

The conclusion is quite different for the second test set. We observe a significant difference in BIM scores. The large difference found in the BIM scores indicates a clear preference for the interlaced chain. The large difference in the BIM figures can be ascribed to: 1) less data has to be compressed, as the interlacing process sub-samples the video, and 2) de-interlacing causes ‘interleaving’ of blocking artifacts as the interpolated lines are reconstructed from the current and neighbouring pictures.

6.1.1 Subjective assessment

As for some conditions, the PSNR and BIM difference have opposite signs, it remains uncertain which of the options is best. We, therefore, set up subjective assessment of the output sequences. The assessment provides an option to weight the objective figures with respect to the subjective quality. We adopted the ‘stimulus-comparison’ method conform the CCIR 500-4 recommendation [16], and originally used two groups of expert and non-expert viewers, respectively. It turned out, however that the difference between these groups was negligible. Therefore, we combined the results of all 23 observers. Figure 7 shows the over-
all results.

We can draw two conclusions from figure 7: 1) the progressive-coding chain is to be preferred for the nearstationary sequences containing abundant vertical detail at bit rates of 4 and 6 Mb/s, whereas 2) the interlaced-coding chain is always the preferred one for sequences containing less detail and stronger motion.

If we compare the subjective scores of figures 7b with the objective results from figure 6b at 6 Mb/s, it is remarkable that even a relatively small difference in the blockiness is subjectively found to be dominant over a large difference in the PSNR. As illustrated in figure 8, it can be seen that this blockiness can be very annoying, especially for the lower bit rates, while it is still visible at the higher bit rates. We also conclude that earlier conclusions in the progressive versus interlaced coding debate based on comparing only PSNR scores, have to be considered irrelevant.

6.2 Transmitter side versus receiver-side de-interlacing

The results that we obtained with receiver versus transmitter side de-interlacing are shown in figure 9. The figure shows the advantage of receiver-side de-interlacing over transmitter-side de-interlacing. Both

---

**Figure 6**: Performance difference between the progressive and the interlaced coding chain.

**Figure 7**: Results of subjective test showing the performance difference between the progressive and the interlaced coding chain. The error bar indicates the 95% confidence interval.

**Figure 9**: Performance difference of the receiver side and transmitter-side de-interlacing.
the PSNR and the BIM, and an informal subjective evaluation, indicate a clear preference for the receiver-side de-interlacing. We thought it redundant, therefore, to confirm this with a formal subjective test.

Again, the low BIM scores for the first test set had to be expected as the BIM takes into account the reduced visibility of blockiness in detailed image parts. The large difference in the BIM scores for the second test set are ascribed to the same two reasons as mentioned previously. Obviously, the PSNR and BIM figures decrease with increasing bit rate. In case of very high bit rates (bit rate >> 6 Mb/s) the codec hardly causes a loss of information. Consequently, the performance obtained with transmitter side and receiver side de-interlacing become equal in the end.

We conclude that receiver side de-interlacing is advantageous over transmitter side de-interlacing.

6.3 Transmitter side versus receiver-side picture-rate upconversion

Figure 10 shows the results of transmitter side versus receiver-side picture rate up-conversion (25 to 50Hz), and figure 11 shows the same comparison for a conversion from 24 Hz to 60 Hz. These results indicate a very clear preference for the receiver-side picture-rate upconversion for both the PSNR and the BIM measure. We therefore felt it redundant to confirm this with a formal subjective assessment.

The difference of the BIM figures is only significant for the bit rate 2Mb/s score on the second test set. As before, the decrease of the PSNR and BIM differences with increasing bit rate is rather obvious.

Figure 10: Performance difference of the receiver and transmitter side picture rate up-conversion of 25 Hz to 50 Hz.

7 Conclusions

We conclude that video format up-conversions (for film motion portrayal improvement and de-interlacing) give a better picture quality when performed at the receiver side.

We also conclude that the preference for interlaced or progressive encoding depends on the statistics of the programme material. Particularly, we found a different conclusion for near stationary sequences with abundant vertical detail (slight preference for progressive coding), and sequences with less abundant detail but more challenging motion (clear preference for inter-
laced coding). Earlier publications definitely overstate the progressive case by exaggerating the importance of the peak signal to noise ratio, and focussing on test sequences with abundant vertical detail.

![Graph](image)

Figure 11: Performance difference of the receiver and transmitter side picture rate conversion of 24 Hz to 60 Hz.

References


