

# INTERLACE AND MPEG – CAN MOTION COMPENSATION HELP?

J. O. Drewery  
British Broadcasting Corporation, UK

## 1. INTRODUCTION

Interlaced scanning was invented as a way of halving the transmission bandwidth of the video signal without appreciably sacrificing the portrayal of motion or detail. The cost came to light only later when scanning spots became smaller and consequently resolution improved; the cost appeared as an inter-line flicker at the picture frequency on stationary objects. Nowadays, the source scanning aperture is deliberately tailored to minimise this effect.

Developments in technology have allowed us to decouple the scanning operations at the source, display and in the transmission path, albeit with a loss, to some degree, in quality. The question is, therefore, do we still need interlaced scanning, now that we can save bandwidth in other ways?

## 2. THE CURRENT PRESSURES TO ABANDON INTERLACED SCANNING

The current pressures to adopt non-interlaced scanning are coming from two directions. The first is from the field of compression, specifically in conjunction with the MPEG coding system. This system is basically an internationally agreed tool kit for reducing the bit rate of raw video data and has achieved spectacular results, precipitating the digital transmission revolution which is now gathering pace.

The tool kit contains, amongst other things, the ability to adapt its processing according to the scanning

format of the input video. This adaptation is not, however, limitless and only a restricted set of formats is catered for. This set includes both standard and high definition numbers of lines and both interlaced and non-interlaced formats but the combination of the standard number of fields per second with the standard number of lines, *non-interlaced* is not included. However, some workers in the field have purported to show that MPEG would be far more efficient if it were operating with this format. We will return to this matter later.

It should be noted that the field of compression has a long history, involving the techniques of DPCM, transform coding, sub-sampling and entropy coding in various combinations. The MPEG system brings them all together, in conjunction with motion estimation, which is a relative newcomer. Experience of early compression systems showed that the most powerful method of compression was the reduction of sample rate, obtained by sub-sampling, giving careful attention to the spectral region that could be supported by such a reduction. Although interlaced scanning is such a sub-sampling technique, early compression systems never regarded it as such, but rather a factor that could not be altered, and it has only been the flexibility of the MPEG system that has allowed it to be challenged.

The other direction from which the pressure comes to abandon interlace is that of the IT industry. The transmission of pictures by IT has its roots in the electronic graphics systems invented in the late 70s and early 80s. These systems were, initially, based on the manipulation of still images, describing the picture in terms of  $x$  and  $y$  co-ordinates. In such a regime, the

concept of interlace was irrelevant; only the pixel addresses and values were important. Moreover, as pixels were initially allowed to take only a small number of values, typically two for characters, the requirement for more resolution implied more pixels. Character generators were sold on the basis of their clock frequency, rather than the quality of their characters. In vain did video engineers try to influence the situation by pointing out that varying the position of an object could be brought about by filtering, allowing the pixels to take on more values. In fact it was quite amusing for the video fraternity to observe the electronic graphics engineers rediscovering the truths about video, such as aliasing, which had been known since the 30s.

As video graphics came from a different place, it was inevitable that a divide, of sorts, has continued to exist. This is now manifest in the different attitudes to scanning formats. Most VDUs use non-interlaced scanning with a field frequency in the range 70-80 Hz and many VDUs still have two-value pixels which require a spatial resolution finer than that for HDTV. These 'excesses' can be tolerated because the displays are not concerned, primarily, with the transmission of moving images derived from the real world and because there is no local bandwidth problem.

One reaction to this plethora of scanning formats has been the setting up of the TRANSIT project under RACE whose remit is to investigate conversion between formats whilst preserving the best image quality. Some work carried out under this project will be reported later.

### 3. OTHER PRESSURES TO ABANDON INTERLACE SCANNING

At IBC86, Clarke presented a paper[1] recommending the use of non-interlace scanning in studio processing and transmission. Although conceding that interlace could be regarded as a means of display improvement, he argued that up-conversion could be better done by starting from a non-interlaced format. He adduced the

further argument that, starting with non-interlaced scanning, up-conversion can be separable in the vertical and temporal directions, that is, the one would not affect the other. One corollary alleged to flow from this would be that derivation of motion vector information, a key component in improved up-conversion, would be easier because of the 'coherent temporal aliasing'.

If non-interlace scanning were adopted then signal processing operations such as still frame would be easier and besides, camera sources were alleged to be unable to generate vertical detail beyond the Nyquist limit for the field, because the spot size of the low-velocity scanning beam would extend to the line pitch of the field.

These arguments now begin to draw out other facets – the temporal aliasing inherent in video scanning and the deficiencies of cameras.

### 4. A CLOSER LOOK AT THE COMPRESSION ARGUMENT

Work reported elsewhere[2] has shown that, starting with pictures scanned in a non-interlaced fashion, a MPEG system, working at a given bit rate, delivers a better picture to a non-interlaced display than one preceded by a non-interlace to interlace down-converter and delivering to an interlace display. With reference to Fig. 1, this is the route

625/50/1:1 – MPEG – 625/50/1:1

compared with

625/50/1:1 – convert2 – MPEG – 625/50/2:1

This is not entirely surprising, since, if the MPEG codec were replaced with a piece of copper wire, the second route would have lost information compared with the first. However, as the bit rate is reduced, one might expect the coding loss to dominate over the conversion loss, and that is, indeed, what happens. It

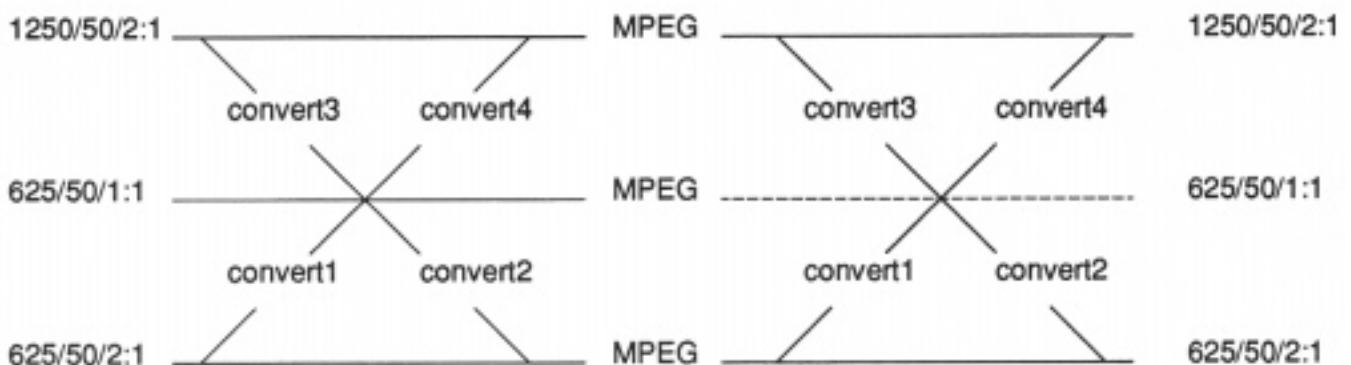


Fig. 1 - Possible routes for testing scanning formats in combination with MPEG coding.

should be pointed out that the filtering involved in the down-conversion is crucial; merely omitting lines would lead to a very unsatisfactory situation.

It would be instructive to compare the routes

625/50/1:1 – MPEG – convert2 – 625/50/2:1

and

625/50/1:1 – convert2 – MPEG – 625/50/2:1

for the same bit rate, for this would show whether or not the MPEG system is at a disadvantage when coding interlace signals.

As it is, the original experiment shows that, *for high bit rates*, the better way of delivering a non-interlaced signal to the viewer is by way of straightforward MPEG coding and non-interlaced display rather than down-converting to interlace and MPEG coding with a lower compression rate. It is therefore tempting to say two units of MPEG compression is better than one unit of interlace compression plus one unit of MPEG compression. Therefore interlace is a poorer method of compression than MPEG.

Unfortunately, matters are not as simple as that. For, by presupposing the format of the input signal, this experiment does not tell us about the best way of conveying information about the real world to the viewer. Laying aside the exact modelling of the MPEG codec, given that a 50 Hz non-interlaced mode does not exist, suppose the routes in Fig.1

1250/50/2:1 – convert3 – MPEG – convert4 – 1250/50/2:1

and

1250/50/2:1 – MPEG – 1250/50/2:1

had been investigated. The input signal to the MPEG coder would have had the same bandwidth as the 625/50/1:1 signal and so the compression ratio would have been the same. Would the quality of the final displayed pictures have been the same, or better, or worse than that in the experiment?

My contention is that both methods would have delivered a better quality, the second better than the first, for the reasons that will now be discussed. In other words, the experiment has not proved the case for abandoning interlace.

## 5. THE CASE FOR INTERLACE AT THE SOURCE

Irrespective of coding efficiency, one has to get a

signal out of a source and into a display. Sources are continuous (vacuum tubes) or discrete (CCDs) and have to be addressed in some way to provide a signal. There is then a bandwidth limit at this point. Similarly, displays are continuous (CRTs) or discrete (LCDs, flat panels) but have to be addressed, with a corresponding bandwidth limit. Given the bandwidth limit, the question arises as to what is the optimum strategy for each of these devices, that is, how may the bandwidth resource be optimally disposed?

For stationary material the appropriate source scanning mechanism is interlaced as it contains more information than a non-interlaced scan of the same line frequency. For example, a 1250/50/2:1 scan contains more information than a 625/50/1:1 scan. As images contain a lot of material that is stationary, interlaced scanning thereby increases the chances of conveying more information, on the average.

When objects move it is arguable whether we can know more or less about them if the scan is interlaced or not since motion of an object with respect to a fixed scan is the same as motion of a scan with respect to a fixed object. With an interlaced scan, if an object moves vertically at the strobe speed, so that the lines of successive fields fall on the same place, we know less about it than when it is stationary. On the other hand, if it moves at other speeds we know more, provided we have access to enough fields. This happens with a non-interlaced scan too, except that we know least about it when it is stationary, which seems the wrong way round.

Moreover, by introducing horizontal scanning in the form of sampling, the concept of interlace can be extended to two dimensions. This mimics what the human eye does. The image on the retina is incredibly crude but the eye is constantly shifting it about, by a process called nystagmus, so as to explore it more fully with its limited resources. The bandwidth resource becomes a sample density resource.

This idea is just sub-sampling in another context. For example, it lies behind the proposal from the EUREKA 95 consortium[3] for the intermediate scanning standard which was dubbed 'high definition quincunxial', HDQ, intended as a stepping stone from interlaced to non-interlaced scanning at high definition. As shown in Fig. 2, it consists of even samples on even lines and odd samples on odd lines in the first field and the remaining samples on the second field. It has the same sample density as normal interlaced scanning sampled orthogonally, but appears subjectively much better because the spectrum associated with inter-line flicker acquires a horizontal component.

By extension to two dimensions, it becomes easier to avoid the strobe problem because sampling structures

1	2	1	2	1	2
2	1	2	1	2	1
1	2	1	2	1	2
2	1	2	1	2	1
1	2	1	2	1	2
2	1	2	1	2	1

Fig. 2 - Sample pattern of HDQ.

1		3		1		3	
	2		4		2		4
3		1		3		1	
	4		2		4		2
1		3		1		3	
	2		4		2		4

Fig. 3 - Sample pattern of MUSE.

can be devised having strobe motions that do not correspond to simple translational motion. This also becomes easier as the interlace order of the structure increases. For example, the two-field structure of HDQ has an obvious strobe in the diagonal directions but the four-field structure of MUSE[4], shown in Fig. 3, has no translational strobe. In fact the MUSE structure is the optimum four-field structure, from the point of view of alias amplitude.

From a practical point of view, it can be argued that, if the source is discrete, such sub-sampling techniques would result in an inefficiency in the use of resources since each site would be used for only a fraction of the time. This can be countered by observing that it is possible to form a signal by combination of adjacent site values *on the device*, as is already done to avoid inter-line flicker, as mentioned earlier. In this way, all site values are used in every field, and there is no sacrifice of signal-to-noise ratio. The combination simply acts as a crude pre-filter.

Linking this to the situation for continuous sources, a scanning spot extending over a field line width, or more, does not prevent the signal from containing information beyond the Nyquist limit for the field, although such information will be attenuated.

## 6. THE CASE FOR INTERLACE AT THE DISPLAY

At the display, the fields are used to reconstruct an approximation to the original real world. The display is required to act as a post filter to the sampling operation. In practice, the filter characteristic offered by a CRT display is roughly Gaussian in the vertical direction and virtually flat in the temporal direction. A discrete display, such as a LCD, can offer a more controlled spatial characteristic but it is arguable what kind of temporal characteristic is needed, certainly not one corresponding to a rectangular hold.

Because the display therefore cannot act as a perfect

spatial low-pass filter, display up-conversion to more lines gives a better subjective result. If the signal is non-interlaced, it can be converted to interlace with twice the number of lines requiring no extra bandwidth. For example, a 625/50/1:1 signal gives a better subjective result when converted to a 1250/50/2:1 display. On the other hand, although conversion of an interlaced signal to a non-interlaced display with the same number of lines can eliminate flicker, it is better converted to an interlaced display with double the number of lines, either option requiring double the bandwidth. For example, a 625/50/2:1 signal looks subjectively better when converted to a 625/50/1:1 display but even better when converted to a 1250/50/2:1 display.

Such up-conversion can be achieved with a fixed interpolation aperture which has to balance the preservation of the baseband spectrum supported by the old sampling structure with the need to avoid aliasing in the new sampling structure. Fig. 4 shows the spectral situation for conversion to 1250/50/2:1 from 625/50/1:1 whilst Fig. 5 shows conversion from 625/50/2:1. From these it can be seen that the first situation is simpler than the second as it involves only vertical filtering, cutting at 312 c/ph, as pointed out by Clarke. No compromises are involved, provided the source filtering ensures that there is no vertical aliasing in the primary signal. Compromises are, however, involved in the second situation where stationary vertical detail can be preserved only at the expense of losing moving vertical detail at 25 Hz. This case has been extensively treated by Weston and Devereux. [5,6]

The compromises involved in interlace-to-interlace up-conversion can be mitigated by adapting the interpolation, dependent on the presence or absence of motion. Unfortunately, such a technique can lead to switching artefacts near the edges of moving objects and a better solution is to use motion vector information, describing the field-to-field displacement of objects within the picture, to improve the interpolation process. Such information must, however, be highly

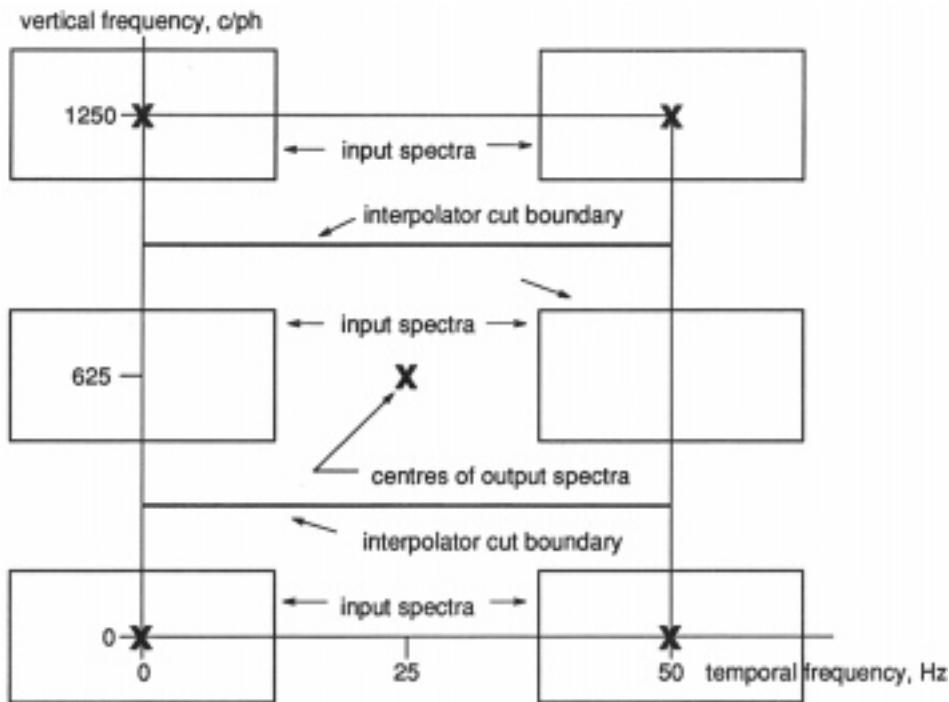


Fig. 4 - The conversion from 625/50/1:1 to 1250/50/2:1 in the spectral domain.

accurate for this application.

Some work on this has been done by BBC Research & Development as part of the RACE TRANSIT programme, mentioned earlier. Motion vectors were derived by the phase-correlation technique[7] which is capable of high accuracy and which is now available in commercial products. Some improvements to the first generation phase correlation motion estimator have been investigated. One concerns the procedure for assigning trial motion vectors measured by the

phase correlation process to individual pixels. Instead of simply choosing the trial vector which minimises the magnitude of the displaced field difference, the new method chooses a trial vector on the basis of maximum probability, taking into account the size and shape of the peaks in the correlation surface and a knowledge of the likelihood of certain types of motion. Another improvement uses gradient techniques to allow the limited sub-set of trial vectors to be expanded to provide a continuum of trial vectors. This helps in areas of complex motion, such as rotation.

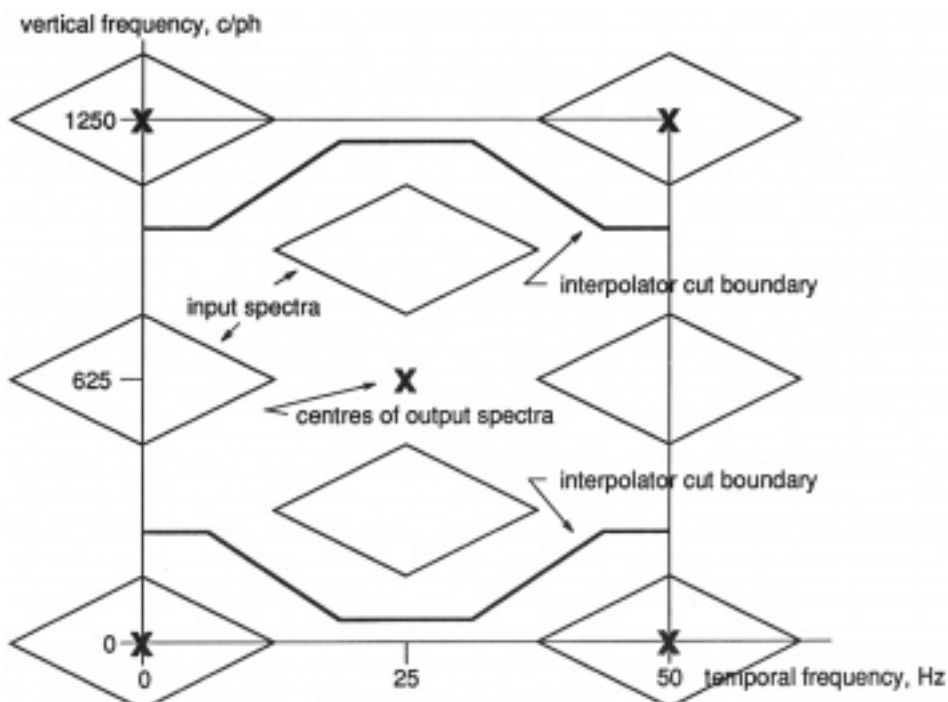


Fig. 5 - The conversion from 625/50/2:1 to 1250/50/2:1 in the spectral domain.

The motion information so derived was used to compensate the action of a three-field interpolation aperture, operating on a 625/50/2:1 signal. This signal, itself, had been derived from a 625/50/1:1 signal, obtained by scanning film, shot at a frame rate of 50 Hz, simulating the action of a shuttered camera with an integration time of 1/100 sec. The picture material consisted of a revolving bicycle wheel in front of a revolving disc with black and white sectors. As such, it represented an extremely testing source of moving material.

The 625/50/2:1 signal was converted to both 625/50/1:1 and 1250/50/2:1 signals using both fixed and motion-compensated algorithms and displayed appropriately. The results showed that the application of motion compensation gave a modest improvement over the fixed interpolation methods.

## 7. THE ROLE OF MOTION COMPENSATION

It has been shown that, for a given bandwidth, it is more efficient to use interlace scanning in the source to convey information about the real world and it is beneficial to up-convert the resulting signal to an interlace format with more lines at the display. Moreover, the derivation of motion vectors is an important part of the second process.

It will be recalled that motion compensation is also a key component of the MPEG system whose vectors will be derived from the incoming video. Now the proponents of non-interlace scanning argue that it enables us to know more securely about the motion of objects. With interlaced scanning, they argue that motion is inextricably mixed up with vertical detail. Whilst this may be true for simple schemes of motion detection, it does not hold for sophisticated methods which can take account of the fact that successive fields are scanned at different positions.

Whatever the outcome of this argument, it must be admitted that derivation of motion vectors is still at an unsatisfactory stage. How else could a MPEG coder, deriving motion information by a block matching technique, spend 80% of its channel capacity sending motion information when fed with a still picture! This is because noise in the incoming signal allows spurious block matching to take place in areas of low contrast, resulting in a decoded signal, covered in blotchy noise. The coding error may be impressively small whilst the picture is subjectively unacceptable.

Given that motion vector information is also useful for applications such as slow-motion, noise reduction, film unsteadiness correction, etc. this suggests that it should be derived as early as possible in the broadcast chain and carried along with the video signal. The

practicability of this would depend, amongst other things, on the degree of compression that could be applied to the vector signal which, in turn, would depend on its quality.

## 8. CONCLUSIONS

This paper has sought to show that the choice of scanning formats should be based on source and display psycho visual properties and not on ease of compression attained by any particular system when the development of such systems is still in a state of flux. In other words the compression 'tail' should not wag the transducer 'dog'.

## 9. REFERENCES

1. Clarke, C.K.P. Improved display techniques based on sequentially-scanned studio and transmission standards. IBC 84, IEE conference publication No. 240, pp. 210-214.
2. Hartwig, S. The influence of interlacing on source coding efficiency and motion estimation. IBC 94.
3. Chatel, J. Towards a world studio standard for high definition television. IBC 88, IEE conference publication No. 293, pp. 8-11.
4. Ninomiya, Y., Ohtsuka, Y., Izumi, Y. NHK Lab. Note No. 304, Sept. 1984.
5. Weston, M. UK patents GB 2 151 431 B & GB 2 184 628 B, Video signal processing. Eur. patent 0 266 079.
6. Devereux, V.G. Standards conversion between 1250/50 and 625/50 TV systems. IBC 92, IEE conference publication No. 358, pp. 51-55.
7. Thomas, G.A. Television motion estimation for DATV and other applications. BBC Research Dept. Report No. 1987/11.

## 10. ACKNOWLEDGEMENT

The author wishes to thank Steve Dancer of BBC Research & Development for his contribution to the work described here and the BBC for permission to publish this paper.

John Drewery  
BBC Research & Development Department,  
Kingswood Warren, Tadworth, Surrey, UK.