B B C

R&D White Paper

WHP 053

December 2002

The Film Look: It's Not Just Jerky Motion…

A. Roberts

The Film Look: It's Not Just Jerky Motion…

Alan Roberts

Abstract

There have been many attempts made to generate television pictures that have "The Film Look" while being generated electronically. Most have concentrated on mimicking the temporal appearance of film, while others have concentrated on contrast handling. Some have done both, and more. The issues of sharpness and depth of field are rarely tackled. This document investigates all of these major aspects of the film look and analyses the properties of film and electronic image generation to find out exactly how they are different and why, and how that difference can be minimised. The "BBC setup" conditions now adopted in some television cameras exploit this work, notably in PSC camcorders used for HDTV image capture. Although aimed primarily at HDTV cameras and their use in attempting to mimic and thereby encourage their replacement of film usage, the principles explored here could apply equally to all forms of video cameras provided that their control ranges are sufficiently flexible.

Key words: camera, colorimetry, gamma, knee, aperture correction, detail enhancement, depth of field, shuttering, film look.

The Film Look: It's Not Just Jerky Motion…

Alan Roberts

If you ask a programme-maker for an opinion on "The Film Look", you will probably hear comments on contrast range, motion judder, sharpness and depth of field. You may also hear about differential focus and grain. Occasionally someone will comment that film "doesn't do red well", or that video cameras always have a "video look". Some of these comments have basis in colour science, others are personal interpretations of aspects of the colour science but without using the "correct" terms. A previous White Paper[1] dealt with the colour science behind the various options for preferred camera setup to mimic film. Since writing that document, there have been developments in the understanding of camera operations which have lead to better camera design, and better use of existing designs. This document builds upon the contents of the earlier paper in the light of later experimental work and camera development.

The major differences between film and video divide neatly into two areas of interest, the amplitude transfer characteristic which deals with contrast range, and the modulation transfer function which deals with everything else. These will be now dealt with separately.

1 Amplitude Transfer Characteristic (or Gamma γ)

There is a commonly held belief that film can handle a contrast range of at least 10 stops (1024:1) whilst video can do maybe only as little as 5 stops (32:1). Neither of these statements is particularly true, and the comparison is grossly misleading. In practice, there is little to choose between film and video, as the following analysis will show.

A reference display is assumed, obeying a power law (*Gamma*) of 2.35 such that:

$$Light \propto Voltage^{2.35}$$

The value of 2.35 is typical for cathode ray tube displays. In practice, all displays produce some light, even when the drive signal is at zero, and so the displayed contrast will not be infinite. Also, the effect of stray light will act as a base level in an uncontrolled way, so the transfer characteristic is much more complex:

$$Light = Stray + k(Voltage - Offset)^{2.35}$$

where *Offset* is any error in the setting of black level and *Stray* is the combined effect of stray light and the residual light produced when the drive signals actually are all zero. The effect of including values for *Stray* and *Offset* is to produce a curvature in the transfer characteristic even when plotted as a log/log curve (where it should be straight, with a slope of 2.35). In practice, the value of *Offset* should be less than ±2% (of voltage, i.e. less than ±0.01% of light output, implying 10,000:1 display contrast rage) if a PLUGE signal is used during display line-up. For the purposes of this investigation, the values for *Stray* and *Offset* are assumed to be zero, so the display can have infinite contrast since there will be zero light output when the drive signals are zero.

1.1 Film Transfer Characteristic

There is no shortage of published data on film performance, and it is particularly easy to analyse since a negative film obeys the following law over a wide range:

$$LOG(Density) \propto LOG(Exposure)$$

It is only the extent and slope of this range, and what happens outside it, that characterises film. The data shown here (Figure 1) is taken from Hunt[2] who uses it to illustrate the performance of negative film. There is a central exposure range of approximately two decades over which the curves are linear, with a slope of about 0.9, meaning that within that range:

$$Density \propto Exposure^{0.9}$$

Outside that range, the curves compress the extremes, covering about 4 decades in all. From this, we can conclude that film linearly copes with 100:1 contrast (6.5 stops), and can deal non-linearly with another 5:1 (2.2 stops) at either end of the contrast range, a total of around 2,000:1 (11 stops). At the extremes of the range, contrast is significantly crushed, to an extent which makes detail extraction difficult if not impossible. The typical appearance of these curves is often described as "lazy S", for obvious reasons. The significant difference in densities for the three layers is due to the construction of the film. The blue layer (sensitive to blue light but linked to a yellow dye in the processing and therefore shown as yellow in Figure 1) is on top, the red layer is at the bottom, and is seen through the blue and green layers.
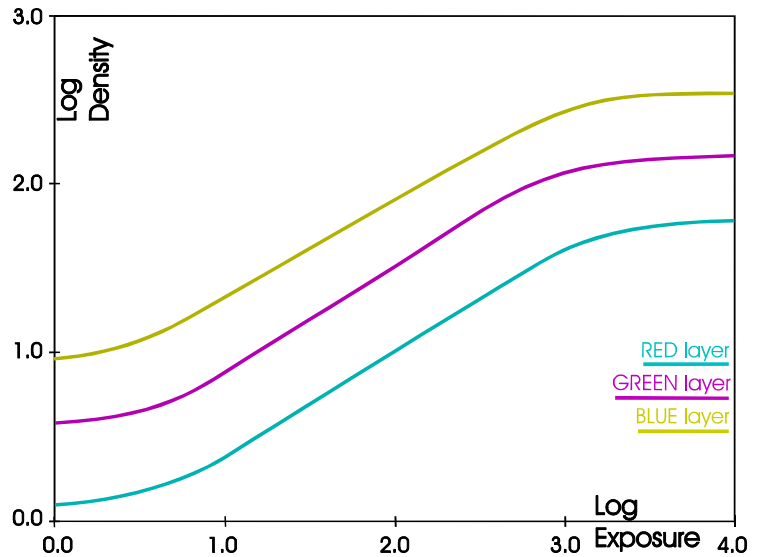

Figure 1 - film negative charcteristic

However, this is only the performance of the camera negative, to see exactly how the film performs on television, we have to perform all the normal production operations. The signals from the telecine channel are proportional to transmittance and must be colour-balanced at some value such that R=G=B, and gamma-corrected, in this instance to perfectly compensate for the reference cathode ray tube display:

$$Light \propto \left(10^{-Density/2.35} \right)^{2.35}$$

The negative sign inverts the signal, to obtain a positive image from the film negative. The values of 2.35 neatly cancel. In practice, the Colourist would modify the gamma curve to allow for the negative's nominal value of 0.9, and would differentially adjust the RGB curves to taste. None of that is allowed for in this analysis since it is part of the artistic adjustments that are made. Plotting these results logarithmically (Figure 2) again reveals the lazy S curve shape that is typical of film performance. The white balance point has been arbitrarily fixed in the early part of the onset of white crushing, leaving approximately three decades (10 stops) of near linear characteristic, and this is the *exposure contrast* range (exposure white/black).
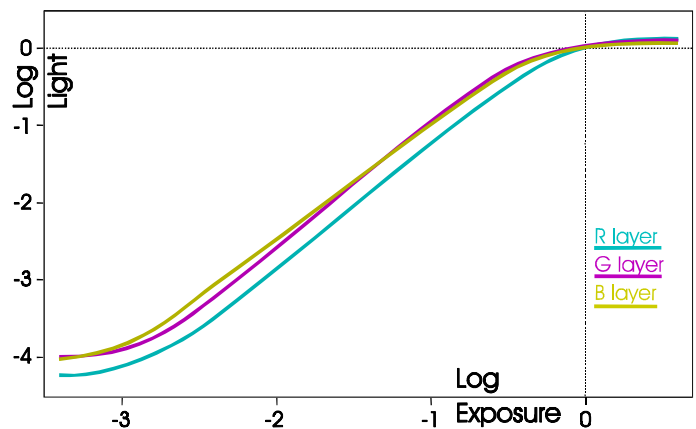

Figure 2 - film negative characteristic

2

The *display contrast* range (displayed white/black) is also about 4 decades, about 10,000:1, assuming a perfect display.

Contrast is the ratio of maximum to minimum value and reduces to a fraction:

$$contrast = value_{max} \Big/ value_{min}$$

but it is more meaningful to express it as a ratio to unity:

$$contrast:1 = \left(value_{max} / value_{min} \right):1$$

In the television industry, is usually quoted as *contrast*:1 (e.g. 32:1, implying that the darkest resolvable signal level is 1/32nd of the maximum). The film industry prefers logarithmic scales and refers to contrast in stops, where one stop is a power of 2, thus 5 stops means 2^5 or 32:1. When the numbers get really big, it is easier to deal in decades, thus 2 decades means 10^2 or 100:1

1.2   Video Transfer Characteristic

It is hard to understand exactly how the commonly held belief came about, that video cameras have a contrast range of between 30:1 and 50:1. The available evidence does not support this belief. To explain this situation, we must first explore the theoretical performance of contrast laws, and then illustrate the validity of that analysis with measurements from real cameras.

1.2.1 Standard Video Transfer Characteristics

There are two types of curve in common use (Figure 3) that offset the input and those that offset the output. Both types have an exponent of around 0.45 and are limited to gains of about 4.5 near black. If a pure power law were used, it would require the gain of the amplifiers to be infinite when the signal level is zero, clearly this is impossible so an offset is imposed such that a line with finite slope joins the curve tangentially at a reasonably low level, typically between 1.5% and 3%. In order to join the linear part to the curve tangentially, the curves can be offset on either axis. The 18% Exposure point is interesting because it represents the normal exposure level for film, using an 18% reflectance grey card.
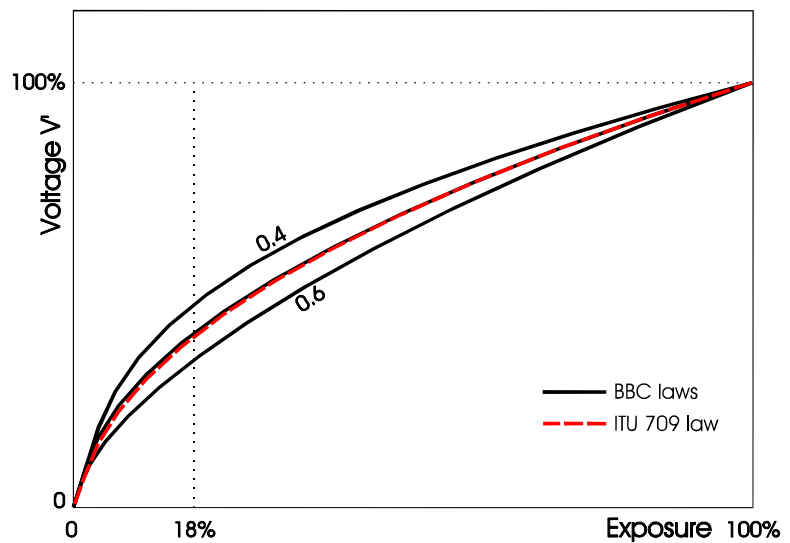

Figure 3 - standard video gamma curves

The BBC standard[3] has three curves, each with maximum gain at black of 5, but with varying power laws and offsets, the central curve follows this equation, offset on the *Exposure* axis:

$$\text{if } Exposure \le 0.020202 \text{ then } V' = 5Exposure$$

$$\text{else } V' = \left(\frac{Exposure - 0.01011}{1 - 0.01011} \right)^{0.5}$$

3

while the ITU HDTV law[4] has an offset on the V' axis:

$$\text{if } Exposure \leq 0.018 \text{ then } V' = 4.5 Exposure$$
$$\text{else } V' = (1.099 Exposure)^{0.45} - 0.099$$

Evidently, the ITU curve is very similar to the BBC 0.5 curve, but this form of plot is not particularly revealing and does not help in any quest to discover the achievable contrast range. To accomplish that, we must apply the display gamma of 2.35 and replot the curves on logarithmic axes (Figure 4).

The lowermost points on each curve are chosen to represent 1 quantum level in an 8-bit system above black. So the *electrical contrast* is 216:1 since black level is 16 and white is 235. But that is not a valid measure of *display contrast*, for that we must use the same criterion as for film, i.e. white/black, which in this case is about 5.3 decades (about 200,000:1), assuming the same perfect display. The *exposure contrast* range (exposure white/black) is about 3 decades (1,000:1).
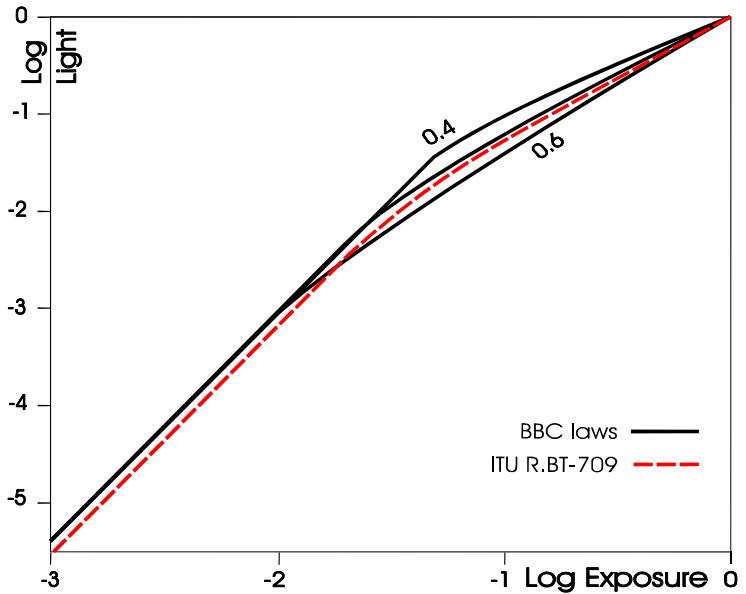

Figure 4 - standard video gamma laws

All these curves have straight parts at the bottom, where the gain is constant (4.5 or 5), the ITU curve lies under the BBC curves because it has lower gain. The break points lie near the middle of the curves, leaving about 1.5 decades (about 30:1) on the "gamma law" part of the curve, which is curved in this plot because of the axis offset in the equations. This is probably the cause for the belief that video has only a 30:1 contrast range. Since the part above the straight line is curved in the log/log plot, colour reproduction cannot be accurate, as it can be for film, but the artistic input of the Colourist can still be used to extract the appropriate appearance for most programme-making genres. Thus it seems permissible to claim that the standard video gamma-correction curves can have an *exposure contrast* range of about 10 stops (1,000:1), albeit non-linearly. This is limited only by camera noise, requiring the signal-to-noise ratio to be 60dB or better.

So far, the extra functions of Black Stretch and Knee have been neglected. Clearly, if these are allowed, then the *exposure contrast* range can be extended. To see how these affect contrast we must use measurements from real cameras.
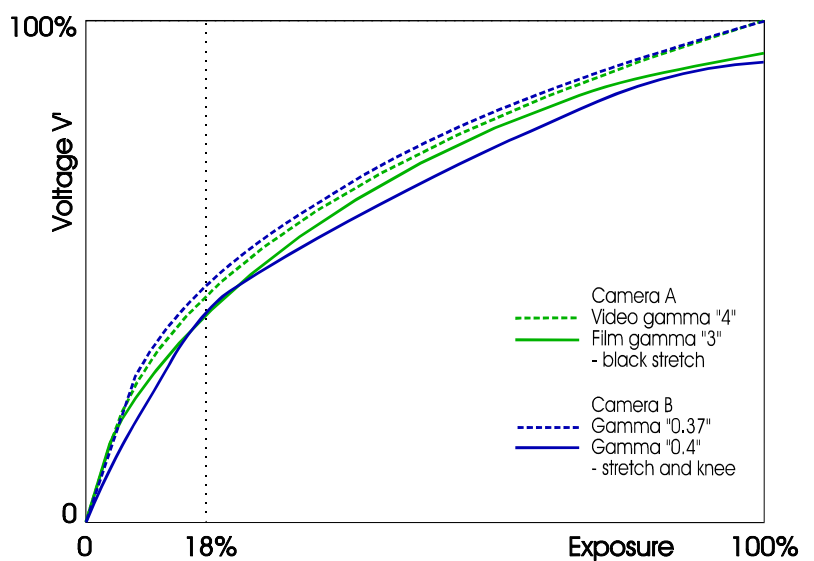

Figure 5 - camera gamma curves

1.2.2 Camera Measurements

For this document, two High Definition cameras (similar models from two manufacturers) were set up in an attempt to make them function as film-look replacements for 35mm film-shooting for television. Two settings are shown for each camera, one for video, one for film (Figure 5), using values from the standard tables in the camera menus. The curves are tracings from waveform-monitor photographs of a saw-tooth test signal. The two video curves (Camera A video 4, and Camera B 0.37) are close approximations to the BBC 0.4 law and are acceptable for normal video shooting. Camera A offered a set of gamma curves designated as "Film" type as an alternative to conventional video versions, which evidently reduces the gain near white to increase the exposure range.

Since the manufacturer's published data sheets claimed that the cameras' ccds used only about 1/5th or 1/6th of the dynamic range for normal output, setup conditions were sought to exploit the extra range, using the knee function. In Camera A, combining optimal knee settings with the camera's own film-style gamma-correction produces a curve that has no obvious break points (knees), yet can capture up to 600% of the nominal peak white level. Camera B has more obvious break points, both near white (~85%) and where the gamma curve becomes linear (~5%), but still can capture about 400% of nominal peak white. Adding a Black Stretch increases the capture range in both cameras.

The results are plotted log/log (Figure 6) in order to estimate the *exposure contrast* range. Clearly, this is now about 4 decades for both cameras when using their preferred film settings (A – Film 4, B - 0.4) with knee and black stretch applied. The performance is markedly different at very low levels (below 1%), and is limited by camera noise; at standard gain this was

Figure 6 - camera gamma curves

measured at about 54dB below white for both cameras, limiting the noise-free *exposure contrast* to about 500:1. However, since grain is often deemed to be a part of the "film-look", it may be appropriate to claim that this noise is part of the "look", and continue to claim 10,000:1.
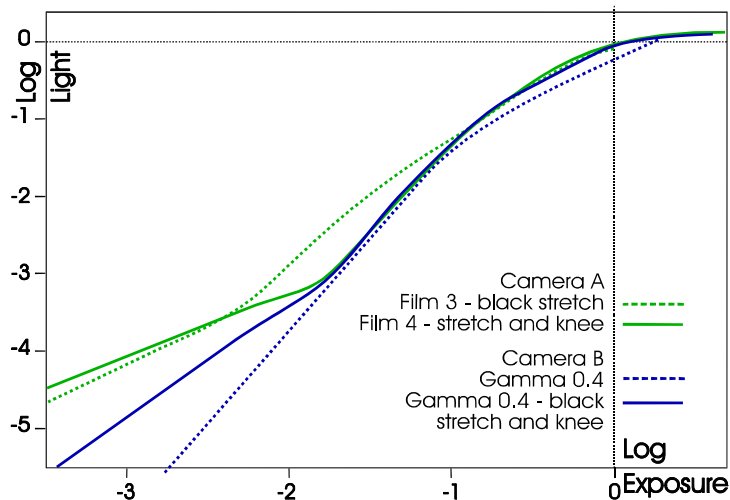
1.3   Comparisons: video versus film

It is not particularly revealing to plot these curves as Voltage versus Exposure, the overall performance of the system is far more interesting.

The linear plot of standard curves versus film, Exposure to Light (Figure 7), shows that the BBC 0.4 law is linear over the bulk of its range, but has a curl near zero. That curl is typical of all gamma-correction curves that do not have infinite gain near zero, and so is a fundamental feature of all video systems. The visual effect of this is most pronounced in colours that are already near full saturation; saturation is enhanced and hues are shifted towards the nearest primary or secondary colour.
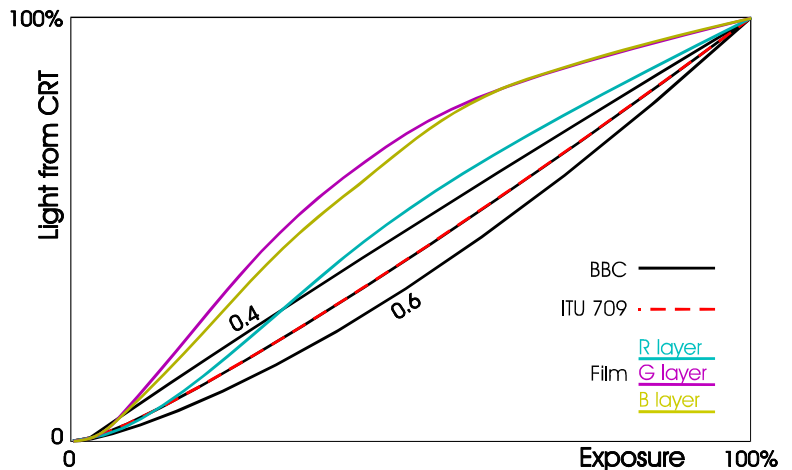

Figure 7 - video vs. film

In extreme cases this produces cartoon-like colouring, W.N.Sproson[5] gives a good analysis of this effect.

The BBC 0.5 and 0.6 laws, and the ITU law, all curve below the linear slope of the BBC 0.4 law; this has a similar effect to the low-end curl, but it applies over much more of the contrast range, and to colours with mid-to-low saturation. This effect, probably above all others, defines the "Video Look".

The film laws (plotted up to the assumed white balance point) are all much more curved, because the power law of 0.9 has not been compensated in this analysis. But significantly, they also curl near zero in the same way as the video gamma curves do, so the same hue shift and saturation changes will happen. Also, the curves for Red Green and Blue are significantly different from each other, therefore colour rendering will be highly inaccurate unless corrective measures are taken. For these reasons, film, when scanned for television, must always be colour corrected. By comparison, video signals are most frequently not corrected at all unless there are gross errors, or the Colourist wishes to make artistic alterations.

The logarithmic plot (Figure 8), shows the performance of the cameras versus film. Clearly, both the *exposure contrast* (4 decades, 10,000:1, 13 stops) and the *display contrast* are actually slightly greater for the High Definition cameras than for film, and Camera B actually achieves more *display contrast* than either Camera A or the film sample. However, camera noise can be expected to limit the achievable *exposure contrast* to about 11 stops, 2,000:1, very similar performance to that of film.

Both the High Definition cameras have much more curved characteristics than does film. For this reason, it is sensible to expect that the video camera signals would need to be adjusted by a Colourist, and that similar performance to that of film should be achievable.

Figure 8 - film vs. video

It is difficult to demonstrate the achieved contrast range in this document, but it is fairly easy to illustrate the way in which highlights are handled by using the knee function. Figure 9 shows the "before and after" performance for Camera A when exposed to BBC Test Card 57, a grey scale. The knee function was out of circuit for the normal exposure. In the over-exposed picture, the brightest patches are still just separable, even though they have been heavily compressed to cope with 12dB of extra gain, equivalent to 2 stops of over-exposure.
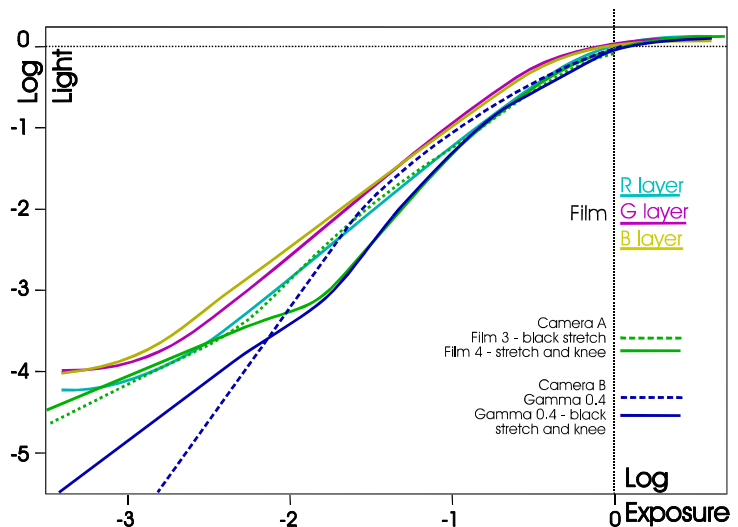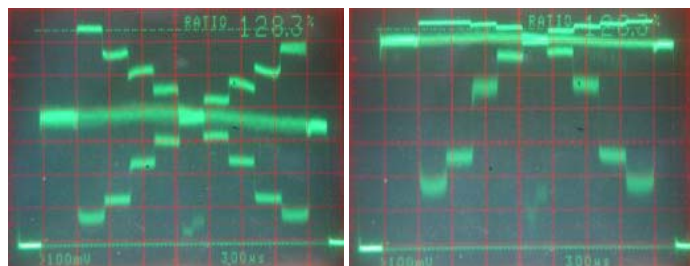

Figure 9 - Test Card 57
Normal exposure (left), 2 stops over-exposed (right)

2   *Modulation Transfer Function (MTF)*

The Modulation Transfer Factor is a measure of a system's ability to reproduce a sinusoidal modulation at a specified frequency, it is the gain at that spatial frequency. The Modulation Transfer Function is the distribution curve of that factor over a range of spatial frequencies. Confusingly, both are known by the same abbreviation, *MTF*. In this document, *MTF* will always describe the function (curve), and *mtf* will refer to the

factor (single value). In video parlance, gain is a near equivalent of *mtf* and horizontal frequency response is a near equivalent to *MTF*, but *MTF* should relate to physical sizes or measurements rather than temporal frequencies (Hz) that involve time, and applies circularly rather than vertically or horizontally. Thus, reference will be made to cycles/mm and cycles/picture height or width, rather than to Hz, and therefore is independent of the scanning standard.

MTF is always measured using linear signals, it relates to optical performance rather than electrical, although the curves can be displayed in the usual variety of confusing ways. Also, it should always be measured using sinusoidal frequency gratings, this avoids the confusing effects of harmonics that arise when using square-wave test signals, and of possible alias products from sampling.

For this investigation, BBC/ITVA Test Card 60 was used, a transparency containing patches of sinusoidal transmittance, calibrated in MHz at frequencies which relate directly to analogue television. Simple scaling was used in order to test cameras at the higher frequencies used in High Definition. The conversion from MHz at SDTV into cycles per picture width is quite simple:

$$Cycles/picturewidth = 52 Frequency$$

since the active duration of the SDTV line is 52µseconds. So, if the test chart is framed to occupy 50% of the image width, the patch marked at, say, 2MHz produces approximately 200 cycles per picture width (actually 208), and so on.

2.1 Why MTF matters

MTF is responsible for the detail in the picture. It is the means of carrying detail from the scene to the display. It provides all the visual clues for placement of object components of the image, because it is the texture and edges of objects that provide the clues for image location. It therefore also provides all the clues for the portrayal of change of object location, i.e. movement. Since movement is responsible for "film judder", that is also controlled by the *MTF*. It also contributes to Depth of Field, thus, if any video system is to mimic film performance, it must mimic the *MTF* reasonably well.

2.1.1 Film Judder

Judder is manifest in two ways, both resulting from whole-image motion or object motion within the image. If the motion rate is low, then the viewer has only a feeling that the picture is "restless", but with higher motion speeds, the image can break up into two separate images, moving together. Both these effects are generated in the eye, rather than on-screen, and are caused by the double-displaying of images. Film shot at, say, 25 frames per second, delivers 25 complete pictures per second to the viewer. If each picture is shown only once, then all motion will be seen as smooth or "fluid", but the display will flicker amazingly and can cause epilepsy in severe cases. To overcome this, each picture is shown twice to the viewer (for film projection this is the whole image, for television the displayed images are fields, each of only half the frame lines), to make an image stream at 50Hz, close enough to the limit of flicker perception. For any moving image content shown twice, one showing must be at the wrong time, or in the wrong place.

Figure 10 shows how film photography records and displays a motion stream. The camera exposure is for only 50% of the frame duration, the shutter must close to allow the film to move to the next frame. The smooth path through each set of displayed images (A and B, with short
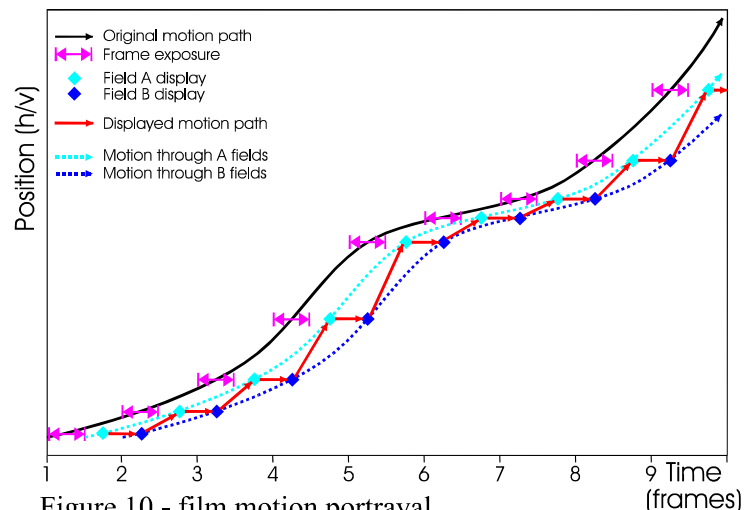

Figure 10 – film motion portrayal

7

duration display exposure) is disrupted by the repetition. What the eye actually sees is the red motion path. If the spatial displacement from frame to frame is small, then the disruption is also small and the viewer sees only a vague restlessness. But if the spatial displacement is large, then the eye separates the two showings of each frame into separate motion paths because it cannot resolve the erratic motion path (the red line) as the motion of a single object. Single objects are seen as adjacent pairs of identical objects, separated by the distance through which they have moved between adjacent exposure frames. This is judder.

Judder is inevitable in any display system that shows source images more than once, but the perception of it depends on the sharpness of the images. If the edges of the object are blurred, then the eye will be less able to identify its location and so judder will be less visible. The sharpness of object edges is controlled by the system *MTF*, and that's why *MTF* is important. Exposure duration also affects judder because shorter exposure increases the sharpness of moving objects and makes them easier to locate.

Experienced film-makers deal with this problem in several ways:

- Avoid rapid camera movement. If the image doesn't move, it can't judder.
- Avoid rapid object movement. Pan the camera to follow the action; it won't judder but the background will. The viewer is supposed to be concentrating on the action rather than the background anyway, so this works.
- Use short Depth of Field to soften the moving parts of the image. They will judder less. Again, the viewer is not supposed to be concentrating on the background anyway, so this is no loss.

2.1.2 *Depth of Field (DoF)*

It may seem odd to include *DoF* in a discussion of motion judder, but it plays a significant part in the subjective appearance of film-shot programming. One of the most often heard complaints about video, made by film people, is that the *DoF* is too big. It would be helpful to investigate this a little further. *DoF* is the range of object distances between which everything is apparently in focus. The term *disc of confusion* is also of interest because it defines the resolution limit of the system. It is convenient to express it as the angle, θ (in radians) subtended at the lens by the smallest object that is distinguishable from a point. Thus, any object subtending a smaller angle will be seen in the image as being in focus i.e. a point.

DoF is defined in terms of the *hyperfocal distance*, *h* (the distance between which and infinity everything appears to be in focus if the lens is focussed at infinity) and the object distance *u*. Now, $h = a/\theta$ where *a* is the lens diameter. Using that definition, the equation for *DoF* is:

$$DoF = hu\left(\frac{1}{h-u} - \frac{1}{h+u}\right)$$

The focal length of the lens is not directly involved, only the aperture, the disc of confusion and the distance. In this analysis, the aperture is a physical dimension, not the *F* stop, but we habitually refer to apertures by *F* numbers since they directly relate to exposure, so another definition would be helpful, $F = a/f$ where *f* is the focal length of the lens. At this point, the image size also becomes important since it linearly affects the aperture, in order to keep *F* constant as image size changes, *a* must scale linearly with image size.

To put all this in context, if *F* is kept constant, then the *DoF* doubles if the image size halves. Table 1 gives the image dimensions for some relevant cameras and systems.

Table 1		Width			Height		
		mm	pixels	Cycles/mm	mm	pixels	Cycles/mm
SDTV	Video ccd, 2/3"	9.60	720	37.5	5.40	576	53.3
SDTV	Film Super 16mm (camera)	12.52	720	28.8	7.42	576	38.8
SDTV	Film 35mm (projected)	20.96	720	17.2	11.33	576	25.4
HDTV	Video ccd, 2/3"	9.60	1920	100.0	5.40	1080	100.0
HDTV	Film Super 16mm (camera)	12.52	1920	68.7	7.42	1080	72.8
HDTV	Film 35mm (projected)	20.96	1920	45.8	11.33	1080	47.7

Using the horizontal dimensions, it is obvious that a 2/3" video camera will have greater *DoF* than a 35mm film camera, in the ratio 20.96:9.6, about 2.2:1, for the same *F* number. To achieve the same *DoF* with the same lens angle, the 2/3" camera lens focal length must reduce by 2.2:1 and the aperture *a* must remain the same, thus the *F* number must reduce by 2.2:1 as well. Many of the best 35mm movie lenses are designed to operate optimally at their maximum aperture, and this can be as much as F1.6, so there is no chance of a 2/3" video camera achieving the same *DoF*, another way must be found.

It is technically possible to use a 35mm movie lens with a 2/3" camera (one manufacturer of lenses routinely does this), but the angle of view and *DoF* do not mimic 35mm. One solution is to mount the 35mm movie lens such that it creates its image on a ground-glass screen outside the camera, and to use relay optics to carry that image to the ccd sensors. Then, the lens would deliver the same angle of view and *DoF* as it would on a 35mm movie camera. Such a product exists in prototype at the time of writing this document. While it may solve the obvious problems, it also introduces new ones, in that some of the advantages of shooting on video rather than film are lost, e.g. the lower cost of lenses and the smaller, more wieldy cameras.

2.2   Film MTF

Again, borrowing generic data from Hunt rather than using data for a specific film stock, Figure 11 shows the *MTF* of a typical 35mm negative film. In this document we are not concerned with the performance of specific film stocks, but more in the general properties of them. As usual, when dealing with film attributes, the curves are plotted on logarithmic scales. Clearly, the red layer has poor response relative to blue and green. This is because of the physical structure of the film, rather than any chemical properties. Typical film stock has the blue-sensitive layer on top (shown yellow in the figures), and the red-sensitive layer on the bottom. Light destined for the red layer must pass through all the upper layers first, not just the two colour sensitive layers, but the colour filters and other blocking layers as well. Optical diffusion happens in these upper layers, resulting in this difference in performance. Also of note, is that the responses all increase above 1.0 before falling, this is due to adjacency effects within the layer. Again, Hunt provides excellent descriptions of this effect and its causes and so no further explanation will be given here.

The logarithmic plots favoured by film-folk can conceal a significant amount of information. Figure 12 shows the same film *MTF* data re-plotted on linear scales.

The horizontal scale is calibrated in cycles/picture width and is independent of sensor and display sizes. The secondary scale, across the top of the graph, is a calibration in cycles/mm, assuming that
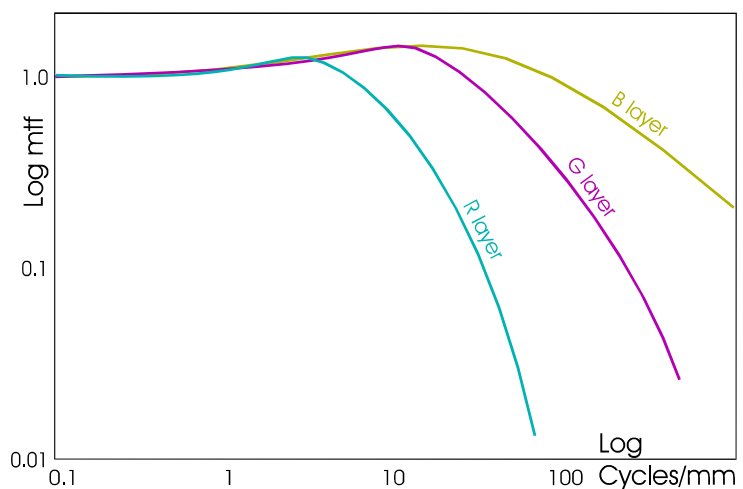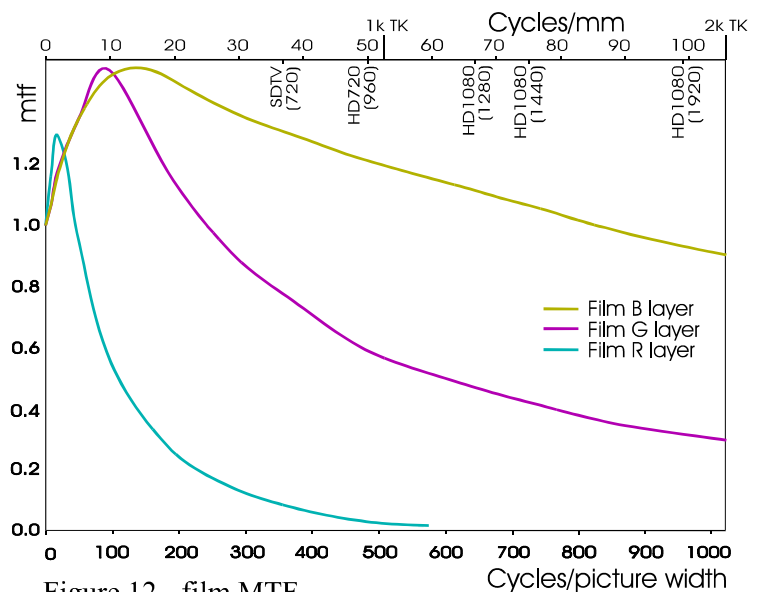

Figure 11 - film MTF

Figure 12 - film MTF

9

the image is formed on a conventional 2/3" 16:9 sensor. There are also indicators to show the resolution limits of several imaging standards, from Standard Definition television (SDTV, 720-pixel) to high-resolution film scanning (2048-pixel, known as 2k). Table 1 shows the image dimensions and resolution limits of some relevant systems for 16:9 widescreen images, and the values for this scaling are taken from that table.

In the linear plot, the performance differences between the film layers is much more evident, and it explains the often-heard comment about film - "it doesn't do red well". Clearly, the diffusion in the upper film layers attenuates much of the detail in red, and it is probably this effect that is liked about film's reproduction of skin tones, it softens faces. Even in SDTV, where the resolution limit is about 50 cycles/mm, the red layer response has fallen to about 8% by this value. At HDTV resolutions, the response is very low.

For any video camera to mimic this performance, it should reproduce the *MTF* of film as well as the transfer characteristic.

2.3   Video MTF

Since camera and image motion are important parts of the film look, the static and dynamic *MTF* of the camera must be considered. Since the dynamic *MTF* is a modification of the static *MTF*, static considerations must come first.

2.3.1 Static MTF

The horizontal frequency responses of video systems are well defined and documented. A good example is the template defined in ITU R.BT-709[4] for HDTV systems (Figure 13). The curves shown are arbitrarily fitted into the templates and should not be taken as representative of real filters.
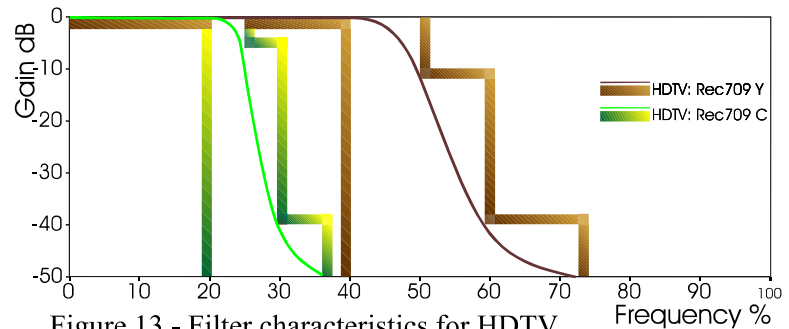

Figure 13 - Filter characteristics for HDTV

The response is essentially flat up to 30MHz, and is required to have attenuation greater than 40dB at 44.25MHz. Thus the response of the system is flat up to 40% of the sampling frequency (74.25MHz). The limiting resolution of the standard is defined by half the pixel count, 37.125MHz or 960 cycles per picture width. Most other television system specifications limit their resolutions in the same way, so the *MTF* of television systems can be regarded as flat up to 80% of the limiting resolution.

Unfortunately, the frequency response defined by these filters is that of the video system and not that of the camera. The camera's response should be measured in the linear light domain, not in the gamma-corrected signal. It is almost impossible to obtain a linear-light version of these filters, simply because the application of gamma-correction alters the gain of the camera with signal level. At low levels (below about 2%) the gain is 4.5 or 5 or maybe higher, but at 90% the gain can be as low as 0.25. Thus the linear-light frequency response depends on the signal amplitude. Since this problem is insoluble, we are forced to use the curves shown in Figure 13, and to assume that the errors in so doing are small.

The response of the camera itself is much less well documented. However, for a ccd sensor whose pixels are exactly square and abut their neighbours (i.e. no light is lost), then the frequency response is effectively that of the pixel aperture and follows the standard Sin(*x*)/*x* curve for any sampling process:

$$mtf = Sin\left(2\pi f / f_s\right) / \left(2\pi f / f_s\right)$$

where f is frequency and f_s is the sampling frequency (74.25MHz for HDTV). Figure 14 shows this response for luma[1] (Y) and chroma[2] (C), for which the response falls to zero at half the pixel counts of 960 and 480 respectively (for a 1920 pixel camera, with normal 4:2:2 sub-sampling). This is the native response of the camera and takes no account of either the lens or of any optical pre-filters used to suppress frequencies above 960 cycles per picture width that would give rise to alias components if allowed to reach the ccd. Figure 14 also shows the same curves multiplied by the response of the electrical filter shown in Figure 13, and this is the actual horizontal resolution curve that applies to our hypothetical HDTV camera. And so it is the horizontal *MTF* of the camera, ignoring the lens.

It is this effect, the Fourier transform of the pixel aperture, that is ideally "connected" by *AK*, aperture correction.

Again, there is the thorny issue of gamma-correction, because the filter shape is applied to gamma-corrected signals, but the $\mathrm{Sin}(x)/x$ curve is derived for linear signals. Strictly speaking, the multiplications used in Figure 14 cannot be done, because a linear response has been multiplied by a gamma-corrected filter. Since the effect of the filter is clearly marginal, the error is likely to be marginal as well, so the curves of Figure 14 can be taken as the horizontal MTF of the camera.
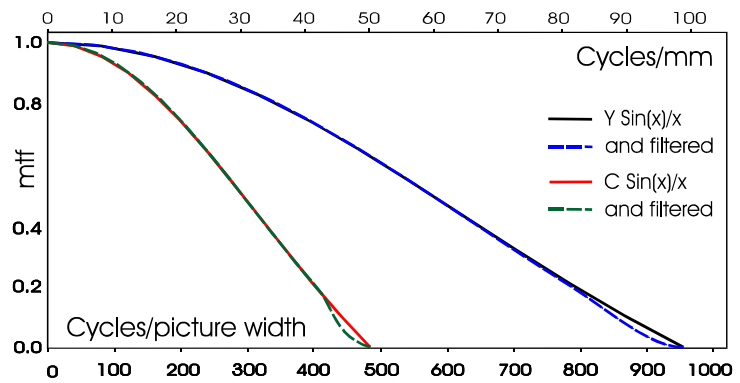

Figure 14 - Sin(x)/x response of video camera

The vertical response is defined in exactly the same way, described by the same curve shape, reaching zero at half the line count. If the camera is being operated in an interlaced mode, then the *interlace factor*, often mistakenly called the *Kell factor* will shift the limiting frequency down to about 70% of half the line count. The normal scanning process in the ccd sums adjacent line pairs to produce each output line, and this should halve the maximum reproducible frequency, but the alternate phasing of the summation which creates the interlaced field pairs restores a proportion of the frequency response above 50%. The *Kell factor* applies to the entire scanning process, whether interlaced or not, and effectively limits the resolution to about 70% of half the line count in each field or frame. Thus the vertical frequency response of a progressively scanned camera is about 70% of half the line count (*Kell*) while that of an interlaced camera is about 50% of half the line count (*Kell* times the *interlace factor*). A camera being used to mimic film, should always be operated in a progressive scan mode, and so it should have a limiting resolution of about 70% of half the pixel count horizontally and 70% of half the line count vertically.

There are exceptions to this rule for vertical resolution, it is becoming common to see *Extended Vertical Resolution* in interlaced cameras, a process which provides the vertical resolution of progressive scanning by avoiding the line summation. This makes the pictures look sharper, at the expense of one stop of exposure, and increased interline twittering due to the high-frequency content.

2.3.2 Dynamic MTF

When anything in the image moves, or the whole image itself moves, the *MTF* changes unless the exposure time is infinitesimal. For normal film-style shooting, the exposure time will be 50% of the frame interval since that is what a film camera would do. The same analysis process applies, using the same equations, but the pixel count is effectively reduced according to the image motion. Suppose that the camera moves horizontally at such a rate that picture points move by exactly 2 pixels per frame period (i.e. a picture point

---

[1] The term "luma" is used here rather than luminance, since luminance has a specific meaning in colour science. The luma signal in television is constructed by adding proportions of the gamma-corrected RGB signals, whereas true luminance can only be constructed by adding the appropriate proportions of the *linear*, non-gamma-corrected RGB signals

[2] The term "chroma" is used to refer to the colouring information, usually called chrominance. Chroma forms a satisfying rhyming couple with luma.

traverses the image in 19.2 seconds, 480 frames at 25fps), then each picture point is now sensed by two pixels instead of one during the exposure time. Thus the effective ccd pixel count is now 960 instead of 1920 for the moving content. The camera *MTF* shape is not just squeezed horizontally such that the zero point is at 480 pixels instead of 960, it must be multiplied by the original curve because that is still the *MTF* of the actual camera. The full equation for the *MTF* then becomes:

$$mtf = \frac{Sin\left(2\pi f \Big/ \left(\frac{2f_s}{n}\right)\right)}{\left(2\pi f \Big/ \left(\frac{2f_s}{n}\right)\right)} \frac{Sin(2\pi f / f_s)}{(2\pi f / f_s)}$$

where *n* is the number of pixels of motion per frame period. Personal observation, over many years of watching films both on television and in movie theatres, suggests that the general uneasiness and grittiness of motion at low speeds breaks into two-track motion, as shown in Figure 10, at about 3 seconds per picture width. For the 1920 pixel HDTV camera this amounts to 640 pixels per frame period, so information is smeared over 320 pixels during the exposure period. Figure 15 shows how motion of as little as 20 pixels per frame period (10 pixel smear, 3.9 second per picture width) affects the camera *MTF*.
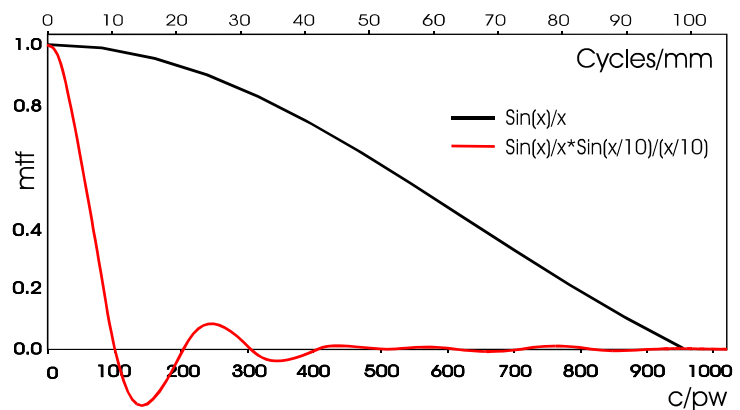

Figure 15 - Sin(x)/x responses of video camera

The red curve has its first zero crossing at 96 cycles per picture width, one tenth of the normal camera resolution, but it has at least two negative lobes of significant amplitude. In practice, the actual RGB signals do not go negative (i.e. below black level) because the *MTF* curve represents the amplitude response of the camera to detail components in colours that are otherwise normally visible. Where the curve goes below the axis, these detail components are reproduced in antiphase to their original content, i.e. they are frequency alias components. In practice, this means that not only are the frequencies aliased, i.e. reproduced at lower frequency, but that the moving detail in the first negative lobe goes in the opposite direction. In the next lobe (positive) the detail goes the right way but at twice the speed, in the next negative lobe the detail goes the wrong way at three times the speed, and so on. Thus these alias components represent moving detail at many frequencies going the wrong way or at the wrong speed or all three, and apply an extra confusion to the eye as well as the uneasiness or judder familiar to film-viewers. Ideally, they need to be suppressed more if film judder is to be replicated well.

2.3.3 Video Aperture Correction (AK) and Detail Enhancement (DE)

All HDTV cameras, and most broadcast or professional cameras, have a range of controls that can manipulate the performance to some degree. *AK* or *DE* should always be available. The original purpose of *AK* was to compensate for the smooth roll-off seen in the Sin(*x*)/*x* response shown in Figure 14 by artificially boosting higher frequencies in a noise-free way. It was also used to partially correct for the plastic deformation of the scanning spot in vacuum tube cameras. Thus, *AK* could be fixed at the design stage, and leave no controls for the user. Sadly, this happens only in consumer camcorders, where the amount of boost is almost always too high and at too low a frequency anyway.

AK survives in ccd cameras where it is often used to abuse otherwise good pictures. Figure 16 shows what misuse of *AK* or *DE* can do. The original photograph was taken (of Loch Lomond) with a digital stills camera, 2048 by 1536 pixels. The picture was artificially sharpened to illustrate the effect, and then

sub-sampled to 307 by 255 pixels for inclusion here. All processing was done in Paint Shop Pro. These examples are not intended to represent what is normally done in video cameras, only to illustrate the sort of harm that can be done by injudicious use of the controls. Clearly, the right-hand example is over-enhanced because there are hard edges added around contrast edges. Even the centre example is probably overdone, since it has a slightly un-real look.

Evidently a good tool, originally designed to correct a failing in early television cameras, is now being used for the wrong purpose. Strictly, the functions of *AK* and *DE* should be separated, since *AK* could be fixed by the manufacturer, to maintain the camera resolution up to a sensible limit, leaving *DE* as a user adjustment. But it should be emphasised that the operation of these controls is usually aimed to flatten the overall

Original, no enhancement   Medium enhancement   Excessive enhancement

Figure 16 - detail enhancement

response of the camera and lens, rather than to mimic any specific performance requirement.

The normal controls for *AK* or *DE* (the terms are often used interchangeably) include:

- Amount, positive going only, with zero as the minimum setting.
- Centre frequency of correction, adjustable over most of the frequency range.
- Q or bandwidth of the correction, affecting the spread of the correction over the frequency range.
- Clipping, to limit the amount of correction that can be applied.
- Noise coring, to prevent low level signals or noise from being emphasised.
- Horizontal/Vertical ratio, or separate controls for vertical enhancement.
- Extraction, the control signal may come from the luma channel or any combination of the RGB channels.

Ideally, a camera should have separate controls marked as *AK* and *DE*. Under those conditions, the functions can be separated, so that *AK* raises the response at high frequencies, which is what it was originally designed for, and *DE* used to control the middle frequency response. It would be nice if that control could go negative as well as positive, i.e. to lower the *MTF* in the middle frequencies, for that would help with motion judder, since it is the middle frequencies that contribute most to that effect, the alias lobes seen in Figure 15.

2.3.4 Camera measurements

The response of a High Definition camera was measured, with gamma and all other functions that could affect the results, switched off. A reference zoom lens, designed for 2/3" HDTV, was used at medium zoom and aperture. As mentioned earlier, the test signal was BBC/ITVA test chart 60, a back-illuminated transparency. It was set up at a variety of distances/lens angles, to provide the camera with sinusoidal test patches at relevant frequencies. Figure 17 shows the
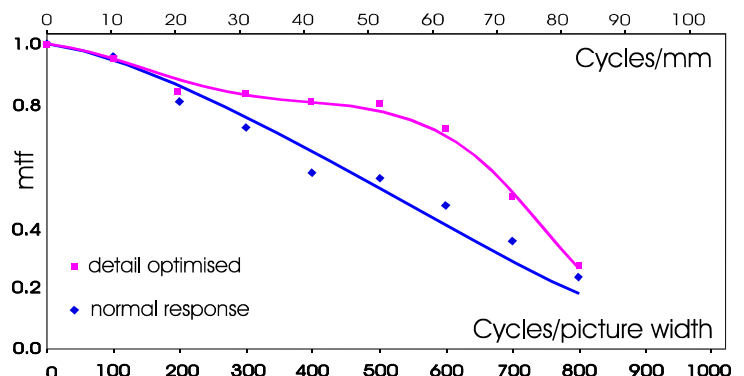

Figure 17 - camera horizontal response

13

measurement results. The result of setting the *DE* to "optimum" is also shown. The intention was to raise the higher frequencies enough to approximate a flat response where possible, but without producing any of the ringing effects visible in Figure 16. Pictures looked clean and sharp but not artificially so. The normal response of the camera is very similar to the theoretical *MTF* of the camera (Figure 14) and so the lens must be having only a small effect on the overall results, we are seeing the response of the camera. The lens is not a limiting factor on camera performance. This is a surprising conclusion.

Clearly the effect of the *DE* is maximal at about 580 cycles/picture width. Setting this frequency lower would have raised the lower frequencies more at the expense of the higher frequencies. This setting is a good compromise for a "video look".

It is interesting to compare this result with the response of a 35mm stills photography lens (a prime), again borrowing data from Hunt. This is shown in Figure 18. The curve shape is very similar to that of the un-enhanced video lens in Figure 17, but that curve is the *MTF* of both the lens and the camera, not just of the lens. The implication must be that the HDTV video lens produces much more detail than the stills lens would, at the same image magnification. Thus, the stills lens, although a good one in its own field, would not make good television pictures. Video lenses are more expensive than stills lenses, mostly because the image format is much smaller so the *MTF* when expressed in cycles/mm must be much better.

2.3.5 Comparisons, video vs. film

Since the lens is a common feature of both film and video recording, it can be omitted from comparisons between film and video cameras. Figure 19 compares the film *MTF* of Figure 12 with the theoretical video *MTF* shown in Figure 14. It is reasonable to use the film G layer to represent luma and the R layer for chroma. Clearly, film beats video at higher frequencies, and produces significant detail at frequencies beyond the limits of HDTV. However, video is better in the middle frequency range. Application of *DE*, as shown in Figure 16 will only increase this effect, so it would be sensible to reduce the middle frequency response in order to make a better match to film. Fortunately, one HDTV camera permits the *DE* setting to go below zero, i.e. to reduce the response instead of enhancing it. Figure 20 shows how this could work, by taking the enhancement shown in Figure 17
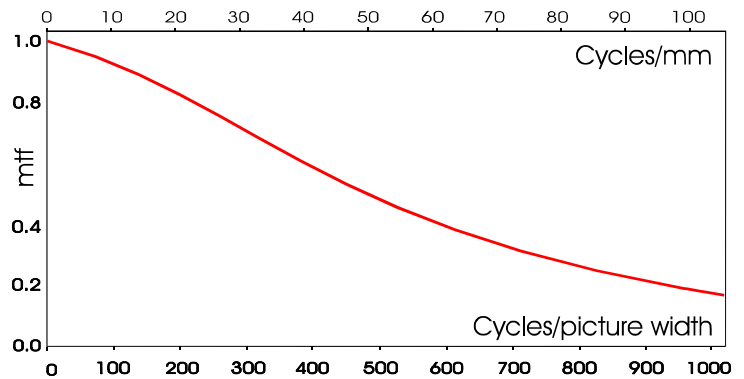

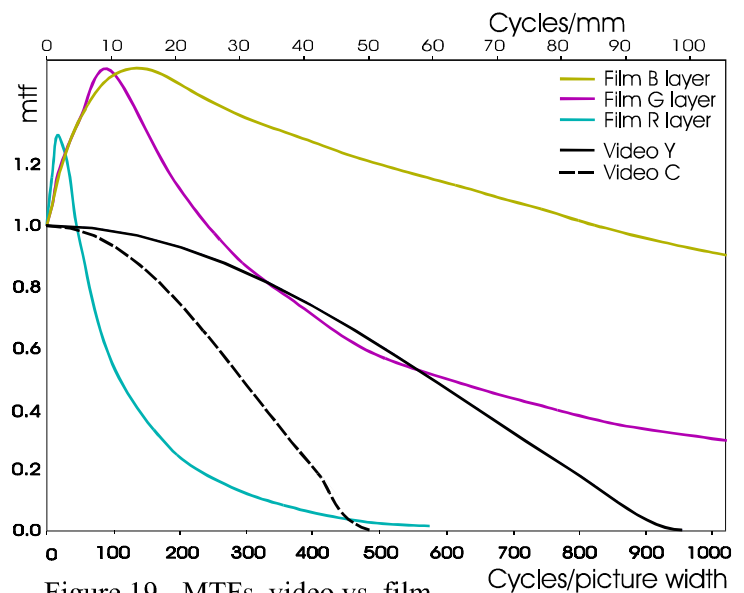Figure 18 - MTF of typical 35mm stills lens
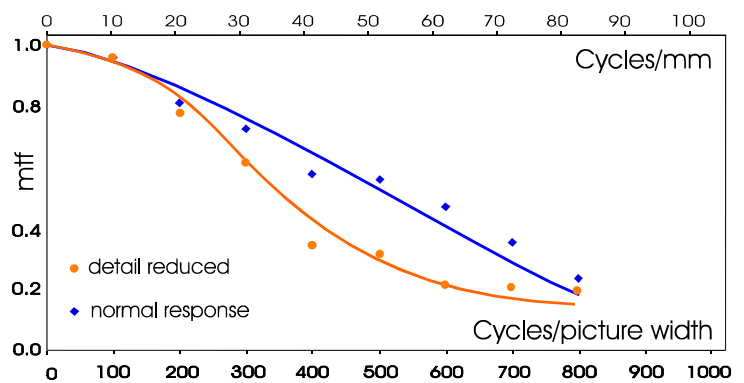

Figure 19 - MTFs, video vs. film

Figure 20 - "filmised" camera horizontal response

but subtracting it rather than adding. The "filmised" MTF curve much more closely resembles that of film. It would probably also be beneficial to make it work at rather lower frequencies, so that it attenuates the motion aliases seen in Figure 15, this would also prevent undue attenuation of the higher frequencies that are needed to make the pictures still look sharp when they are not moving.

The chroma response of video is clearly much higher than the R layer of film except at very low frequencies. Although not affecting the general sharpness of the picture, the extra resolution will better reproduce coloured textures, such as skin details. This may not be desirable in the attempt to mimic film, but most HDTV cameras have a menu option to reduce resolution in a definable colour zone by applying a form of negative *DE*.

3   *Conclusions*

It is now possible to get video cameras to mimic film performance in a convincing way. But several aspects of the video camera's performance must be preset for this to happen, listed here in probable order of importance.

- Film motion (judder): operate the camera in progressive scan mode at the desired frame rate (25fps for television). The shutter should be about 50% to mimic the 180° of the normal film camera. The resulting performance gives motion judder that closely resembles that of film, but see below as well.

- Contrast range: preset the gamma-corrector, black stretch and knee controls to capture 11 stops range. HDTV cameras can achieve this, as can some high end SDTV cameras. Setting these controls optimally is a laboratory operation and cannot be done reliably in the field. The resulting performance captures about the same range as does film, albeit without the straight-line performance that film can deliver over its central exposure range.

- Modulation Transfer Function: video camera cannot match film in this aspect. Usually all controls are switched off, but in some cases Aperture Correction can be set to maintain or increase sharpness in fine detail and Detail Enhancement can be set to affect middle-to-low frequencies. If the Detail control can be made to reduce rather than enhance detail then motion judder should better approximate to the performance for film. Obviously this requires separate controls and for the Detail control to go negative, but this is true in at least one recent HDTV camera. Setting this is, again, a laboratory operation and cannot be done reliably in the field. The resulting performance more closely matches that of film, particularly in reducing motion judder and enhancing (by reducing) the apparent depth of field, both of which are often regarded as being too harsh when derived from a progressively scanned video camera.

- Depth of Field: video cameras with 2/3" sensors have very similar depth of field to that of super 16mm film, which is much more than that of 35mm film. It is impossible to reduce the depth of field to match 35mm simply because lenses cannot be built with that specification. But it is possible, with ingenuity, to use 35mm film lenses on video cameras. A recently announced optical adaptor permits the use of 35mm film lenses on video cameras, maintaining the view angle and depth of field on the lens. It has not yet been field tested by the BBC.

- Skin tones: video has better colour resolution in red than does film. In some circumstances it may be useful to deploy a feature of recent cameras to detect and soften any skin tone. Although best set under laboratory conditions, this is one setting that can perhaps be done in the field, since it may need to be changed from shot to shot. However, it requires high resolution monitoring to see the effects, and it is very easy to overdo it. Therefore it should be regarded as a laboratory operation.

Other factors that contribute to a "film-look" are perhaps best regarded as faults of film, rather that features, and perhaps should best be ignored.

Video cameras, when set up according to these principles, can be thought of as electronic film. The camera performs the function of the film camera, negative processing and telecine. The process of designing the settings can be thought of as designing a film stock. This can lead to a straightforward way of using video cameras to perform film-type functions, without the photographer needing to worry about the settings. Modern cameras have a huge number of controls, albeit often buried deep in a menu structure, and can seem

daunting at first sight. But careful design of the setup can remove most of the photographer's concern about how the camera performs, and allow him or her to get on with the job of making programmes.

4 References

1   A.Roberts, "The Colorimetric and Resolution Requirements of Cameras"; BBC R&D White Paper WHP-034, 2001
2   Dr.R.W.G.Hunt, "The Reproduction of Colour"; Voyager Press, sixth edition 2002 (earlier editions from Fountain Press are still valid).
3   BBC specification TV2248 "Colour Television Cameras"
4   International Telecommunications Union (ITU) ITU-R.BT709 "Parameter values for the HDTV standards for production and international programme exchange".
5   W.N.Sproson, "Colour Science in Television and Display Systems"; Adam Hilger Ltd, Bristol, 1983.