

(51) International Patent Classification:
H04N 5/202 (2006.01)(21) International Application Number:
PCT/GB2015/050334(22) International Filing Date:
6 February 2015 (06.02.2015)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
1403448.2 27 February 2014 (27.02.2014) GB(71) Applicant: **BRITISH BROADCASTING CORPORATION** [GB/GB]; Broadcasting House, London W1A 1AA (GB).(72) Inventors: **BORER, Tim**; c/o British Broadcasting Corporation, BBC Research & Development, South Lab, BBC Centre House, 56 Wood Lane, London W12 7SB (GB).
COTTON, Andrew; c/o British Broadcasting Corporation, BBC Research & Development, South Lab, BBC Centre House, 56 Wood Lane, London W12 7SB (GB).(74) Agent: **REDDIE & GROSE LLP**; 16 Theobalds Road, London WC1X 8PL (GB).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

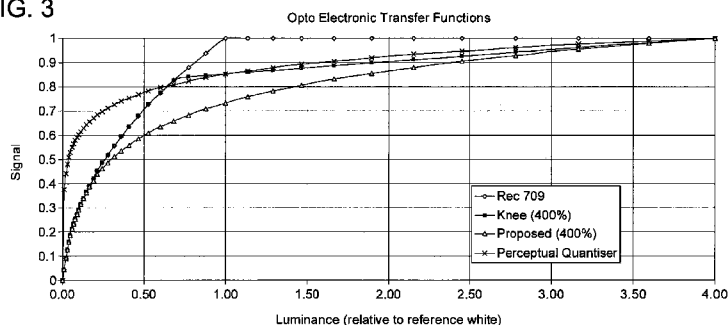
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))

(54) Title: METHOD AND APPARATUS FOR SIGNAL PROCESSING FOR HIGH DYNAMIC RANGE VIDEO

FIG. 3



(57) Abstract: A method of processing a video signal from a source to produce an output signal comprises converting between a luminance value and signal value using a conversion that has a power of the luminance value for the lower range of luminance values and a logarithm of luminance value for an upper range of luminance values. Such an arrangement allows the signal to be compatible with existing transfer functions using a power law over the bottom part of the range and to maintain some compatibility with existing systems over the upper part of the range. A signal produced by the method may therefore be used both with high dynamic range and standard dynamic range displays.

METHOD AND APPARATUS FOR SIGNAL PROCESSING FOR HIGH DYNAMIC RANGE VIDEO

BACKGROUND OF THE INVENTION

5 This invention relates to processing a video signal from a source, involving converting between a luminance value and signal value.

 A particular non-limiting example is conversion of a luminance value from a camera to a "voltage" signal value for subsequent processing. For many years, a power law with exponent 0.5 (i.e. square root) has ubiquitously been used in
10 cameras to convert from luminance to voltage. This opto-electronic transfer function (OETF) is defined in standard ITU Rec 709 (hereafter "Rec 709") as:

$$V = \begin{cases} 4.5L & \text{for } 0 \leq L < 0.018 \\ 1.099L^{0.45} - 0.099 & \text{for } 0.018 \leq L \leq 1 \end{cases}$$

15 where:

 L is luminance of the image $0 \leq L \leq 1$

 V is the corresponding electrical signal Note that although the Rec 709 characteristic is defined in terms of the power 0.45, overall, including the linear portion of the characteristic, the characteristic is closely approximated by a pure
20 power law with exponent 0.5.

 Combined with a display gamma of 2.4 this gives an overall system gamma of 1.2. This deliberate overall system non-linearity is designed to compensate for the subjective effects of viewing pictures in a dark surround and
25 at relatively low brightness. This compensation is sometimes known as "rendering intent". The power law of approximately 0.5 is specified in Rec 709 and the display gamma of 2.4 is specified in ITU Rec 1886. Whilst the above processing performs well in many systems, for various historic reasons discussed later, we would like to improve the dynamic range of signals beyond current
30 standard displays.

 Various attempts have been made to improve upon the existing standard. For example, Dolby Laboratories have proposed a new OETF to provide higher dynamic range for video and movie production and distribution. This OETF is

based on the human contrast sensitivity model. The proposed "ideal" OETF is approximated by a model given in the equations below. Here V is the signal value and Y is the brightness.

5

$$V = \left(\frac{c_1 + c_2 Y^n}{1 + c_3 Y^n} \right)^m$$

where

$$n = \frac{2610}{4096} \times \frac{1}{4} \approx 0.15930176$$

$$m = \frac{2523}{4096} \times 128 = 78.84375$$

$$c_1 = c_3 - c_2 + 1 = \frac{3424}{4096} = 0.8359375$$

$$c_2 = \frac{2413}{4096} \times 32 = 18.8515625$$

10

Over a wide dynamic range, from $Y=0.001$ to 1 , the proposed Dolby OETF, quantised to 10 bits, has a remarkably constant Weber fraction of about 1% (the difference in brightness represented by adjacent quantisation levels, as a fraction of the brightness, is known as the Weber fraction). Extending the dynamic range to encompass $Y=0.0001$, a dynamic range of 10000:1, the Weber fraction increases only slightly to 2%. However the Dolby proposal extends the dynamic range to $10^7:1$, from 0.001 cd/m^2 to 10^4 cd/m^2 . This is achieved by increasing the Weber fraction to 4% at $Y=10^{-5}$, 9% at $Y=10^{-6}$, and a huge 29% at $Y=10^{-7}$. This can be justified by equating the value of $Y=1$ to be 10^4 cd/m^2 , because then the tiny values of Y ($<10^{-4}$) correspond to very low light level to which the eye is much less sensitive. At these low light levels the threshold for just noticeable differences is a much higher than at more typical levels of illumination.

25

Dolby's proposal to link the camera OETF to the absolute luminance level has potentially far reaching consequences. Previously the photographic, movie and television industries have always worked with relative, rather than absolute, luminance levels. Changing to absolute luminance levels will require significant changes to the way television is produced and viewed. Whilst the objective of

achieving a wider dynamic range is admirable, the practicality of such a change has yet to be established. Nor is it clear that such a large dynamic range is needed for image display, since the simultaneous dynamic range of the eye is probably less than 10^4 . Furthermore the dynamic range of film and the best
5 electronic cameras is also about 10^4 . And a dynamic range of 10^4 (100 times more than conventional television) may be achieved with a simple 10 bit log OETF with a Weber fraction of less than 1%.

Given the need to produce a video signal with higher dynamic range, and
10 the limitations of the Rec 709 OETF, many electronic camera manufacturers have designed their own OETF (mostly 10 bit). These include: Filestream [Thomson] (Thomson), S-Log [Sony] (Sony), Panalog [Galt][Panavision] (Panavision), Log C [Brendel] (Arri), Canon Log [Thorpe] (Canon, only 8 bit).

15 Modern digital motion imaging sensors can originate linear video signals having dynamic ranges up to 80dB or more requiring A/D conversion up to about 14 bits. Examples of cameras that support such dynamic range are the Cannon EOS C300 and the Arri Alexa. This dynamic range is similar to the simultaneous dynamic range of the human visual system, which is about 10000:1. That is,
20 humans can simultaneously, in the same scene, see brightness variations of this range, for example between shadows and highlights. Such dynamic range far exceeds the dynamic range of printed material (less than 100:1), of (now obsolete) CRT displays (less than 100:1) and, in practice, of modern flat panel displays (which, despite claiming huge dynamic ranges, are limited by the low
25 dynamic range signals supported on their interfaces).

For practical and historical reasons digital video, for the overwhelming majority of cases, is limited to 10 bits in professional video production systems, and to 8 bits for consumer equipment and computer graphics. For video
30 programme production we need to preserve as much dynamic range as possible to provide latitude for processing (e.g. "colour grading") during "post production". For consumer applications we would like to preserve dynamic range to simultaneously support details in shadows and highlights. By preserving more dynamic range we hope to provide a more compelling and immersive experience
35 to viewers on newer displays that could, potentially, support higher dynamic range.

SUMMARY OF THE INVENTION

We have appreciated the need to improve upon existing standards to increase dynamic range. We have also appreciated, though, the need to maintain compatibility with existing equipment in the broadcast chain as well as allowing for diversity of viewing devices (screens), user preference and viewing environment.

In broad terms, the invention provides a method of processing a video signal from a source, comprising converting between a luminance value and signal value according to a conversion in which: a power law is used for a lower range of luminance values in which the signal value is a function that includes a power of the luminance value; and a log law is used for an upper range of luminance values in which the signal value is a function that includes a log of the luminance value.

The use of the two differing functions over the two ranges has various advantages. Over a lower range of luminance values, the power law function may be substantially the same as the Rec 709 standard noted above. As a consequence, any existing display system processing a signal so produced in accordance with the inverse function of Rec 709 will produce an acceptable result. A level of backward compatibility to existing equipment is therefore maintained. Over an upper range of luminance values, the dynamic range may be extended, when processed by equipment using the inverse of the log law function, but the signal can still be processed in this range by any existing display system using the inverse function of Rec 709 and produce an acceptable result. Again, a level of compatibility is achieved.

The problem as to how to encode a high dynamic range signal into only an 8 or a 10 bit signal is therefore addressed. This becomes possible because the human visual system has a non-uniform sensitivity to light, which allows a non-linear transfer characteristic as described above to be used. Unlike previous proposals, processing according to an embodiment of the invention is somewhat compatible with the ubiquitous ITU Rec 709 characteristic.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention will be described in more detail by way of example with reference to the accompanying drawings, in which:

5

Fig. 1 is a pair of images showing a "contour" effect when quantisation is insufficient to render subtle shading changes;

10

Fig. 2 is a graph showing an example "Knee" characteristic as used in some cameras to extend dynamic range;

Fig. 3 is a graph showing a comparison of known opto electronic transfer functions in comparison to a function of the present invention;

15

Fig. 4 shows a graph of Weber fractions comparing known opto electronic transfer functions to a function for the present invention;

20

Fig. 5 shows a schematic system diagram for a known proposed opto electronic transfer function and the signals processed;

Fig. 6 shows the arrangement of Figure 5 but with the improvement of the present embodiment of the invention showing a reduction in signals processed;

25

Fig. 7 is a schematic diagram of a processing component of the embodiment of the invention; and

Fig. 8 is a schematic diagram of a display device of the embodiment of the invention.

30

DESCRIPTION OF A PREFERRED EMBODIMENT OF THE INVENTION

35

We will first discuss the historic and practical use of non-linearities in television, by way of background. This includes a discussion of the psycho-visual reasons for the use of non-linearities and the different reasons for non-linearity in analogue and digital systems. We will then discuss a new opto electronic transfer function (OETF) embodying the invention, which is intended to support significantly higher dynamic range whilst retaining a high degree of compatibility with the conventional "Rec 709" non-linearity. The mathematical basis of the proposal is explained along with its psycho visual rationale.

45

For many years the dynamic range of television displays was limited to a about 100:1 by CRT technology. A non-linear "gamma" curve was used to equalise the effect of noise at different brightnesses in analogue TV systems. With the advent of digital TV the same gamma curve also allowed video to be quantised to 8 bits with minimal visible contouring. Modern displays potentially support higher dynamic range images, but remain limited to 100:1 dynamic range by existing infrastructure and standards, particularly for interfaces to TVs. The conventional Rec 709 gamma curve does not support higher dynamic range (or at least brighter images), even if extended to 12 bits, because of the expectation that the 8 most significant bits are equivalent irrespective of the precision of the signal. Film and cameras have long been able to capture higher dynamic range, with modern film and electronic cameras supporting dynamic range up to 14 stops, i.e. >10000:1. The Rec 709 gamma curve is often modified in cameras, though the use, of a "knee" to extend the dynamic range and prevent the signal saturating. This is adequate for live television but for higher quality production higher dynamic range is achieved through the use of numerous non-standard, quasi logarithmic, OETFs.

The non-linearity in television was originally introduced to make the effects of noise more uniform at different brightness. The CIE (International Committee on Illumination) specifies a function, lightness or L^* , which closely approximates human vision's lightness response [CIE 1976]. It is, more or less, a power function with exponent 0.42. As a result of this non-linear visual response the same level of noise is much more visible in dark regions of an image than in bright regions. In an analogue television system, a non-linearity is required to make the subjective effect of noise uniform for regions with different brightness. Hence the signal was non-linearly compressed, with a power law of approximately 0.42, at the camera, and expanded again at the display to produce an approximately linear system overall but with more or less uniform visibility of noise. Early television engineers took advantage of the non-linear characteristic of CRT displays achieve this, since the non-linearity of a CRT closely approximates a power law of 2.4 (and 2.4 is approximately the reciprocal of 0.42). These power laws are commonly referred to as gamma laws. So the gamma of a CRT display is about 2.4 (and is specified in ITU Recommendation 1886), and

the overall gamma of the system described in this paragraph is 0.42×2.4 , which is approximately unity.

5 In practice a power law with exponent 0.5 (i.e. square root) is ubiquitously used in the camera. Combined with a display gamma of 2.4 this gives an overall system gamma of 1.2. This deliberate overall system non-linearity is designed to compensate for the subjective effects of viewing pictures in a dark surround and at relatively low brightness. This compensation is sometimes known as “rendering intent”. The power laws of 0.5 is specified in Rec 709 and the display
10 gamma of 2.4 is specified in ITU Rec 1886. In relation to an embodiment of the invention, the rendering intent implied by the overall system gamma is not very important as focus is on the OETF (Opto Electronic transfer Function) of the source such as camera.

15 With the advent of digital video a problem is noted as to how many bits were required to represent the gamma corrected signal from the camera without introducing additional artefacts? The artefact introduced by using insufficient bits is spurious boundaries introduced in smooth regions of the picture. This artefact is known as “banding”, “contouring” or “posterisation”. Figure 1 provides an
20 extreme example of effects of contouring.

A BBC Research Department Report from 1974 [Moore] provides both theoretical analysis and subjective experiments to estimate the number of bits required to quantise a gamma corrected video signal. Given a signal with 100:1
25 dynamic range, i.e. with black level (zero signal) luminance set at one 1% of the maximum (70cd/m^2) contouring was barely perceptible in a worst case scenario (a grey level ramp) using 9 bits, which could be reduced by two bits by dithering the digital signal. These results were in line with their theoretical analysis. They also tested real pictures, which they found required only 6 bits with or without
30 dither. However only 3 images were used which may not have been representative of critical modern images. The report recommended using 7 bits but noted that visible contouring could still arise under some circumstances. Subsequently widespread use of 8 bit signals has shown that 8 bits is sufficient for 100:1 dynamic range signals using a gamma curve non-linearity.

Ten bit quantisation has been increasingly used in video production. Unfortunately in practice, using Rec 709 gamma, this is of limited help in increasing the dynamic range of the image that can be supported. This is because it is expected that the 8 most significant bits (MSBs) of a 10 bit Rec 709 signal may be treated as an 8 bit signal. This often happens when, for example, the signal is coded for transmission (many transmission paths are only 8 bits) or when the signal is sent to a television or computer monitor (most televisions and monitors only display 8 bits). So the additional two bits are the least significant bits of the signal. They can reduce the minimum black level, but they can't increase the brightness of the scene, even though it is often more desirable to increase the maximum scene brightness. Similarly a 12 bit Rec 709 signal, which is specified in ITU Recommendation 2020, does not increase the maximum scene brightness that can be supported either.

Camera manufacturers would like to be able to support higher dynamic range and, particularly, brighter scenes. Their camera sensors have a dynamic range significantly greater than 100:1 and, if they could provide this dynamic range through to the display, pictures from their camera would look better. But with a standard Rec 709 signal about 100:1 dynamic range is all that can be achieved. To try to circumvent this limit camera manufacturers commonly modify the standard Rec 709 transfer characteristic and add a "knee" to the characteristic. In order to extend contrast at the bright end, camera manufacturers introduce a knee in the OETF transfer characteristic. The knee puts a break point in the OETF and compresses the contrast above the break point to fit into available signal range. An example of a knee characteristic is shown in figure 2 (for a 10 bit signal).

In this example the knee is at 100% reference white and the headroom available in the signal (levels 941 to 1019 in SMPTE standards 259 & 292) is used to increase the available exposure to 300% of reference white. Often the break point is set at about 85% of reference white output level (corresponding to an input exposure of about 70% of peak white). With an 85% break point the exposure may be extended up to 500% or 600% of reference white. Indeed [Roberts] suggests that, by using a knee, exposure may even be extended to 800% of reference white.

There is a price to pay for using a knee in the Rec 709 OETF. To achieve improvements in exposure, and maintain image quality, the camera must be carefully configured, and the output quality may easily be wrecked by inappropriate adjustments. Even when the camera is well adjusted there is a risk of visible contouring (discussed in more detail below). The characteristics of the knee depend on the camera and are not standardised. This makes it difficult to take full advantage of the enhanced dynamic range because it is difficult to undo the knee characteristic to return to a linear light format for processing. Consequently whilst a Rec 709 OETF with knee would typically be used for live TV programmes alternative approaches may be used for non-live content such as dramas.

We have appreciated that the dynamic range of the video signal may be increased by using an alternative OETF. In particular, we note that the gamma curve specified in Rec 709 was designed to produce an approximately uniform perception of video noise in an analogue signal. So it was designed to approximate the subjective lightness curve experience by the human visual system. But in quantising a video signal the objective is to avoid contouring, not to provide uniform perception of noise. So in quantising a video signal the important characteristic is the human visual system's ability to distinguish similar values of brightness.

In quantising a video signal we wish to avoid contouring, and so it is the likelihood of detecting the difference between adjacent quantisation levels that is important (not lightness). The just noticeable difference in brightness is governed by Weber's law mentioned above (modified to the De Vries-Rose law at low luminance as discussed below). Weber's law is that detectable difference between brightnesses (or, more generally, other stimuli too) is proportional to the brightness. That is that just noticeable difference in brightness is a constant fraction of the brightness, known as the Weber fraction. For brightness the Weber fraction of cone cells in the eye is between 2% and 3%, which means that subjects can reliably detect a change of between 2% and 3% in brightness.

We have appreciated that Weber's law suggests that a logarithmic OETF (signal $\propto \log(\text{relative luminance})$) would provide the maximum dynamic range whilst rendering quantisation steps imperceptible. A Weber fraction of 2% means we could quantise a 100:1 dynamic range, without perceptible contouring, using
5 233 quantisation levels, i.e. 8 bits. The aforementioned BBC report suggests only 9 bits are needed worst case, for a Rec 709 OETF, and fewer bits in practice. So for a required dynamic range of 100:1 there is little to be gained from using a logarithmic OETF, which would be incompatible with pre-existing TV equipment.

10 For movie production the formulation of the film stock ensures a more or less logarithmic response to light. For decades this has been digitally scanned using a linear analogue to digital converter (ADC). The resulting "Cineon" format is essentially a 10 bit signal with a logarithmic OETF. Film provides about 14 stops (i.e. 2^{14}) of dynamic range, which is satisfactorily captured in the 10 bit
15 Cineon format.

The present embodiment provides benefits in terms of dynamic range using a logarithmic transfer characteristic that has not previously been adopted for video production and distribution. We note that, hitherto, the dynamic range of
20 television signals has been limited to 100:1, which may not require a logarithmic curve, and that a logarithmic curve would appear to be incompatible with the installed television infrastructure. If a higher dynamic range for the end viewer is required in new television standards a conventional gamma curve, even extended to 12 bits, is no longer adequate.

25 We will refer to the processing as used in embodiments of the invention as a compatible OETF for enhanced and high dynamic range. Modern displays support higher brightness, higher dynamic range and lower noise than conventional, obsolete, CRT displays. Yet they remain limited by conventional gamma non-linearity, and the standard interfaces based on it, to the dynamic
30 range of about 100:1 supported by CRT displays.

In order to exploit the potential of modern displays a signal format, based on a new OETF, is needed. We propose a new OETF for high dynamic range video which is broadly compatible with Rec 709. Clearly no new OETF can be
35 completely compatible with Rec 709 (else it would be Rec 709). However the new

OETF should have similar characteristics to Rec 709, and so facilitate compression, video processing and allow the display of good quality pictures on existing displays.

- 5 A new OETF would allow video to provide greater impact and a more immersive experience. It should support a dynamic range of more than 1000:1 and preferably significantly more. Even a dynamic range of 1000:1, using a logarithmic, OETF, would require a Weber fraction of about 3% if the signal is quantised to 8 bits. And a 3% Weber fraction is barely, if at all, adequate to avoid
10 visible contouring in the displayed image. Which suggests a new OETF may need more than 8 bits of precision.

- Any new OETF should, we appreciate, also interoperate with existing standards and infra-structure. Ideally a new signal format should be able to be
15 carried over existing video connections (including compression) and be displayed, in a broadly compatible fashion, on existing 8 and 10 bit displays. Television will increasingly encompass multiple formats (e.g. SD, HD and UHD) and frame rates. So a new signal format should also, ideally, be compatible with processing such as spatial up and down conversion and frame rate ("standards")
20 conversion.

- Not all OETFs would be equally compatible with existing standards and infra-structure. Any reasonable 10 bit OETF may be carried over existing 10 bit interconnects. However the more the signal differs from the conventional Rec 709
25 characteristic the less compatible it will be with existing video compression, processing and, particularly, displays. Video compression is optimised for the characteristics of conventional video, so that new formats are likely to require greater bit rate and/or exhibit more artefacts. Video processing is typically performed on non-linear signals so a new signal format may degrade the quality
30 of such processing. And clearly a signal that is radically different from Rec 709 will be significantly distorted on a conventional display. Potential display distortion include altered brightness, changed colours, reduced or excessive sharpness and, perhaps, contouring. For this reason both (quasi) logarithmic OETFs and Dolby's perceptual quantiser OETF, are likely to present considerable practical

difficulties if used with conventional infra-structure. Clearly there are advantages to a new format that is broadly similar to Rec 709.

The construction of the proposed OETF is based on Rec 709 but departs
 5 from both this and the “knee” characteristics in cameras mentioned above. Rec 709 is a two part curve with a linear part near black and a power law (gamma curve) for the majority of the input range. It is designed so that the value and the gradient of both curves match at the transition between them. Camera makers modify the OETF by adding a third section near white, by using a “knee”, to
 10 increase dynamic range and avoid clipping the signal. Unfortunately the section added above the “knee” risks introducing contouring artefacts (assuming an inverse ETOF) by failing to take full account of the psychovisual aspects of vision.

The present embodiment adds a further (third) part to the Rec 709 curve,
 15 at higher input luminance, to extend the dynamic range without re-introducing the risk of contouring. The new, upper portion, of the curve is a logarithmic function to allow for the non-linear response of the eye. And, as with the linear portion of the Rec 709 curve, the output value and the gradient of both curves are designed to match at the transition between them. Because this new proposal has some
 20 similarities to the widely used knee characteristic in cameras we can be confident that it is broadly compatible with Rec 709.

As a reminder, the OETF defined in Rec 709 is:

$$25 \quad V = \begin{cases} 4.5L & \text{for } 0 \leq L < 0.018 \\ 1.099L^{0.45} - 0.099 & \text{for } 0.018 \leq L \leq 1 \end{cases}$$

where:

30 L is luminance of the image $0 \leq L \leq 1$

V is the corresponding electrical signal

In Recommendation ITU-R BT.2020 (Parameter values for ultra-high definition television systems for production and international programme
 35 exchange) the same equation is specified as:

$$E' = \begin{cases} 4.5E, & 0 \leq E < \beta \\ \alpha E^{0.45} - (\alpha - 1), & \beta \leq E \leq 1 \end{cases}$$

“where E is voltage normalized by the reference white level and proportional to the implicit light intensity that would be detected with a reference camera colour channel R, G, B; E' is the resulting non-linear signal.

$\alpha = 1.099$ and $\beta = 0.018$ for 10-bit system

$\alpha = 1.0993$ and $\beta = 0.0181$ for 12-bit system”

Although not explicitly stated α and β are the solution to the following simultaneous equations:

$$\begin{aligned} 4.5\beta &= \alpha\beta^{0.45} - (\alpha - 1) \\ 4.5 &= 0.45\alpha\beta^{-0.55} \end{aligned}$$

The first equates the values of a linear function and the gamma function at $E = \beta$, and the second equates the gradient of the two functions also at $E = \beta$, thereby ensuring a smooth transition between the two parts of the curve.

The present embodiment adds a third, logarithmic, portion to the transfer function at higher values of luminance such that (equation A):

20

$$V = \begin{cases} 4.5L & \text{for } 0 \leq L < \beta \\ \alpha L^{0.45} - (\alpha - 1) & \text{for } \beta \leq L \leq \mu \\ \eta \ln(L) + \rho & \text{for } L > \mu \end{cases}$$

We have used the V and L notation from Rec 709 to avoid the confusion of having both dashed and undashed versions of the same variable (as in Rec 2020). Here L is still normalised to the same reference white level, but now L can exceed unity, i.e. this transfer function supports luminance greater than reference white. α and β are as defined in Rec 2020. μ , the breakpoint between the gamma and logarithmic sections of the curve, determines the maximum value of L for which $V \leq 1$, which is discussed in more detail below. The values of η and ρ are determined by equating the derivatives and the values of the gamma and log curves at the breakpoint μ .

Equating the derivatives yields:

$$\eta = 0.45\alpha\mu^{0.45}$$

And equating the values of the curves at μ (and substituting for η) yields:

$$\rho = \alpha\mu^{0.45}(1 - 0.45 \ln \mu) - (\alpha - 1)$$

5

The breakpoint between the gamma and log parts of the curves must be selected. It determines by how much the input luminance level may exceed reference white. It also determines how compatible the proposed transfer function is with Rec 709/Rec 2020. A low value of μ gives a higher dynamic range but a poorer compatibility with Rec 709. So the choice of μ is a compromise. With a little algebra the value of luminance at the maximum output $V=1$, may be found to be:

$$L_{\max} = \mu \exp\left(\frac{1}{0.45}\left(\frac{1}{\mu^{0.45}} - 1\right)\right)$$

15

We propose that a maximum value of luminance (when $V=1$), relative to reference white, should be $L_{\max}=4$. This implies a breakpoint of $\mu=0.12314858$. The corresponding output level is $V=0.33$ but, in practice, Rec 709 and this proposal are very similar for the whole of lower half of the output range. Taking into account headroom above $V=1$, which is allowed in the 10 bit coding scheme defined in Rec 709 & Rec 2020, this choice of breakpoint allows more than 600% of reference white to be coded before clipping is required.

25 A 10 bit signal conforming to the proposed OETF provides a sixteen fold increasing in dynamic range compared to an 8 bit Rec 709 signal. The maximum luminance is increased by a factor of 4 as above. Increasing the signal depth from 8 to 10 bits reduces the difference between successive quantisation levels by a further factor of 4, thereby increasing the dynamic range a low luminance ("in the blacks").

Figure 3 shows a comparative plot of several OETFs. The Rec 709 curve (partially hidden by the "knee" curve), clips at reference white. Clipping is avoided up to 400% of reference white by using the camera knee characteristic illustrated.

35 The proposed OETF offers the same extension in dynamic range but with a

smoother curve. At the bottom end it is (by design) very similar to the Rec 709 curve. At the upper end it is similar to the knee curve. Based on industry experience of cameras with knee curves we may be confident that the proposed curve has a high degree of compatibility with Rec 709. The known “perceptual quantiser” OETF is included for comparison. Over the range plotted it is extremely similar to a pure logarithmic transfer function with a (10 bit) Weber fraction of about 1% (see below). As a logarithmic transfer function it is clearly highly incompatible with the Rec 709 curve. However the perceptual quantiser is designed to be used over a much wider dynamic range so incompatibility is unavoidable. The comparison is only intended to show that a perceptual quantiser or, equivalently, a pure logarithmic transfer characteristic is not compatible with Rec 709.

To gauge the likely subjective quality of this proposal it is instructive to consider the Weber fraction, that is, the fractional steps between quantisation levels. This is illustrated in figure 4 for the 4 OETFs compared above. The Weber fractions for all OETFs are comfortably below the 2% to 3% threshold of visibility (except at very low relative luminance). The Rec 709 Weber fraction decreases monotonically up to reference white. Its low values indicate that coarser quantisation could be used, without being visible, for higher relative luminances. And this is what the knee function response does. But unfortunately it introduces a sharp increase in Weber fraction at the knee point and also has too low a Weber fraction at high relative luminance. The proposed OETF prevents the Weber fraction from reducing to unnecessarily low values and by doing so allows a higher dynamic range. The Perceptual Quantiser has a more or less constant Weber fraction of 0.09 across the range plotted and so corresponds closely with a pure logarithmic transfer function over this range.

At low values of luminance the Weber threshold of visibility is replaced by the De Vries-Rose law. This says the threshold of visibility is proportional to the square root of the brightness rather than to brightness. The brightness corresponding to the transition between De Vries-Rose and Weber depends on conditions, such as the size, frequency and duration of the (visual) signal. Typically the transition brightness is between .04 and 25 cd/m² [Sezan]. This transition brightness approximately corresponds to the breakpoint between the

gamma and logarithmic sections of the curve in the proposed OETF. Consequently higher Weber fractions in Rec 709 (and therefore in this proposal), at low values of luminance, do not result in visible contouring. Furthermore, if the peak brightness of a display using the proposed OETF is approximately a few
5 hundred cd/m^2 , then the OETF approximately corresponds to the psychovisual sensitivity of the eye.

Reducing the depth of the signal from 10 to 8 bits increases the Weber fractions by a factor of 4. The Weber fraction for the proposed OETF would
10 increase to 2.4%. This is still only at or below the threshold of visibility, particularly if dither is added or there is (a low level of) noise in the image. Consequently an 8 bit version of this OETF should be able to produce a higher dynamic range image without visible artefacts. It should be noted that an 8 bit HDR signal would include little exposure latitude to support post processing such
15 as grading, and so would be unsuitable for production applications. Similarly artefacts from end user compression are likely to be more visible with an 8 bit OETF. Nevertheless this would allow a higher dynamic range image to be transferred to a display via an 8 bit interface. Note that an 8 bit version of a Rec 709 OETF with a knee would have a peak Weber fraction of about 5%, which
20 would be clearly visible on some pictures.

A 10 bit signal using the proposed OETF provides a dynamic range of at least 1600:1 (10.6 stops), compared to 100:1 for Rec 709. This is sufficient for consumer displays and for some video production. However film and modern
25 electronic cameras can support dynamic ranges up to 14 stops. A higher dynamic range signal is needed to support high quality video and movie production.

We further suggest using the proposed OETF with a 12 bit signal to support an extremely high dynamic range image suitable for video and movie
30 production. Conventionally 2 "precision" bits are added as the least significant bits to extend an 8 bit Rec 709 signal to 10 bits. This nominally extends the dynamic range by a factor of 4 (because the bottom part of the Rec 709 curve is linear). However, as noted above, there is no increase in the maximum luminance that can be captured relative to reference white. If the proposed 10 bit
35 signal were extended to 12 bits in the same way, by adding two LSBs, it would

not adequately support the needs of video and movie production. Such a 12 bit signal would support a dynamic range of only about 13 stops and a maximum luminance of 4 times reference white.

5 It is proposed to extend the 10 bit signal to 12 bits by adding one MSB plus one LSB. This means doubling the range of the signal (V or E' in the equations above), from 0 to 2, and increasing its precision by a factor of two. A signal value $V=2.0$ with the proposed OETF corresponds to a (relative) luminance of about 1000 times reference white. The additional bit of precision also extends
10 the dynamic range by a factor of two, i.e. 1 stop, "in the blacks". Overall a 12 bit signal, constructed in this way would have a dynamic range of more than 20 stops. This is many stops more than the best film or electronic cameras providing plenty of scope for future improvements and, potentially, plenty of exposure latitude to meet the most demanding production requirements.

15

For video production in standard definition (Rec. ITU-R BT.601) or HD (Rec 709) both 8 and 10 bit signals are valid formats. 10 bit signals have been increasingly used for video production (although 8 bit signals are still sometimes used) but only 8 bits are currently ever used for end user distribution. To allow
20 interoperation of 8 and 10 bit equipment simple conversion between the 8 and 10 bit signals is essential. If 8 bit signals continue to be used for HDR (which is possible as discussed above) this will continue to be true. Simple conversion is afforded through the addition of precision bits to the 8 bit signal as the LSBs (as discussed above). This remains true for the proposal herein.

25

There is no requirement, in extending an HDR video format to 12 bits, to use only precision bits. There is virtually no 12 bit video equipment in existing broadcast infrastructure. So there is no requirement for backward compatibility. Indeed, as already noted, the use of only precision bits would not support the
30 needs of high quality production.

Using the 12 bit format proposed here would support easy interoperation with 10 (and 8) bit equipment. There are no proposals to use 12 bits for end user distribution. So the 12 bit format would only be used for production to provide
35 exposure latitude during grading and post production. Today's non standard

higher dynamic range formats (such as Log C or Panalog discussed above) are graded on monitors with 10 input bits (higher bit depth monitors are virtually unavailable). Similarly the final grading of HDR video will be performed using 10 bit monitors (because consumer monitors will be no more than 10 bit). This is facilitated by using the same OETF with a wider signal range ($V=0$ to 2) as proposed here. The 10 bit signal for the monitor could be derived by simply omitting the MSB and the LSB of the 12 bit signal. Of course, sometimes, the 12 bit signal would exceed the range of the 10 bit signal (precisely the reason for using the 12 bit signal). But in this case the over range signal would be easily seen because the signal would “wrap round” to black. Such out of range signals would be corrected during grading (one of the purposes of grading) to produce the final 10 bit output. In practice, to provide a more robust conversion from 12 to 10 bits, it would be more appropriate to clip the 12 bit signal to limit its range and round the signal, rather than simply omitting the MSB and LSB.

15

So far, we have focused on the opto-electronic transfer function of the video system. To complete the picture a brief discussion of the corresponding electro-optical transfer function (EOTF) for the display follows. The electro-optical transfer corresponding to Rec 709 is specified in Recommendation ITU-R BT.1886 (Rec 1866). Rec 1886 includes adjustment for legacy “brightness” and “contrast” controls, but discounting those the EOTF is specified as:

20

$$L = V^{2.4}$$

That is the screen luminance, L , is a pure power law, with a exponent (gamma) of 2.4, of the signal, V .

25

The overall gamma of the system, from scene luminance to screen luminance, is 1.2. As noted above, the OETF specified in Rec 709 closely approximates a pure power law with exponent 0.5 once allowance is made for the linear part of the curve near black. Combining this with the power law from Rec 1886 yields the overall system gamma of 1.2. This overall system non-linearity is intentional and designed to compensate for viewing conditions.

30

To be compatible with Rec 709 & Rec 1886 an electro-optical transfer function corresponding to the proposal herein should also provide an overall system gamma of 1.2. This implies the following EOTF (Equation B):

$$L = \begin{cases} V^{2.4} & \text{for } 0 \leq V \leq \xi \\ \exp\left(\frac{1.2(V - \rho)}{\eta}\right) & \text{for } V > \xi \end{cases}$$

5 where ξ is signal value V corresponding to the breakpoint luminance μ ($=0.12314858$) given by:

$$\xi = \alpha \mu^{0.45} - (\alpha - 1) = 0.32906248$$

Figures 5 and 6 demonstrate the reduction in processing needed with an embodiment of the invention in comparison to a proposed transfer function that departs from the Rec 709 standard. Figure 5 shows the process in which two signals need to be produced to remain compatible with an existing standard dynamic range screen as well as a true high dynamic range screen. A raw image is captured at capture step 10 and may be previewed at preview step 12 and passed to an imaging screen for editing at step 14. At this point, a separate colour grading 16 may be performed on the HDR and SDR images and separate output files 18 produced in a format such as MXF which are then separately provided to a play out server at step 20. A HDR and SDR transmission encoder 22 then provides two signals, a first standard dynamic range signal and a second different signal from which the HDR signal may be derived using the SDR signal. This is then sent to a broadcast chain 24 such as a satellite at step where both signals must be broadcast. As can be seen, there is redundancy in the chain as two separate signals must be prepared and transmitted. Separate SD and HD receivers 26,28 are then able to receive the respective signals/

25

Figure 6 shows the contrast with the improvement of the present invention. As with Figure 5, an image is captured at step 10 but at step 12 a new transfer function according to the present disclosure is applied prior to editing in a single editing screen at step 14 and colour grading at step 16. A single MXF file is produced 18 and sent to a transmission server 20 prior to encoding 22 and satellite uplink at step 24. The single signal is now receivable and can be displayed by both the SDR and HDR receivers in the manner described above. A significant saving is thus achieved.

30

Figures 7 and 8, respectively show an image capture device such as a camera, and a display device such as a TV screen embodying the invention. The camera of Figure 7 includes a detector CCD or CMOS 30, processor 34 and memory 36 and an output 38. The memory 36 holds the EOTF function as described above, particularly in Equation A, by which the processor processes the incoming luminance signal. The display device of Figure 8 has an input 40 for receiving a signal, processor 44 and memory 42 and display 46. The memory 42 holds the OEFT function such as in Equation B herein.

Figure 9 shows a further embodiment of the OETF that provides for a lower range of luminance values that the signal value is derived using a first function that includes a power of the luminance value, and for an upper range of luminance values the signal value is derived using a second function that includes a log of the luminance value. The compressive function of this embodiment outputs the same value as is input (linear) until the signal level reaches a threshold, or breakpoint. Thereafter, for larger values of the input signal, the output is a logarithmic function that compresses the input. The second stage of the OETF applies a power law function as defined in Rec 709. Therefore, overall, for the lower part of the input signal domain the OETF is a power law defined, for example, by ITU Rec 709 (because combining the linear function and power law gives a power law). For the upper part of the signal domain the OETF is a power of a logarithm.

The choice of compressive function is explained below. A compressive function may be obtained by generating a two part curve in which the lower part of the curve is the straight line:

$$y = x$$

And in which the upper part of the curve is a logarithmic function:

$$y = \beta \ln x + \delta$$

That is the compressive function is:

$$y = \begin{cases} x & 0 \leq x \leq \alpha \\ \beta \ln x + \delta & \alpha < x \end{cases}$$

The function is intended to have a domain $[0:k]$ and compress this to a range $[0:1]$.

- 5 The desired function is continuous and has a continuous first derivative. This provides a pair of simultaneous equations, below, from which the curve may be defined.

$$\alpha = \beta \ln \alpha + \delta$$

$$1 = \frac{\beta}{\alpha}$$

- 10 These equations may be solved to yield:

$$\beta = \alpha$$

$$\delta = \alpha(1 - \ln(\alpha))$$

- 15 The breakpoint, α , determines the amount of compression that the function provides. For a given breakpoint the compression factor, k , is defined by:

$$1 = \alpha \ln k + \alpha(1 - \ln(\alpha))$$

This may be solved to yield the compression factor as a function of the breakpoint, that is:

20

$$k = \alpha \cdot \exp\left(\frac{1}{\alpha} - 1\right)$$

So, for example, with a breakpoint of 0.25 this function provides compression by a factor of, approximately, 5.02.

25

CLAIMS

1. A method of processing a video signal from a source to produce an output
signal, comprising converting between a luminance value and signal value
5 according to a conversion in which:
for a lower range of luminance values the signal value is derived using a
first function that includes a power of the luminance value; and
for an upper range of luminance values the signal value is derived using a
second function that includes a log of the luminance value.
10
2. A method according to claim 1, wherein the first function, for the lower
range, is compatible with existing systems.
3. A method according to claim 1 or 2, wherein the first function, for the
15 lower range, is in accordance with ITU-R BT.601, ITU-R BT.709, ITU-R BT.2020
or similar standards.
4. A method according to any preceding claim, wherein the second function,
for the upper range, extends the dynamic range.
20
5. A method according to claim 4, wherein the second function, for the upper
range extends, the dynamic range beyond that of the first function used for the
lower range.
- 25 6. A method according to any preceding claim, wherein the signal is
compatible with existing systems such that existing systems may produce an
acceptable image from the signal.
7. A method according to any preceding claim, wherein the lower range and
30 upper range are contiguous in luminance values and signal values at an
intersection.
8. A method according to claim 7, wherein the gradient of the first function
substantially matches the gradient of the second function at the intersection.
35

9. A method according to claim 7 or 8, wherein the intersection is chosen as to maintain compatibility with existing systems whilst extending the dynamic range of the signal.
- 5 10. A method according to any preceding claim, wherein the upper range of luminance values extends to an integer power of 2 multiple of existing ranges.
11. A method according to any preceding claim, wherein for a range below the lower range of luminance values the signal value is derived using a third
10 function that is a linear relationship of luminance value, wherein the functions are in accordance with equation A herein.
12. A method according to any preceding claim, wherein the signal value is quantised to an 8 bit or 10 bit bit depth.
- 15 13. A method according to any preceding claim, wherein the signal value is quantised to a 10 bit bit depth and reduced to an 8 bit signal for acceptable display on existing systems.
- 20 14. A method according to any preceding claim, wherein the signal value is quantised to a 12 bit bit depth and reduced to a 10 bit signal by clipping so as to remove the most significant bit, and by removing the least significant.
- 25 15. A method according to claim 7, wherein the point of intersection is a controllable variable to optimise a balance between compatibility with existing systems and dynamic range.

16. A method of processing the output signal of any preceding claim, comprising processing the signal with an inverse function of a conversion in which:

5 for a lower range of luminance values the signal value is derived using a first function that includes a power of the luminance value; and

for an upper range of luminance values the signal value is derived using a second function that includes a exponential of the luminance value.

10 17. A method according to claim 16, wherein the functions are in accordance with equation B herein.

18. A display operable to process the output signal of any of claims 1 to 15, having means for processing the signal with an inverse function of a conversion in which:

15 for a lower range of luminance values the signal value is derived using a first function that includes a power of the luminance value; and

for an upper range of luminance values the signal value is derived using a second function that includes a exponential of the luminance value.

20 19. A method according to claim 18, wherein the functions are in accordance with equation B herein.

20. A transmitter comprising means arranged to undertake the method of any of claims 1 to 15.

25

21. A camera comprising means arranged to undertake the method of any of claims 1 to 15.

22. Apparatus being part of a studio chain comprising means arranged to 30 undertake the method of any of claims 1 to 15.

23. A method according to any of claims 1 to 15, wherein the first function is a power of the luminance value and the second function that is a power of a logarithm.

35

1 / 7



5-bit indexed (32 colours)



24-bit RGB (16.7 million colours)

FIG. 1

2 / 7

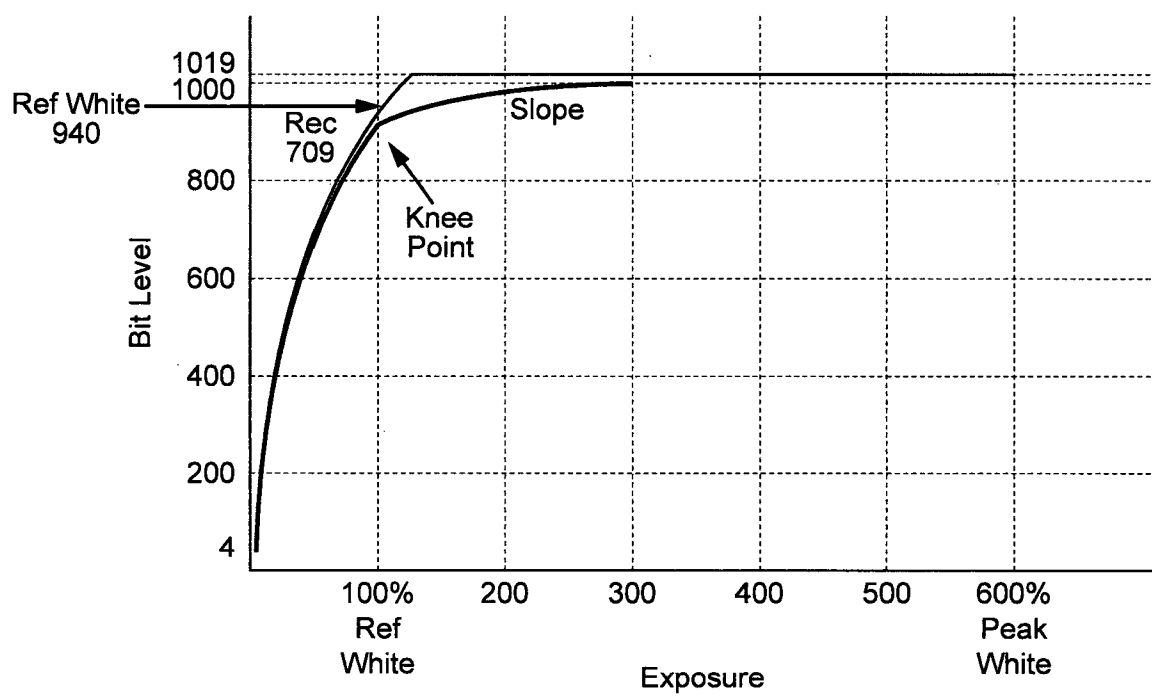


FIG. 2

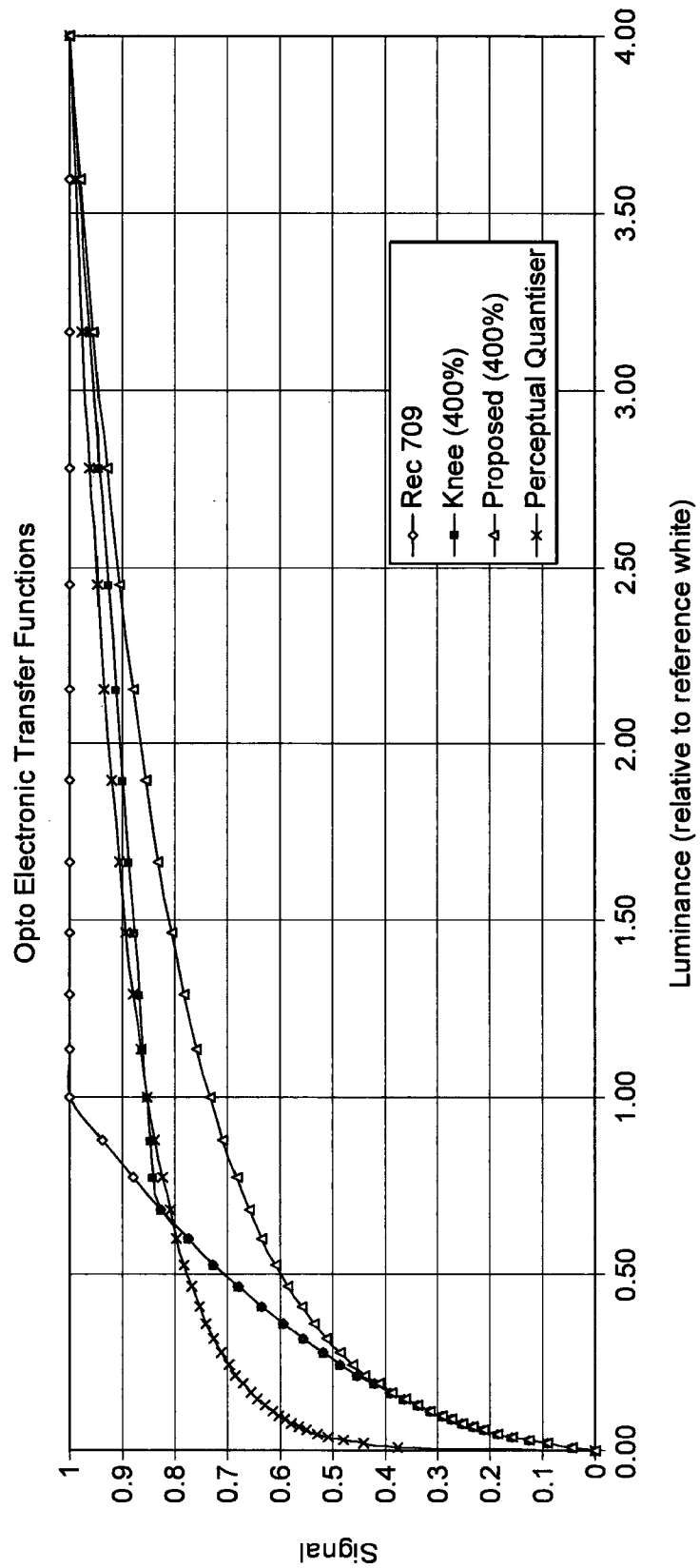


FIG. 3

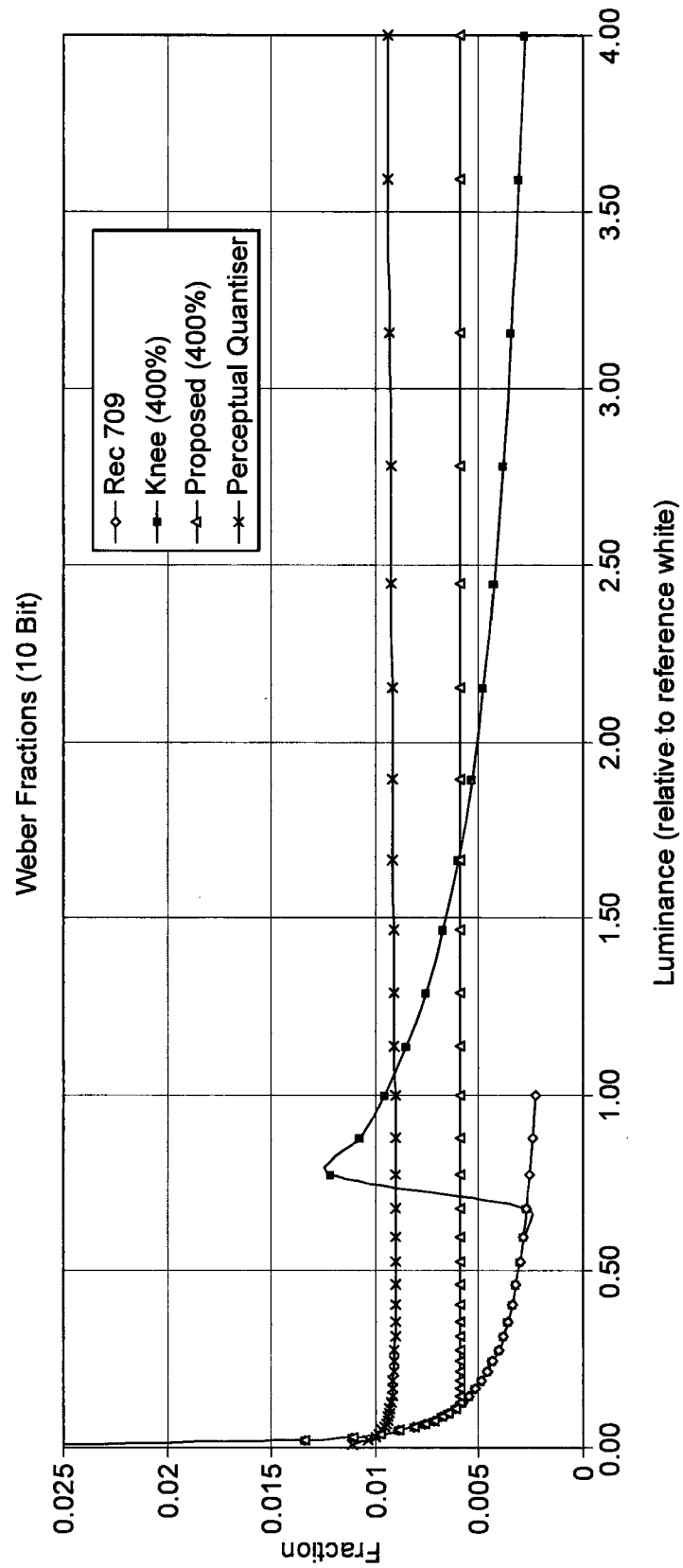


FIG. 4

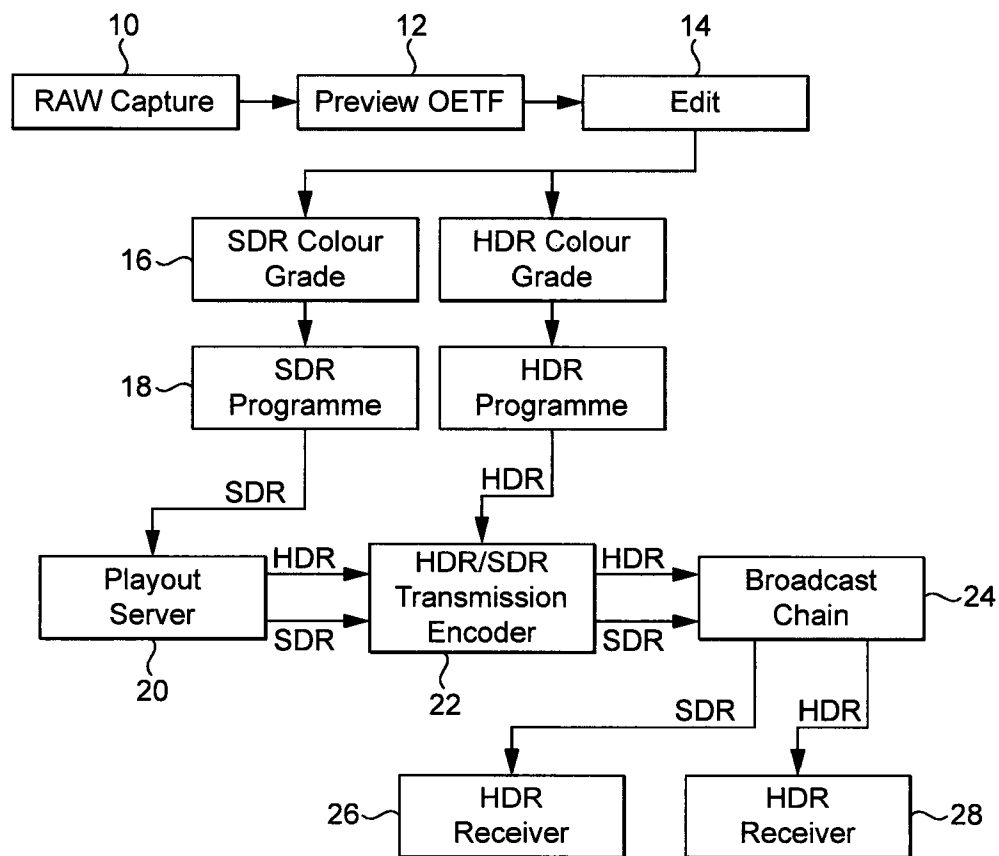


FIG. 5

6 / 7

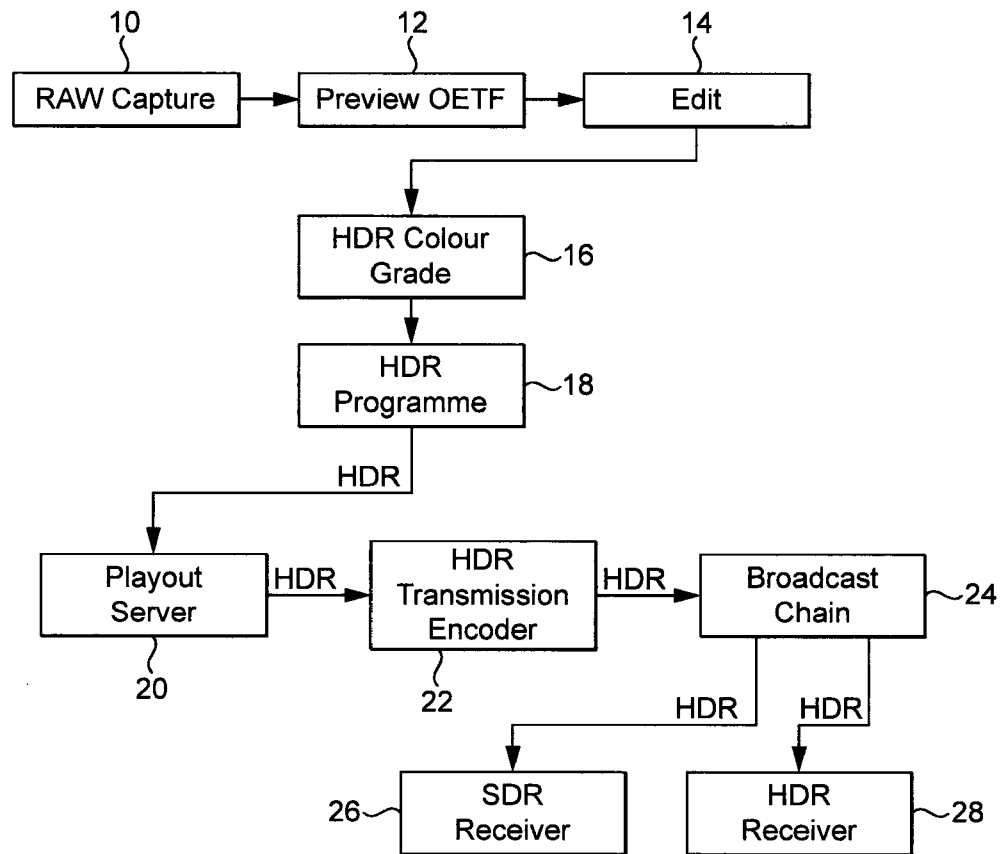


FIG. 6

7 / 7

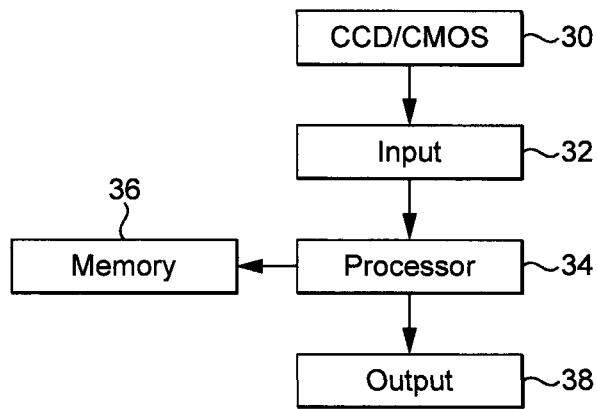


FIG. 7

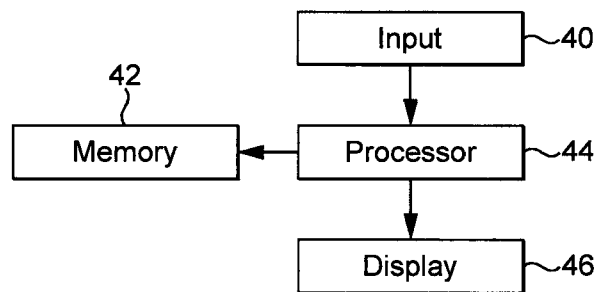


FIG. 8

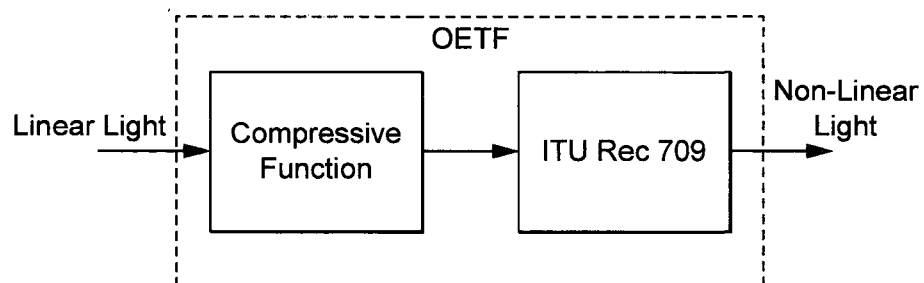


FIG. 9

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2015/050334

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04N5/202
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2013/039730 A2 (DOLBY LAB LICENSING CORP [US]; MILLER JON S [US]; NEZAMABADI MAHDI [US] 21 March 2013 (2013-03-21) paragraph [0031] - paragraphs [0046], [0053]; figures 1,2	1-13, 15-23
A	----- US 2004/196394 A1 (NAKAJIMA FUMIO [JP]) 7 October 2004 (2004-10-07) paragraph [0027] - paragraph [0053]; figures 1,2,4	14
X	----- US 2004/196394 A1 (NAKAJIMA FUMIO [JP]) 7 October 2004 (2004-10-07) paragraph [0027] - paragraph [0053]; figures 1,2,4	1-13,15, 20-23
A	----- US 8 279 308 B1 (BORG LARS U [US]) 2 October 2012 (2012-10-02)	14,16-19
X	----- US 8 279 308 B1 (BORG LARS U [US]) 2 October 2012 (2012-10-02) column 3, line 32 - column 7, line 59; figures 1,3,4,5	1,4,5,7, 8,12,15, 18-23
	----- -/-	



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

27 April 2015

Date of mailing of the international search report

12/05/2015

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Seytter, Fritz

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2015/050334

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	F. DRAGO ET AL: "Adaptive Logarithmic Mapping For Displaying High Contrast Scenes", COMPUTER GRAPHICS FORUM, vol. 22, no. 3, 1 September 2003 (2003-09-01), pages 419-426, XP055099276, ISSN: 0167-7055, DOI: 10.1111/1467-8659.00689 Sections 3.3, 4; figures 5,7 -----	23

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/GB2015/050334

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2013039730 A2	21-03-2013	US 2014341272 A1 WO 2013039730 A2	20-11-2014 21-03-2013
US 2004196394 A1	07-10-2004	JP 4368127 B2 JP 2004312269 A US 2004196394 A1	18-11-2009 04-11-2004 07-10-2004
US 8279308 B1	02-10-2012	NONE	