Research White Paper

WHP 169

September 2008

High Frame-Rate Television

M Armstrong, D Flynn, M Hammond, S Jolly, R Salmon

BRITISH BROADCASTING CORPORATION

BBC Research White Paper WHP 169

High Frame-Rate Television

M Armstrong, D Flynn, M Hammond, S Jolly, R Salmon

Abstract

The frame and field rates that have been used for television since the 1930s cause problems for motion portrayal, which are increasingly evident on the large, high-resolution television displays that are now common. In this paper we report on a programme of experimental work that successfully demonstrated the advantages of higher frame rate capture and display as a means of improving the quality of television systems of all spatial resolutions. We identify additional benefits from the use of high frame-rate capture for the production of programmes to be viewed using conventional televisions. We suggest ways to mitigate some of the production and distribution issues that high frame-rate television implies.

This document was originally published in the proceedings of the IBC2008 conference.

Additional key words: static, dynamic, compression, shuttering, temporal

White Papers are distributed freely on request.

Authorisation of the Head of Broadcast/FM Research is required for publication.

BBC Research White Paper WHP 169

High Frame-Rate Television

M Armstrong, D Flynn, M Hammond, S Jolly, R Salmon

1 Introduction

The frame rates used for film and television have been fixed for the best part of a century. A belief has arisen (eg Ferguson and Schultz (1)) that the frame rates chosen are close to an upper limit, and that little improvement can be expected from an increase. In this paper we will challenge this view, reporting on some experimental work that shows that the use of higher frame rates for capture, storage, transmission and display offers clear advantages at the resolutions associated with SD and HDTV. We will also explain why the frame rates currently in use will increasingly limit the quality of television pictures if the size of displays and/or the resolution of television systems continue to grow.

2 Historical Overview

2.1 Origin of Frame Rates

In the days of silent cinema, frame rates were not standardised, and projectionists were advised to vary the speed according to the subject matter portrayed. Operators were said to “render” a film similar to a musician rendering a piece of music (Richardson (2)). With the development of sound-on-film processes in the 1920s, film speeds and hence frame rates standardised at the now ubiquitous 24 fps. To avoid visible flicker, a double or treble-bladed shutter was used to display each image two or three times in quick succession. A downside of this technique is that moving objects being tracked by the eye appear as two or three overlapping images or appear to jump backwards and forwards along their line of motion: an effect also known as “film judder” (Roberts (3)).

The 30-line opto-mechanical television system developed by Baird and the BBC in the late 1920s and early 1930s ran at 12.5fps (Baird (4)). After broadcast trials against an improved 240-line (progressive-scan) Baird system, the interlaced Marconi-EMI television system (now known as “405-line”) was adopted by the BBC in 1937. These systems were described contemporaneously as "high-definition television". The Marconi-EMI system and all subsequent TV standards have used a field rate that is the same as the mains frequency (50Hz in Europe).

The reasons given contemporaneously (BBC (5)) for synchronising the frame rate of television to the mains frequency were to avoid "beating" against the 100Hz brightness fluctuation in AC-driven studio lights and the 50Hz fluctuation induced by poor ripple-suppression in the HT generation circuitry of early CRT televisions (Engstrom (6)). The 60Hz mains frequency used in the USA similarly led to a 60Hz field rate in their television systems (Kell et al (7)). In addition, these rates are slightly above the 40Hz minimum that was found necessary to avoid visible flicker in the displayed image on contemporary television screens (6).

At that time, it was considered sufficient (Zworykin and Morton (8)) for the frame rate to be high enough merely to exceed the threshold for “apparent motion” – the boundary above which a sequence of recorded images appear to the eye as containing moving objects rather than being a succession of still photographs. Priority was not given to the elimination of motion artefacts such as smearing and jerkiness. Contemporary tube cameras suffered from image retention, which may have limited the benefits of a higher rate anyway.

A final benefit of choosing a field rate equal to the mains frequency is simple interoperability with cinematic film recording. In 50Hz countries, since the speed difference between 24fps and 25fps is generally imperceptible, a frame of film can be represented as two successive fields of video. In 60Hz countries alternate frames of film have to be represented as three successive fields (a frame and a half) of video, a process known as “3:2 pull-down” which introduces further judder artefacts.

In summary, it appears that the field rates originally determined for television (and kept ever since) were chosen to meet the following criteria:

- Greater than the perceptual threshold for apparent motion.
- Higher than the threshold frequency at which flicker was imperceptible on contemporary televisions.
- Simpler conversion to and from cinematic film.

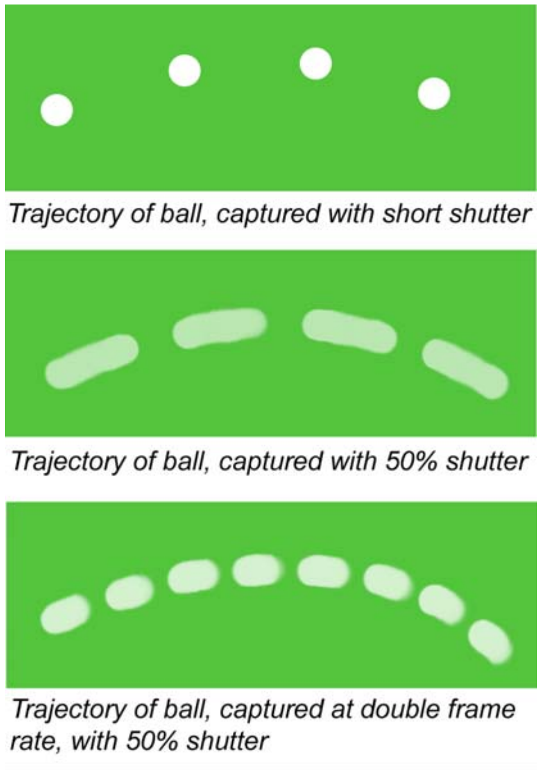

Figure 1 – Effects of frame rate and shuttering upon motion portrayal.

2.2 Early work on HDTV frame rates

With research into HDTV commencing in the 1970s, the question of the appropriate frame rate for the new television standard was open for re-evaluation. The Japanese broadcaster NHK was the leader in this field, and the 1982 summary of their HDTV research to date by Fujio et al (9) identifies “frame frequency” as a parameter to be determined. There appears to be no published research from them on the subject, however, and the field rate of NHK’s 1125-line interlaced HDTV standard remained essentially unchanged from the NTSC standard it replaced, at 60 fields per second.

NHK was not the only organisation researching HDTV at that time, however. The question of frame rate, amongst other parameters, was investigated by the BBC’s Research Department. Stone (10) performed a number of experiments with a tube camera and a CRT monitor, both modified to support non-standard field rates and other parameters set by the vertical deflection waveform.

The issue of increased flicker perceptibility on increasingly large and bright television sets was well known by the 1980s, and taking a leaf out of cinema’s book, the use of higher refresh rates was being considered to compensate (Lord et al (11)). Stone recognised that increasing the frame rate of television would not only reduce the visibility of flicker, but that it would also improve the portrayal of moving objects. He carried out subjective tests and found that for fast-moving subject material (corresponding to a camera pan at a speed of one picture-width per second), increasing the frame rate to 80Hz resulted in a subjective quality improvement of two points on the CCIR 5-point quality scale (10). The camera and monitor used were only capable of 625-line operation, but the viewing conditions were set up such that the 625-line image simulated part of an 1125-line HDTV picture.

Despite this finding, the eventual standardised HDTV formats retained the 50Hz (and 60Hz, for countries with that mains frequency) frame/field rate that was previously standardised for the original broadcast television formats. In a 1988 article, Childs (12) (then also working for Research Department) attributes this simply to the increases in transmission bandwidth and storage capacity required by the higher rate, over and above those needed for the increase in spatial resolution implied by HD.

As CRT televisions grew larger and brighter, manufacturers started using frame-doubling techniques to reduce flicker. However, the simple techniques initially employed made the portrayal of moving objects worse, by introducing a 50/60Hz “film-judder” effect (Philips (13)).

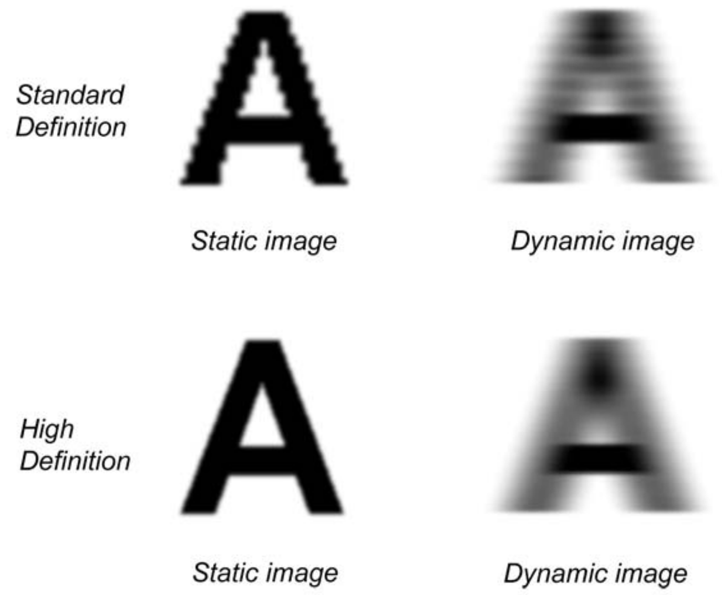

Figure 2 – Illustration of static & dynamic resolution for SD and HD images. The clear difference in static resolution is eliminated by the movement in the dynamic image.

3 Issues with conventional frame rates

Current television field and frame rates cause problems for motion portrayal. Objects stationary within the video frame are sharp, provided they are in focus, but objects that move with respect to the frame smear due to the integration time of the camera’s sensor. Shuttering the camera to shorten the integration time reduces the smearing, but the motion breaks up into a succession of still images, causing jerkiness. The perceptual difference between moving and stationary subjects is increased with the increasingly sharper images due to new television systems with successively higher spatial resolutions, so long as the temporal resolution remains unchanged. We describe the ability of a television system to represent the spatial detail of moving objects as its “dynamic resolution”. The problems of insufficient dynamic resolution – smearing, jerkiness or a combination of the two – are more noticeable with larger displays where the eye tends to follow the motion across the scene.

The problem is illustrated in Fig. 1, in terms of the movement of a ball across a plain background. In the top illustration, the trajectory of the ball is shown as if captured by a video camera with a very short shutter. Each frame would show the ball “frozen in time”, and the motion would appear jerky when the video sequence was replayed. In the middle illustration, the effect of a (half-) open shutter is depicted. The camera integration smears the motion of the ball out over the background, removing any spatial detail and making it partially transparent. These effects would be clearly visible in the final video sequence. The bottom image shows the effect of doubling the frame rate: both the smearing and jerkiness are reduced. A substantial further increase in frame rate would still be required in this example to eliminate their effects, however.

In cinema, which evolved a high resolution-to-frame-rate ratio much earlier than television, production techniques have evolved in parallel to deal with the low dynamic resolution of the medium. Tracking shots and camera moves are commonplace, often used in conjunction with short depths of field, which help by softening backgrounds that if moving at different speeds to the tracked subject would otherwise appear to jerk and judder.

The decision to adopt interlaced video for Standard Definition television resulted in a lower spatial resolution and a higher image repetition rate (and hence a dynamic resolution better matched to the static spatial resolution) than would have been the case in a progressively-scanned system of the same frame rate and bandwidth, and so the problems of motion portrayal were considerably ameliorated.

High-Definition television (by which we mean television with a vertical resolution of 720 or 1080 lines and a field or frame rate of 50/60Hz) has increased the spatial resolution without altering the frame rates used, however. Traditional television production techniques have been constrained by this change. For example, during camera pans to follow the action at sports events, HDTV trial viewers reported nausea as the static portion of the scene changed between sharp (when stationary) and smeared (when panning). The implied constraint of reducing the pan rate is not always practical in live coverage, but in practice compromises such as camera shuttering and deliberate softening of the images can help reduce the problem. Regardless of this, simple maths shows that motion of the camera or of objects within the scene at speeds higher than three pixels per field/frame eliminates all of the additional detail gained by the use of high definition, in the direction of motion. This effect is illustrated in Fig. 2. These problems will be compounded by any future increases in the spatial resolution of television.

Just as shuttering in the camera reduces the extent of smearing, a sample-and-hold characteristic in the final display increases it in a directly comparable fashion. This smearing arises with trackable motion in the displayed video where the eye is following the object across the screen, but where within each displayed image the object remains stationary for duration of the frame or field. This characteristic is to be found in the LCD televisions that are currently taking a dominant share of the market, and is the reason why these displays have a reputation for representing fast-moving material, such as sport, poorly. Manufacturers have recently started to add hardware inside LCD televisions to perform a motion-compensated frame rate doubling, which ameliorates the problem to some extent at the cost of introducing other artefacts when the motion becomes too hard to predict.

In the light of these issues, we propose that higher frame rates be part of any future video format standard, tracking or exceeding any future increases in spatial resolution. This would help redress the imbalance between dynamic and spatial resolutions which exists in current television standards, and is a necessary precursor to further increases in spatial resolution if further undesirable constraints on production techniques are to be avoided.

4 An Investigation into the Effects of High Frame-Rates

To investigate the theoretical advantages of high frame-rate capture and display, in the summer of 2007 an intensive week of experiments was undertaken. Using a Vision Research Phantom V5.1 camera, a series of 25-second sequences were captured at a resolution of 1024x576 and a rate of 300 frames per second. This camera is capable of capturing video at up to 1,200 fps, and at resolutions of up to 1024x1024 pixels, but has only sufficient memory to capture four seconds of video at that resolution and rate. To obtain a TV-standard 16:9 aspect ratio we cropped the vertical image to 576 lines. The Bayer-pattern sensor implies a lower luminance resolution than this, similar in magnitude to the reduction in vertical resolution associated with the use of interlace in standard-definition television. A shooting frame rate of 300fps was chosen to allow for shots in excess of twenty seconds long, and to facilitate down-conversion to 25, 50 and 100 fps video. (300fps also has the advantage of simple down-conversion to 60fps.) Each 25-second sequence took around ten minutes to download from the camera.

A variety of subjects was chosen to explore the advantages of high frame-rate capture and display. These included a roulette wheel and a rotating bicycle wheel, for rotational motion; bouncing balls, table-tennis and juggling, as examples of fast-moving “sports” material, and a fast-panning camera shot with and without a tracked subject.

There are few displays that accept and display video at frame rates higher than around 60fps. CRT computer monitors can in some cases be driven at up to 200fps at reduced resolution, but with a display size much smaller than is normal for HD televisions. For the purposes of our experiments we chose a projector designed for frame-interleaved stereoscopy applications, which could be driven at 100fps at a sufficiently high resolution: the Christie Mirage S+4K. The material was sent to the display over DVI from a dedicated playout PC, reading uncompressed YUV video from a high-speed RAID array.

To create 100fps material, every three successive frames of the 300fps original were averaged to simulate an unshuttered 100fps camera. For comparison purposes, we also averaged every six successive frames to simulate an unshuttered 50fps camera, and then alternated between averaging six and dropping six successive frames to simulate a 25fps camera with film-style 50% shuttering.

Further material was computer-generated by taking a still image and simulating a sinusoidal pan across it, with camera integration to match the frame rates and shuttering choices described above. The still image chosen was the well-known “Kiel Harbour” photograph. The video sequence was rendered at a resolution of 1280x720.

Our observations were as follows. The most striking differences were seen in the panning shots – real and simulated – where the loss of spatial resolution in the detail of the background was particularly marked, particularly in the 720p Kiel Harbour simulated pan sequence. In the standard definition pan shot, lettering that was clearly legible in a static image was unreadable during the pan at frame rates below 100fps. The reduced motion blur on the tracked pan shot also gave a greater sense of realism and “three-dimensionality” as the improved dynamic sharpness of both the moving objects and the background improved the quality of the occlusion depth cue. The table-tennis sequence demonstrated that even 100fps was manifestly insufficient for coverage of this and similar sports when viewed perpendicular to the action. Motion blur was also still in evidence in the juggling sequence at 300fps, played back at 1/3 speed.

It is striking that significant improvements were discernable even at resolutions similar to standard definition television. This implies that high frame-rate capture and display is a technique that can improve the quality of television in its own right, as well as a necessary consideration as the spatial resolutions of proposed television standards continue to increase.

5 Implications of High Frame-Rate Video

Capturing video at a higher frame rate inevitably leads to higher-capacity storage and bandwidth requirements for an uncompressed signal. In practice however, the use of compression for storage and transmission at all stages of the production process is already commonplace in today’s HD workflows. A high frame-rate video signal will contain less frame-to-frame variation and temporal aliasing than a conventional signal, which will facilitate higher compression ratios for the same perceptual quality. Thus a doubling of frame rate (say) is not anticipated to lead to anything like a doubling of bandwidth for a compressed signal. Furthermore, the reduction in temporal aliasing should enable better motion recovery from the video signal and enable new video compression techniques such as the use of three-dimensional transforms.

Producing programmes at a high frame-rate also has the potential to improve the final quality even when broadcast at conventional frame rates, in the same way that HD production can result in exceptionally high-quality pictures even when the production is broadcast in a standard-definition format. The high frame-rate can be regarded as temporal oversampling in this context, which leads to the exciting possibility of allowing temporal aspects of the video such as shuttering and motion blur to be adjusted by the director in post-production, perhaps even selectively within the frame. This would add additional “looks” to the director’s palette, complementing the well-known “video” and “film” looks.

The use of high frame-rates may also permit the automated removal of the effects of flash photography from the recorded scene, along with improved noise reduction. Production at frame rates such as 300fps would allow simple high-quality down-conversion to the frame rates used in both PAL and NTSC countries.

While the early constraints on frame rates imposed by simple CRT displays and tube cameras are no longer current, the problems associated with the brightness fluctuations of mains-powered lighting remain. This is anticipated to be a particular issue for fluorescent tubes, which contain several phosphors of different colours which generally have different decay times. New techniques and technologies will need to be devised to deal with these issues, which could include the increased use of DC lighting, 10 kHz fluorescent tubes etc, and perhaps the automatic detection and correction in the camera for the changes in lighting amplitude and colour.

6 Conclusion

The spatial resolution of broadcast television cameras and displays has reached the point where the temporal resolution afforded by current frame rates has become a significant limitation, particularly for fast moving genres such as sport. BBC Research has successfully demonstrated that increasing the frame rate can significantly improve the portrayal of motion even at standard definition. If the spatial resolution of television standards continues to increase, raising the frame rate to maintain the balance between static and dynamic resolution will only become more important. Even at the spatial resolutions of SD and HDTV, the motion artefacts associated with 50/60Hz screen refresh rates will become increasingly apparent as television display sizes continue to grow.

Even for television pictures transmitted and displayed at conventional frame rates, capturing at high frame-rates can offer some improvement to picture quality through temporal oversampling, giving better control over temporal aliasing artefacts and offering a choice of “looks” to the director at the post-production stage. It also offers improved compatibility with the different conventional frame rates adopted internationally.

We assert that a higher capture and display frame rate leads to a step change in picture quality regardless of the spatial resolution.

7 References

1. Ferguson, K. and Schultz, W, 2008. Predicting Subjective Video Quality. Broadcast Engineering World. February 2008.
2. Richardson, F. R., 1911. Projection Department. The Moving Picture World. December 1911.
3. Roberts, A., 2002. The Film Look: It’s Not Just Jerky Motion… BBC R&D White Paper, WHP 053. December 2002. p.7.
4. Baird, J. L., 1933. Television in 1932. BBC Year Book, 1933.
5. BBC, 1939. Technical Manual, Marconi-EMI System of Television, London Television Station. April 1939.
6. Engstrom, E. W., 1935. A Study of Television Image Characteristics: Part Two, Determination of Frame Frequency for Television in Terms of Flicker Characteristics. Proc IRE, 23 (4). April 1935. pp. 295 to 309.
7. Kell, R. D. et al, 1936. Scanning Sequence and Repetition rate of Television Images. Proc IRE, 24 (4). April 1936. pp. 559 to 576.
8. Zworykin, V. K. and Morton, G. A., 1940. Television: The Electronics of Image Transmission. Wiley. New York.
9. Fujio, T. et al, June 1982. High Definition Television. NHK Technical Monograph, 32.
10. Stone, M. A., 1986. A variable-standards camera and its use to investigate the parameters of an HDTV studio standard. BBC Research Department Report 1986/14.
11. Lord, A. V. et al, 1980. Television Display System. UK Patent GB2050109, 8 May 1980.
12. Childs, I., 1988. HDTV: putting you in the picture. IEE Review, July/August, 1988. p. 261.
13. http://www.research.philips.com/newscenter/dossier/naturalmotion/judder-free.html. Accessed in May 2008.
14. Wade, J. G, 1987. Signal Coding and Processing: an introduction based on video systems. Ellis Horward, Chichester.

Appendix A – Progress Update

We include here a comparison of still images from a follow-up shoot to produce material for demonstration and experimental purposes at 1920x1080p300. This material was first shown at IBC 2008, downsampled to 1400x788 and temporally downconverted to frame rates of up to 100fps as described in section 4. This work post-dates that published in the IBC2008 proceedings.

We present here two still images from that material. The top image shows a crop from a 1920x1080 still frame, taken from a sequence downconverted from 300fps to 50fps. The bottom image shows a still from the same point in time in the 300fps original. The reduction in motion blur is clear, illustrating the increase in picture detail in objects moving relative to the camera that results from the shorter frame duration. If presented at 300fps, this increase in dynamic resolution comes without the disadvantages of higher noise and stuttering motion associated with shortening the frame duration by shuttering a lower frame-rate camera.