



(51) International Patent Classification:
G09G 5/10 (2006.01)

(21) International Application Number:
PCT/EP2017/077161

(22) International Filing Date:
24 October 2017 (24.10.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
16306421.5 28 October 2016 (28.10.2016) EP

(71) Applicant: **THOMSON LICENSING** [FR/FR]; 1-5, rue
Jeanne d'Arc, 92130 Issy-les-Moulineaux (FR).

(72) Inventors: **REINHARD, Erik**; c/o Technicolor, 975, avenue des Champs Blancs, CS 17616, 35576 CESSON-SEVIGNE (FR). **SEIFI, Mozhdeh**; c/o Technicolor, 975, avenue des Champs Blancs, CS 17616, 35576 CESSON-SEVIGNE (FR). **LE PENDU, Mikael**; c/o Technicolor, 975, avenue des Champs Blancs, CS 17616, 35576 CESSON-SEVIGNE (FR). **HAWARY, Fatma**; c/o Technicolor, 975, avenue des Champs Blancs, CS 17616, 35576 CESSON-SEVIGNE (FR).

(74) Agent: **HUCHET, Anne** et al.; Technicolor, 1-5, rue
Jeanne d'Arc, 92130 Issy-les-Moulineaux (FR).

(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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, 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).

Declarations under Rule 4.17:

— as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))

Published:

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

(54) Title: METHOD FOR ADAPTING A LUMINANCE OF A MULTIMEDIA CONTENT, CORRESPONDING COMPUTER PROGRAM PRODUCT AND APPARATUS

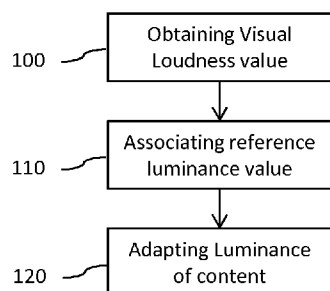


Figure 1a

(57) Abstract: A method is proposed for adapting a luminance of a multimedia content. Such method comprises, for a current pixel of an image of the multimedia content: - obtaining (100) a visual loudness value from at least a luminance value of the current pixel and from at least a luminance value of another pixel of the image; - associating (110) a reference luminance value to the visual loudness value; and - adapting (120) the luminance of the multimedia content by attributing the reference luminance value to the current pixel; the visual loudness being a measure representative of a luminance of the image with respect to a spatial configuration of distinct brightness areas in the image.

Method for adapting a luminance of a multimedia content, corresponding computer program product and apparatus

1. FIELD OF THE DISCLOSURE

The field of the disclosure is that of the image and video processing.

More specifically, the disclosure relates to a method for adapting the luminance of a video content to a reference so as to normalize the rendering.

The disclosure can be of interest in any field where the rendering of different contents originating from different sources is involved. This can be the case for instance in the field of broadcasting where channels need to support contents from various sources, with a range of formats and characteristics.

2. TECHNOLOGICAL BACKGROUND

This section is intended to introduce the reader to various aspects of art, which may be related to various aspects of the present disclosure that are described and/or claimed below. This discussion is believed to be helpful in providing the reader with background information to facilitate a better understanding of the various aspects of the present disclosure. Accordingly, it should be understood that these statements are to be read in this light, and not as admissions of prior art.

Broadcasting channels typically take contents from various sources, with a range of formats and characteristics. This includes for instance movies, live content, television shows and commercials. Here, aside from various format conversions, it is important that the look and feel between programs is consistent for a pleasant user experience.

For that, broadcasters typically strive to maintain visual consistency of the content that they transmit. This means that luminance differences between the various different types of content need to be reconciled, while maintaining the director's intent of each piece of content. There are many tone reproduction and inverse tone reproduction techniques to adjust the luminance profile of content.

However, the emergence of high dynamic range (HDR) imaging may limit the results obtained with existing technics. Indeed, various developments are observed related to HDR:

- whereas conventional standard dynamic range contents have a well-defined peak luminance, assumed to be 100 cd/m^2 (nits), the diversification of camera equipment and grading practices means that contents will have different peak luminances. Further adjustments to video content may be necessary to make content consistent with the look

and feel of the channel. Here, a balance needs to be achieved between channel consistency and maintaining the director's intent of the original content;

- currently, post-production services typically produce two different grades, one for cinema viewing in dark environments, and one for home viewing in dim environments. There is reasonable uniformity in the set-up of the mastering suites across grading studios. However, at least one third grade may be introduced to serve HDR cinema and/or high dynamic range home viewing;
- advertisers are likely to use the headroom afforded by HDR imaging systems in ways that make their content stand out. In essence, it is anticipated that advertisements will regularly appear to be significantly brighter unless it is controlled in a manner that balances the needs of the broadcasting channel, the viewers, and the advertisers;
- to achieve high luminance levels, high dynamic range display devices incorporate power management strategies to limit the power output. In practice this means that the highest achievable luminance on such displays can only be achieved for a certain proportion of the screen, possibly for only a certain amount of time.

Currently, algorithms exist that prepare content either for displays with a higher dynamic range than the content, or for displays with a lower dynamic range than the content. Often, the dynamic range of the content is matched to the dynamic range of the display, with the aid of knowledge of the dynamic range extrema. In practice this means that the black level and the peak luminance (or only the peak luminance) of both content and display are used to guide the dynamic range adjustment.

However, despite its advantages (e.g. the relative pixel values are maintained, which means that all contrasts in an image can be enhanced or reduced in substantially similar ways) it appears that such level adjustments are not sufficient if the look and feel of the content needs to be brought closer together, or needs to be matched.

There is thus a need for a method that allows adapting the look and feel of an image or a video content for matching a reference.

There is also a need for this method to be temporally stable in order to avoid introducing any flicker or artifacts when applied to a video content.

3. SUMMARY

A particular aspect of the present disclosure relates to a method for adapting a luminance of a multimedia content. Such method comprises, for a current pixel of an image of the multimedia content:

- obtaining a visual loudness value from at least a luminance value of the current pixel and from at least a luminance value of another pixel of the image;
- associating a reference luminance value to the visual loudness value; and
- adapting the luminance of the multimedia content by attributing the reference luminance value to the current pixel;

the visual loudness being a measure representative of a luminance of the image with respect to a spatial configuration of distinct brightness areas in the image.

Thus, the present disclosure proposes a new and inventive solution for adapting the luminance of a multimedia content (called a source content) that includes image or video materials to a reference (a reference image for example), for example to ensure consistency between two contents to be broadcast.

For this to be possible, a visual loudness metric representative of the perception of how light or dark an image is seen by a person viewing the content is introduced. It indeed appears that the lightness of a patch of color in an image depends on the luminance of that patch in relation to other areas of the image. It is thus proposed to introduce a visual loudness metric of the images in the content corresponding to a measure representative of the luminance of the image with respect to the spatial configuration of distinct brightness areas in it.

Based on such metric, a luminance of a pixel in the image can be adapted by attributing a reference luminance value to this pixel, this reference luminance value being associated to the visual loudness value of the pixel. In other words, the “global” luminance of the source image can be adapted so as to get an adapted image that presents a visual consistency with the selected reference.

Another aspect of the present disclosure relates to an apparatus for adapting a luminance of a multimedia content. Such apparatus comprises a memory and a processor configured for, for a current image of the multimedia content:

- obtaining a visual loudness value from at least a luminance value of the current pixel and from at least a luminance value of another pixel of the image;
- associating a reference luminance value to the visual loudness value; and
- adapting the luminance of the multimedia content by attributing the reference luminance value to the current pixel;

the visual loudness being a measure representative of a luminance of the image with respect to a spatial configuration of distinct brightness areas in the image.

Such an apparatus is particularly adapted for implementing the method for adapting a luminance of a multimedia content according to the present disclosure. Thus, the characteristics and advantages of this apparatus are the same as the disclosed method for adapting a luminance of a multimedia content.

5 According to one embodiment, the obtaining a visual loudness value comprises:

- obtaining at least two visual loudness components belonging to the group comprising:
 - a highlight visual loudness component V_h , representing an assessment of the contribution of highlight pixels in a luminance of the image with respect to a spatial configuration of areas containing the highlight pixels in the image;
 - 10 ○ a mid-tone visual loudness component V_m , representing an assessment of the contribution of mid-tone pixels in a luminance of the image with respect to a spatial configuration of areas containing the mid-tone pixels in the image; and
 - a dark visual loudness component V_d , representing an assessment of the contribution of dark pixels in a luminance of the image with respect to a spatial configuration of
 - 15 areas containing the dark pixels in the image;
- calculating the visual loudness value as a function of the at least two visual loudness components.

Indeed, the three groups of pixels (i.e. highlight, mid-tone and dark pixels) contribute to the sensation of visual loudness in different ways. Taking into account for their respective contribution separately allows achieving a suitable visual consistency between contents (for

20 example a source content and a reference content).

According to one embodiment, the reference luminance value being associated to the visual loudness value is obtained from a luminance value of a pixel in a reference image.

Thus, the luminance of the image in the multimedia content can be adapted so as to

25 present a visual consistency with a reference image, e.g. in a reference content.

According to one embodiment, the method further comprises, or the apparatus is further configured for, averaging the reference luminance value with the luminance value of the current pixel, providing a modified luminance value, the adapting the luminance of the multimedia content corresponding to attributing the modified luminance value to the current

30 pixel.

Thus, only a partial adaptation of the luminance of the content is performed so that a smooth adaptation of the source content is obtained.

According to one embodiment, the obtaining at least two visual loudness components comprises calculating at least one threshold belonging to the group comprising:

- an upper threshold T_{mh} ; and
- a lower threshold T_{dm} ;

5 the highlight pixels corresponding to pixels with a luminance above the upper threshold T_{mh} , the dark pixels corresponding to pixels with a luminance below the lower threshold T_{dm} and the mid-tone pixels corresponding to pixels with a luminance below the upper threshold T_{mh} and above the lower threshold T_{dm} .

Thus, the three groups of pixels are defined in a simple and robust way.

10 According to one embodiment, the obtaining at least two visual loudness components comprises deriving at least two corresponding characteristic luminance values belonging to the group comprising:

- a characteristic highlight luminance value L_h representative of a luminance of lighter pixels in the image;
- 15 • a characteristic mid-tone luminance value L_m representative of an average luminance of pixels in the image; and
- a characteristic dark luminance value L_d representative of a luminance of darker pixels in the image;

the upper threshold T_{mh} being a function of the characteristic highlight luminance value L_h and
 20 of the characteristic mid-tone luminance value L_m , and the lower threshold T_{dm} being a function of the characteristic mid-tone luminance value L_m and of the characteristic dark luminance value L_d .

Thus, the characteristic luminance of the three groups of pixels is defined in a simple and robust way through such characteristic luminance value.

25 According to one embodiment, the characteristic highlight luminance value L_h , respectively characteristic dark luminance value L_d , corresponds to a first percentage, respectively second percentage, of pixels in the image having a luminance below the characteristic highlight luminance value L_h , respectively characteristic dark luminance value L_d , the first percentage being greater than the second percentage.

30 According to one embodiment, the characteristic mid-tones luminance value L_m corresponds to a geometric mean luminance of pixels in the image.

Thus, the characteristic highlights, mid-tones and darks luminance values are derived based on robust statistics, i.e. on statistics that are not sensitive to outliers (such as minimum and maximum values in an image).

According to one embodiment, the visual loudness value is extrapolated from the at least two visual loudness components and from the at least two corresponding characteristic luminance values.

Thus, a visual loudness value can be derived in a simple and robust way for each pixel in the considered image, based only on the visual loudness components associated to the three groups of pixels and on the corresponding characteristic luminance values.

According to one embodiment, the obtaining at least two visual loudness components comprises:

- determining a mask M_h representative of a location of the highlight pixels in the image; and
- splitting the mask M_h when a density of the highlight pixels in the mask M_h is below a highlight pixels density threshold, the splitting providing at least one sub-mask presenting a higher density of highlight pixels than the mask M_h ;

the highlight visual loudness component V_h being a function of a density of highlight pixels of the at least one sub-mask and of a size of the at least one sub-mask.

Thus, areas of high density of lighter pixels can be obtained in order to derive a highlight visual loudness component that takes into account for the spatial configuration of those areas.

According to one embodiment, the function is a non-linear function of the size of the at least one sub-mask.

Thus, the inability of certain HDR displays to reach certain luminance levels for regions of a certain size can be taken into account in the determination of the highlight visual loudness component.

According to one embodiment, the splitting is enforced hierarchically and recursively on previously obtained sub-masks according to a quad-tree structure.

According to one embodiment, the mid-tone visual loudness component V_m is a function of a barycenter of a luminance of the mid-tone pixels with respect to the characteristic highlight L_h and/or dark L_d luminance values.

According to another embodiment, the dark visual loudness component V_d is a function of a barycenter of a luminance of the dark pixels with respect to the characteristic dark luminance value L_d and/or of a zero luminance reference value.

Thus, the mid-tone and dark visual loudness components can be derived in a simple and straightforward manner from the characteristic highlight and dark luminance values.

According to one embodiment, the function of a barycenter is a non-linear function.

Thus, the contribution of the mid-tone and dark visual loudness components to the overall visual loudness of the image can be controlled more accurately.

According to one embodiment, the multimedia content is a video content, and the obtaining a visual loudness value further comprises adjusting visual loudness components for providing at least two adjusted visual loudness components as a function of the at least two visual loudness components and of at least two previous visual loudness components obtained for a previous image in the video content, the visual loudness value being obtained from the at least two adjusted visual loudness components.

Thus, the disclosed method is temporally stable so that no flicker or other unwanted artefacts are introduced in the video content.

Furthermore, when a sharp transition occurs in the content, or during the time immediately following such a sharp transition (where for instance a transition may be due to switching between a movie and a commercial and back), the method according to this embodiment allows a smooth adaptation of the content so as to follow the speed of human visual adaptation for instance.

Another aspect of the present disclosure relates to a computer program product comprising program code instructions for implementing the above-mentioned method for adapting a luminance of a multimedia content (in any of its different embodiments), when the program is executed on a computer or a processor.

Another aspect of the present disclosure relates to a non-transitory computer-readable carrier medium storing a computer program product which, when executed by a computer or a processor causes the computer or the processor to carry out the above-mentioned method for adapting a luminance of a multimedia content (in any of its different embodiments).

4. LIST OF FIGURES

Other features and advantages of embodiments shall appear from the following description, given by way of indicative and non-exhaustive examples and from the appended drawings, of which:

- Figures 1a and 1b are flowcharts of particular embodiments of the disclosed method for adapting a luminance of a multimedia content;
- Figure 2 is a schematic illustration of the structural blocks of an exemplary apparatus that can be used for implementing the method for adapting a luminance of a multimedia content according to any of the embodiments disclosed in relation with figures 1a and 1b.

5. DETAILED DESCRIPTION

In all of the figures of the present document, the same numerical reference signs designate similar elements and steps.

The general principle of the disclosed method consists in obtaining a visual loudness value for a multimedia content, the visual loudness value representing a measure of how light or dark an image appears with respect to a spatial configuration of distinct brightness areas in the image, in order to adjust the multimedia content for matching the visual loudness of a reference multimedia content.

For that, the disclosed method calculates a visual loudness value from at least a luminance value of a current pixel in an image of the multimedia content (including image or video materials) and of at least a luminance value of another pixel in the same image. Then, the disclosed method associates a reference luminance value to the obtained visual loudness value, and adaptes the luminance value of the multimedia content by attributing the reference luminance value to the current pixel.

Referring now to **figures 1a and 1b**, we present different embodiments of a method for adapting a luminance of a multimedia content.

In block 100, a visual loudness value is obtained, for a multimedia content, from at least a luminance of a current pixel in an image of the multimedia content and of at least a luminance value of another pixel in the same image.

According to the present disclosure, a visual loudness, representative of how light or dark an image appears with respect to a spatial configuration of distinct brightness areas in the image, is expressed as a function of components among three visual loudness components belonging to the group comprising:

- a highlight visual loudness component V_h , representing an assessment of the contribution of highlight pixels in a luminance of the image with respect to a spatial configuration of areas containing the highlight pixels in the image;
- a mid-tone visual loudness component V_m , representing an assessment of the contribution of mid-tone pixels in a luminance of the image with respect to a spatial configuration of areas containing the mid-tone pixels in the image; and
- a dark visual loudness component V_d , representing an assessment of the contribution of dark pixels in a luminance of the image with respect to a spatial configuration of areas containing the dark pixels in the image.

The motivation for this comes from the observation that dark pixels, mid-tone pixels and highlight pixels each contribute to the sensation of visual loudness in different ways.

Indeed, it appears that lightness is a perceptual measure, which can be defined as the attribute of visual sensation according to which the area in which the visual stimulus is presented appears to emit more or less light in proportion to that emitted by a similarly illuminated area perceived as the “white” stimulus (see “Reinhard, E., Khan, E. A., Akyuz, A. O., & Johnson, G. (2008). **“Color imaging: fundamentals and applications”**, CRC Press”).

It was recently found that the lightness of a uniform patch of color depends on the luminance of that patch in relation to other regions, as well as its size (see “Allred, S. R., Radonjić, A., Gilchrist, A. L., & Brainard, D. H. (2012). **“Lightness perception in high dynamic range images: Local and remote luminance effects”**. *Journal of vision*, 12(2), 7”). This is known as the “area rule”, formalized by Alan Gilchrist as: “In a simple display, when the darker of the two regions has the greater relative area, as the darker region grows in area, its lightness value goes up in direct proportion. At the same time the lighter region first appears white, then a fluorescent white and finally, self-luminous.” (see “Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., ... & Economou, E. (1999). **“An anchoring theory of lightness perception”**. *Psychological review*, 106(4), 795”).

In the definition of visual loudness, a variety of computational lightness models may be used, ideally but not necessarily including a measure of size. However, within the context of visual loudness, computational measures of lightness evaluate high luminance pixels, as opposed to medium and low luminance pixels

It moreover appears that the overall impression of a frame or image is in part determined by the luminance level of the mid-tones, in relation to the dark and light parts of

the frame or image. A mid-tones visual loudness component can thus be assessed by deriving measures from the analysis of the histogram of an image.

Last, it also appears that the extent of the dark parts of the frame or image also impacts the impression of visual loudness. However, the magnitude of visual loudness will be inversely proportional to the percentage of pixels that can be considered dark.

Consequently, in one embodiment of the present disclosure, the highlight visual loudness component V_h component is following modern insights into the perception of lightness to assess the contribution of the lightest parts of the image. The mid-tones visual loudness component V_m constitutes an analysis of the histogram of the (possibly down-sampled) image to evaluate the contribution of mid-tones. The darks visual loudness component V_d results from an analysis of the spatial extent of the darkest parts of the scene.

In one exemplary embodiment, the overall measure V of visual loudness of the considered image is the product of the three contributions:

$$V = V_h V_m V_d$$

In other embodiments, a visual loudness calculation is calculated that omits one or two of the components. This can be formalized by setting $V_h = 1$, $V_m = 1$ and/or $V_d = 1$ in the above equation.

In yet another embodiment, the visual loudness may be represented by pairs of numbers, e.g. three pairs of numbers characterizing the visual loudness contributions of the dark, mid-tone and highlight regions in the considered picture. In that case, these pairs of numbers are

$$\begin{pmatrix} L_d, V_d \\ L_m, V_m \\ L_h, V_h \end{pmatrix} \quad (\text{Eq-1})$$

where:

- L_h is a characteristic highlight luminance value representative of a luminance of lighter pixels in the image;
- L_m is a characteristic mid-tone luminance value representative of an average luminance of pixels in the image; and
- L_d is a characteristic dark luminance value representative of a luminance of darker pixels in the image.

Based on at least two of those three pairs of numbers, or control points, a curve mapping any luminance value to a visual loudness contribution can be obtained, so that a notion of visual loudness can be assigned to each luminance value in the image. In other

words, a visual loudness value is extrapolated from the at least two visual loudness components and from the at least two corresponding characteristic luminance values and can be assigned to each luminance value in the image. A visual loudness value can thus be derived in a simple and robust way for each pixel in the considered image, based only on the visual loudness components associated to the three groups of pixels and on the corresponding characteristic luminance values.

There are many possible ways to define curves according to control points, including but not limited to splines such as B-splines or non-uniform rational basis splines (NURBS), piecewise linear curve segments and polynomials. Further, such curves can be represented for instance with 1D Look-Up Tables (LUTs).

We now disclose in relation with block 100a (figure 1b) example derivations for obtaining at least two visual loudness components belonging to the group comprising a highlight visual loudness component V_h , a mid-tone visual loudness component V_m , and a dark visual loudness component V_d .

For that, in block 100a1 (figure 1b), the characteristic highlight luminance value L_h , the characteristic midtone luminance value L_m and the characteristic dark luminance value L_d are derived.

In one embodiment, the characteristic mid-tones luminance value L_m corresponds to a geometric mean luminance of pixels in the image. The geometric mean luminance L_m is a measure of where the majority of the luminance of pixels typically is located within the dynamic range of the image and can be calculated as follows:

$$L_m = \exp \left(\frac{1}{w \cdot h} \sum_{x=1}^w \sum_{y=1}^h \ln(L(x, y)) \right)$$

where w and h are the width and the height of the image and $L(x, y)$ is the luminance of the pixel located at position (x, y) in the image.

In another embodiment, this is the arithmetic mean luminance that is used for deriving L_m , yet leading to simpler calculations.

Conversely, the characteristic highlight luminance value L_h , and the characteristic dark luminance value L_d , may be computed as a percentile, i.e. the luminance value for which $N\%$ of the pixels are below that luminance value. In other words, the characteristic highlight luminance value L_h , respectively characteristic dark luminance value L_d , corresponds to a first percentage, respectively second percentage, of pixels in the image having a luminance below

the characteristic highlight luminance value L_h , respectively characteristic dark luminance value L_d , the first percentage being greater than the second percentage.

For instance, the characteristic dark luminance value L_d may be characterized by setting the value of N to 0%, 1%, 5% or 10%, or any other appropriate value that suitably separates dark luminance values from all other luminance values. In the same way, the characteristic highlight luminance value L_h may be characterized by setting the value of N to a high value, such as 90%, 95%, 99% or 100%.

Consequently, the characteristic highlight, mid-tone and dark luminance values are derived based on robust statistics, i.e. on statistics that are not sensitive to outliers (such as minimum and maximum values in an image).

In block 100a2 (figure 1b), at least one threshold belonging to the group comprising an upper threshold T_{mh} and a lower threshold T_{dm} is calculated.

More precisely, those upper and lower thresholds allow defining the highlight pixels as the pixels having a luminance above the upper threshold T_{mh} , the dark pixels as the pixels with a luminance below the lower threshold T_{dm} , and the mid-tone pixels as the pixels with a luminance below the upper threshold T_{mh} and above the lower threshold T_{dm} .

With such definition, the three groups of pixels are defined in a simple and robust way.

Having that the three characteristic luminance values L_d , L_m and L_h (introduced above in relation with equation (Eq-1)) are representative of luminance values located in dark, mid, and high luminance regions, in one embodiment:

- the upper threshold T_{mh} is calculated as a function of the characteristic highlight luminance value L_h and of the characteristic mid-tone luminance value L_m ; and
- the lower threshold T_{dm} is calculated as a function of the characteristic mid-tone luminance value L_m and of the characteristic dark luminance value L_d .

For example, the lower threshold T_{dm} between dark pixels and mid-tone pixels may be computed as an average between L_d and L_m according to:

$$T_{dm} = \alpha L_d + (1 - \alpha) L_m$$

where α interpolates between the two luminance values. For a value of $\alpha = 0.5$, for instance, the average between L_d and L_m is computed. The upper threshold T_{mh} between mid-tone pixels and highlight pixels may be computed analogously following:

$$T_{mh} = \beta L_m + (1 - \beta) L_h$$

where β interpolates between the two luminance values L_m and L_h as does α for L_d and L_m .

In another example, the threshold T_{dm} is computed by finding the minimum value of the image's luminance histogram between L_d and L_m . The threshold T_{mh} can be computed by finding the minimum value of the image's luminance histogram between L_m and L_h .

Back to block 100a (figure 1b), the highlight visual loudness component V_h is obtained as a function of all the pixels with a luminance value larger than T_{mh} , i.e.:

$$V_h = f_h(S_h), \quad \text{where } S_h = \{L_i \mid L_i > T_{mh}\}$$

In other words, V_h is a function $f_h()$ of the set of pixels S_h that contains all pixels that have a luminance larger than the threshold T_{mh} .

In that perspective, it is proposed to follow recent insights into lightness perception. In models of lightness perception, uniform patches with a given luminance are assessed with respect to both size and level. In particular, larger patches with a given luminance are perceived to be less light relative to smaller patches with the same luminance. This means that to be able to assess lightness, it is necessary to assess the pixels in the set S_h with respect to how they are grouped. In essence, if the highlight pixels in an image form many small clusters, then their perception will be different than if these highlight pixels form a small collection of much larger clusters.

There are several approaches possible to determine clusters for a collection of pixels. In exemplary embodiments, the k -means clustering algorithm (or one of its heuristic variants) could be used to partition an image into k clusters. This algorithm is also known as Lloyd's algorithm. Other approaches include the use of Gaussian Mixture models that are trained with the Expectation-Maximization algorithm, k -medians clustering, Fuzzy C-Means Clustering and k -means++. In these algorithms the number of clusters k needs to be known *a priori*. As in the present case, this number is not known, it could be estimated with a separate algorithm, or using a rule of thumb. Further, these methods do not provide information regarding the size of each cluster.

Suitable algorithms that do not require the number of clusters to be known *a priori* include Density-Based Spatial Clustering of Applications with Noise (DBSCAN), Ordering Points To Identify the Clustering Structure (OPTICS), and hierarchical clustering such as SLINK (for "single-link" clustering method) and CLINK (for "complete link" clustering method).

As the specific application of highlight assessment does not require high accuracy in the detection of clusters, but does require knowledge of the size of such cluster, a different and straightforward approach is described which may serve as an exemplary implementation.

For that, a mask M_h representative of a location of the highlight pixels in the considered image is determined in block 100a3 (figure 1b).

In an exemplary embodiment, the mask M_h is created with the same size as the image, with values set to "0" or "1" dependent on whether the corresponding pixels belong to the set

5 S_h or not, i.e.:

$$M_h(x, y) = \begin{cases} 1 & \text{if } L(x, y) > T_{mh} \\ 0 & \text{otherwise} \end{cases}$$

In block 100a4 (figure 1b), the mask M_h is split when a density of highlight pixels in the mask M_h is below a highlight pixels density threshold t for providing at least one sub-mask presenting a higher density of highlight pixels than the mask M_h .

10 For instance, denoting respectively w and h the width and the height of the image, the mask M_h is split into four smaller masks, each of size $w/2$ by $h/2$ if the number of pixels corresponding to elements with a value of "1" in the mask is lower than the highlight pixels density threshold t , i.e. if the following inequalities hold true:

$$p = \frac{\sum_w \sum_h M_h(x, y)}{w h} < t$$

15 In one embodiment, the highlight pixels density threshold t is suitably chosen in the range $(t_{\text{low}}, t_{\text{high}})$. An example value of t_{low} may be 0.2, whereas a typical example value of t_{high} may be 0.8 for achieving good results at a reasonable computation cost.

In another embodiment, the mask M_h is split either iteratively or recursively into four equal parts, provided the splitting criterion is satisfied. In other words, the splitting is enforced
20 hierarchically and recursively on previously obtained sub-masks according to a quad-tree structure.

In this embodiment, thresholds on the width and height of the (sub-)masks ensure that the depth of the quad-tree stays within reasonable bounds. Example values for these thresholds may be $t_w = t_h = 8$.

25 The resulting quad-tree stores at every node the calculated percentage p . Each node q is located at a given level in the quad-tree and represents a sub-mask. The size s of a leaf node (i.e. the size of the corresponding sub-mask) relates directly to the level l of that node in the quad-tree, according to $s = (w h)4^{-l}$ (assuming that the root node is at level $l = 0$). The height of the quad-tree (i.e. the maximum level) is limited by the two thresholds on the width and height, t_w and t_h respectively, discussed above.
30

Considering that the maximum level in the quad-tree is given by l_{max} , the calculation then proceeds by analyzing all leaf nodes located at levels $0 \leq l < l_{\text{max}}$. These leaf nodes

were not subdivided further, because the stopping criterion $p \geq t$ was reached. In other words, these leaf nodes contain a substantial number of highlight pixels. The set of leaf nodes between levels $0 \leq l < l_{max}$ is denoted $Q = \{q_1, \dots, q_n\}$, where n is the number of leaf nodes in this set. Each leaf node q_i has a size s_i and a percentage of pixels p_i that correspond to a value of “1” in the mask M_h . The size s_i depends on the level l at which this node is located as discussed above.

The visual loudness v_i of each node q_i is determined as follows:

$$v_i = p_i h(s_i)$$

where the function $h()$ maps the size of each leaf node to a scaling factor that is applied to p_i .

This function is typically monotonically increasing. In its simplest form, it is the identity function, $h(s_i) = s_i$, but other functions can be considered.

In one exemplary embodiment, the function $h()$ is a non-linear function of the size s_i of the sub-mask represented by the leaf node q_i . This allows taking into account possible power management issues that could arise in certain display devices. This function then models the inability of a display to reach certain luminance levels for regions of a certain size.

Furthermore, following the findings in lightness perception disclosed in “*Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., ... & Economou, E. (1999). “An anchoring theory of lightness perception”. Psychological review, 106(4), 795*”, it appears that the visual loudness increases as the size of the region s_i increases. It also appears that the visual loudness increases with the number of highlight pixels in each region s_i .

Consequently, in an exemplary embodiment of the present disclosure, the highlight visual loudness component V_h induced by the highlight pixels is obtained as the sum of each region’s contribution normalized by the sum of the region’s sizes s_i involved in the calculation, i.e.:

$$V_h = f_h(S_h) = \frac{\sum_{i=1}^n v_i}{\sum_{i=1}^n s_i}$$

Back to block 100a, the calculation of the mid-tone visual loudness component V_m is obtained as a function of all the pixels with a luminance value larger than T_{dm} , but lower than T_{mh} , i.e.:

$$V_m = f_m(S_m), \quad \text{where } S_m = \{L_i \mid T_{dm} < L_i < T_{mh}\}$$

In general, the luminance histogram of images provides clues as to whether the scene was overall light, medium or dark, irrespective of whether the luminance values of the image signify absolute or relative values. The luminance histogram of a high dynamic range image

can be interpreted as a probability density function. In that case, high dynamic range image histograms tend to deviate from a Gaussian distribution in that they often exhibit high values for their third and fourth moments (skew and kurtosis). In particular, the kurtosis represents the fact that the distribution is long-tailed. It would therefore be possible to analyze the histogram for skew and kurtosis, and make decisions as to whether the image is overall light or dark.

However, in particular embodiments of the present disclosure, the mid-tone visual loudness component V_m is obtained in a more straightforward manner using the characteristic luminance values.

In one embodiment, the mid-tone visual loudness component V_m is obtained as a function of a ratio r expressed in a form of a barycenter as:

$$r = \frac{1}{|S_m|} \sum_{L_i \in S_m} \frac{L_i - L_d}{L_h - L_d}$$

Small values of r indicate that the mid-tones of the image are relatively close to the dark values. This indicates that the image may be dark overall. On the other hand, large values of r indicate the reverse, i.e. that the image may be light overall.

In another embodiment, the ratio r may be defined relative to the characteristic highlight luminance value L_h instead of the characteristic dark luminance value L_d , i.e.:

$$r = \frac{1}{|S_m|} \sum_{L_i \in S_m} \frac{L_i - L_h}{L_h - L_d}$$

Small values of r then indicate that the mid-tones of the image are relatively close to the highlight values. This indicates that the image may be light overall. On the other hand, large values of r indicate the reverse, i.e. that the image may be dark overall.

In yet another embodiment, the ratio r may be defined as a function of a ratio r expressed in a form of a barycenter relative to both the characteristic highlight luminance value L_h and the characteristic dark luminance value L_d .

Whatever the considered embodiment discussed above, the mid-tone visual loudness component V_m is obtained as a function of r , i.e. as a function of a barycenter of a luminance of the mid-tone pixels with respect to the characteristic highlight L_h and/or dark L_d luminance values, i.e. in a very simple and straightforward manner from those characteristic values.

In other embodiments, the ratio r is defined relative of the upper threshold T_{mh} and/or of the lower threshold T_{dm} instead of characteristic highlight L_h and/or dark L_d luminance values.

In another embodiment, the three characteristic luminance values L_d , L_m and L_h are used to assess the contribution of the mid-tones to visual loudness by computing the ratio r as:

$$r = \frac{L_m - L_d}{L_h - L_d}$$

5 In yet another embodiment, the ratio r , which may be computed according to any of the embodiment disclosed above, may be normalized. The normalization may be necessary as pixels with a luminance L_i in the mid-tone luminance region are defined to have values in the range $[T_{dm}, T_{mh}]$, with $L_d < T_{dm}$ and $L_h > T_{mh}$. This means that without normalization r ranges between $n_{low} = \frac{T_{dm} - L_d}{L_h - L_d} > 0$ and $n_{high} = \frac{T_{mh} - L_d}{L_h - L_d} < 1$. The normalization of r is achieved

10 as follows:

$$r_{norm} = \frac{r - n_{low}}{n_{high} - n_{low}}$$

Further, the ratio r_{norm} may be weighted according to a non-linear function, which may be any function that takes a normalized scalar as input and produces a scalar as output. This includes, but is not limited to power functions, logarithmic functions, exponential functions, 15 trigonometric and sigmoidal functions. The advantage of such weighting is that the contribution to overall visual loudness of the mid-tone component can be better controlled. In that case, the ratio r_{norm} weighted according to such non-linear function may be directly interpreted as the mid-tone visual loudness component V_m .

In the same way, the calculation of the dark visual loudness component V_d is obtained 20 as a function of all the pixels with a luminance value lower than T_{dm} , i.e.:

$$V_d = f_d(S_d), \quad \text{where } S_d = \{L_i \mid L_i < T_{dm}\}$$

In one embodiment, the dark visual loudness component V_d is obtained as a function of a ratio r_d expressed in a form of a barycenter referred to a zero luminance reference value as:

25

$$r_d = \frac{1}{|S_d|} \sum_{L_i \in S_d} \frac{L_i}{T_{dm}}$$

In alternative embodiments, the ratio r_d is expressed in a form of a barycenter referred to the characteristic dark luminance value L_d instead of the zero luminance reference value.

In the same way, r_d may be normalized to another quantity than the lower threshold T_{dm} (e.g. the characteristic dark luminance value L_d). The dark visual loudness component V_d 30 is thus obtained as a function of a barycenter of the luminance of the dark pixels with respect to the characteristic dark luminance value L_d and/or of a zero luminance reference value, i.e. in a very simple and straightforward manner from those characteristic values.

As for the ratio r_{norm} , r_d may be weighted according to a non-linear function, which may be any function that takes a normalized scalar as input and produces a scalar as output. This includes, but is not limited to power functions, logarithmic functions, exponential functions, trigonometric and sigmoidal functions. The advantage of such weighting is that the contribution to overall visual loudness of the dark component can be better controlled. In that case, the ratio r_d weighted according to such non-linear function may be directly interpreted as the dark visual loudness component V_d .

In some embodiments, the multimedia content includes a video content, and in block 100b (figure 1b), the at least two visual loudness components obtained in block 100a are further adjusted for providing at least two adjusted visual loudness components as a function of the at least two visual loudness components and of at least two previous visual loudness components obtained for a previous image in the video content.

In one embodiment, a leaky integration is applied to the visual loudness components V_d , V_m and V_h and/or to the characteristic luminance values L_d , L_m and L_h obtained for successive images in the video content.

Thus, the disclosed method is temporally stable so that no flicker or other unwanted artefacts are introduced in the video content.

Furthermore, when a sharp transition occurs in the content, or during the time immediately following such a sharp transition (where for instance a transition may be due to switching between a movie and a commercial and back), the method according to this embodiment allows a smooth adaptation of the content so as to follow the speed of human visual adaptation for instance.

For instance, these parameters may be adjusted such that the overall visual loudness V follows a temporal Cornsweet profile (see "*Cornsweet, T. (2012). 'Visual perception'. Academic press*"). This could be used to increase the visual appearance of temporal contrast, or alternatively to decrease the visual appearance of temporal contrast.

In block 100c (figure 1b), the visual loudness value to be obtained from at least the luminance of the current pixel (i.e. the output of block 100) is calculated as a function of the at least two visual loudness components obtained in block 100a.

However, in the embodiments discussed above in relation with block 100b, the multimedia content includes a video content, and the visual loudness value to be obtained is calculated in block 100c as a function of the at least two adjusted visual loudness components obtained in block 100b.

In any case, whatever the nature of the multimedia content (video or single image), the visual loudness value may be obtained according to the method disclosed above in relation with block 100 and equation (Eq-1) when introducing the representation of visual loudness as pairs of numbers. It means that a curve representation $V_{\text{image}}(x, y) = c_{\text{image}}(L_{\text{image}}(x, y))$ can be derived using for instance an interpolation process based on the at least two pairs of numbers derived from the visual loudness components obtained in block 100a or in block 100c.

The visual loudness value is thus obtained as extrapolated from the at least two visual loudness components (or two adjusted visual loudness components) and from the at least two corresponding characteristic luminance values, based on the luminance $L_{\text{image}}(x, y)$ of the current pixel located at position (x, y) .

In block 110 (figures 1a and 1b), a reference luminance value is associated to the visual loudness value obtained in block 100.

In one embodiment, the reference luminance values are stored in a look-up table with entries corresponding to all possible visual loudness values.

In another embodiment, a reference source of content may be created that has a desirable level of visual loudness. The visual loudness for this reference content is represented by at least two among the three pairs of characteristic luminance values / visual loudness components:

$$\begin{aligned} &(L_{d,\text{ref}}, V_{d,\text{ref}}) \\ &(L_{m,\text{ref}}, V_{m,\text{ref}}) \\ &(L_{h,\text{ref}}, V_{h,\text{ref}}) \end{aligned}$$

From these values a curve representation $V_{\text{ref}} = c_{\text{ref}}(L_{\text{ref}})$ can be derived using the interpolation process disclosed above in relation with block 100 and equation (Eq-1). The function c_{ref} should be monotonically increasing. It is envisaged that the desired visual loudness for an HDR broadcast channel is encoded with this function, and should be applied to all program material to be broadcast over this channel. It is therefore anticipated that it is possible to construct a reference visual loudness presentation that, if represented by a curve, yields a monotonically increasing curve. This reference curve must be monotonically increasing so that it can be inverted.

Reconsidering the similar representation as constructed for the considered current image in block 100c, namely $V_{\text{image}}(x, y) = c_{\text{image}}(L_{\text{image}}(x, y))$ (This representation may or may not constitute a monotonically increasing curve), a reference luminance value belonging

to the reference source of content is associated to the visual loudness $V_{\text{image}}(x, y) = c_{\text{image}}(L_{\text{image}}(x, y))$ according to:

$$L_{\text{ref}}(L_{\text{image}}(x, y)) = c_{\text{ref}}^{-1}(c_{\text{image}}(L_{\text{image}}(x, y)))$$

In this embodiment, the reference luminance value being associated to the visual loudness value associated to the current pixel is obtained from a luminance value of a pixel in a reference image.

Thus, the luminance of the considered image in the multimedia content can be adapted so as to present a visual consistency with a reference image in a reference source of content.

In block 130 (figure 1b), the reference luminance value $L_{\text{ref}}(L_{\text{image}}(x, y))$ is averaged with the luminance value $L_{\text{image}}(x, y)$ of the current pixel in the considered image, providing a modified luminance value.

Indeed, in certain applications it may be desirable to affect a partial match of visual loudness, rather than a full match between source and reference content. This can be achieved by means of averaging the reference luminance value $L_{\text{ref}}(L_{\text{image}}(x, y))$ and the luminance $L_{\text{image}}(x, y)$ of the current pixel located at position (x, y) in the considered image, for providing a modified luminance value $L'_{\text{ref}}(L_{\text{image}}(x, y))$ given by:

$$L'_{\text{ref}}(L_{\text{image}}(x, y)) = \alpha L_{\text{ref}}(L_{\text{image}}(x, y)) + \beta L_{\text{image}}(x, y)$$

where α and β determine the relative contributions of matched and unmatched luminance values of the considered image. In typical applications, the sum of the weights equal 1, i.e.

$$\alpha + \beta = 1.$$

In block 120 (figures 1a and 1b), the luminance of the multimedia content is adapted by attributing the reference luminance value obtained in block 110 to the current pixel.

Consequently, the “global” luminance of the source image can be adapted so as to get an adapted image that presents a visual consistency with the selected reference.

In the embodiment where the reference luminance value $L_{\text{ref}}(L_{\text{image}}(x, y))$ is averaged with the luminance value $L_{\text{image}}(x, y)$ of the current pixel in the considered image for providing a modified luminance value in block 130, this is the modified luminance value $L'_{\text{ref}}(L_{\text{image}}(x, y))$ that is attributed to the current pixel. In that case, only a partial adaptation of the luminance of the content is performed so that a smooth adaptation of the source content is obtained.

Referring now to **figure 2**, we illustrate the structural blocks of an exemplary apparatus that can be used for implementing the method for adapting a luminance of a multimedia content according to any of the embodiments disclosed above in relation with figures 1a and 1b.

In an embodiment, an apparatus 200 for implementing the disclosed method comprises a non-volatile memory 201 (e.g. a read-only memory (ROM) or a hard disk), a volatile memory 203 (e.g. a random access memory or RAM) and a processor 202. The non-volatile memory 201 is a non-transitory computer-readable carrier medium. It stores executable program code instructions, which are executed by the processor 202 in order to enable implementation of the method described above (method for adapting a luminance of a multimedia content) in its various embodiment disclosed in relationship with figures 1a and 1b. Non-volatile memory 201 and volatile memory 203 may be combined into a single entity referred to as memory, configured as combined volatile and non-volatile memory, or just one of the types of memory.

Upon initialization, the aforementioned program code instructions are transferred from the non-volatile memory 201 to the volatile memory 203 so as to be executed by the processor 202. The volatile memory 203 likewise includes registers for storing the variables and parameters required for this execution.

All the steps of the above method for adapting a luminance of a multimedia content may be implemented equally well:

- by the execution of a set of program code instructions executed by a reprogrammable computing machine such as a PC type apparatus, a DSP (digital signal processor) or a microcontroller. This program code instructions can be stored in a non-transitory computer-readable carrier medium that is detachable (for example a floppy disk, a CD-ROM or a DVD-ROM) or non-detachable; or

- by a dedicated machine or component, such as an FPGA (Field Programmable Gate Array), an ASIC (Application-Specific Integrated Circuit) or any dedicated hardware component.

In other words, the disclosure is not limited to a purely software-based implementation, in the form of computer program instructions, but that it may also be implemented in hardware form or any form combining a hardware portion and a software portion.

CLAIMS

1. A method for adapting a luminance of a multimedia content, comprising, for a current pixel of an image of said multimedia content:

- 5 • **obtaining** (100) a visual loudness value from at least a luminance value of said current pixel and from at least a luminance value of another pixel of said image;
- **associating** (110) a reference luminance value to said visual loudness value; and
- **adapting** (120) said luminance of said multimedia content by attributing said reference luminance value to said current pixel;

10 said visual loudness being a measure representative of a luminance of said image with respect to a spatial configuration of distinct brightness areas in said image.

2. An apparatus for adapting a luminance of a multimedia content, comprising:

- a memory; and

15 - a processor (202) configured for, for a current pixel of an image of said multimedia content:

- obtaining a visual loudness value from at least a luminance value of said current pixel and from at least a luminance value of another pixel of said image;
 - associating a reference luminance value to said visual loudness value; and
 - adapting said luminance of said multimedia content by attributing said reference
- 20 luminance value to said current pixel;

said visual loudness being a measure representative of a luminance of said image with respect to a spatial configuration of distinct brightness areas in said image.

3. A method according to claim 1 or an apparatus according to claim 2,

25 wherein said **obtaining** (100) a visual loudness value comprises:

- **obtaining** (100a) at least two visual loudness components belonging to the group comprising:
 - a highlight visual loudness component V_h , representing an assessment of the contribution of highlight pixels in a luminance of said image with respect to a spatial configuration of areas containing said highlight pixels in said image;
 - a mid-tone visual loudness component V_m , representing an assessment of the contribution of mid-tone pixels in a luminance of said image with respect to a spatial configuration of areas containing said mid-tone pixels in said image; and
- 30

- a dark visual loudness component V_d , representing an assessment of the contribution of dark pixels in a luminance of said image with respect to a spatial configuration of areas containing said dark pixels in said image;
- calculating (100c) said visual loudness value as a function of said at least two visual loudness components.

4. A method according to claim 1 or 3, or an apparatus according to claim 2 or 3, wherein said reference luminance value being associated to said visual loudness value is obtained from a luminance value of a pixel in a reference image.

5. A method according to any of claims 1, 3 or 4 further comprising, or an apparatus according to any of claims 2 to 4 further configured for, **averaging** (130) said reference luminance value with said luminance value of said current pixel, providing a modified luminance value,

said **adapting** (120) said luminance of said multimedia content corresponding to attributing said modified luminance value to said current pixel.

6. A method or an apparatus according to claim 3 or 4, wherein said **obtaining** (100a) at least two visual loudness components comprises **calculating** (100a1) at least one threshold belonging to the group comprising:

- an upper threshold T_{mh} ; and
- a lower threshold T_{dm} ;

said highlight pixels corresponding to pixels with a luminance above said upper threshold T_{mh} , said dark pixels corresponding to pixels with a luminance below said lower threshold T_{dm} and said mid-tone pixels corresponding to pixels with a luminance below said upper threshold T_{mh} and above said lower threshold T_{dm} .

7. A method or an apparatus according to claim 6, wherein said **obtaining** (100a) at least two visual loudness components comprises **deriving** (100a2) at least two corresponding characteristic luminance values belonging to the group comprising:

- a characteristic highlight luminance value L_h representative of a luminance of lighter pixels in said image;

- a characteristic mid-tone luminance value L_m representative of an average luminance of pixels in said image; and
- a characteristic dark luminance value L_d representative of a luminance of darker pixels in said image;

5 said upper threshold T_{mh} being a function of said characteristic highlight luminance value L_h and of said characteristic mid-tone luminance value L_m , and
said lower threshold T_{dm} being a function of said characteristic mid-tone luminance value L_m and of said characteristic dark luminance value L_d .

10 **8.** A method or an apparatus according to claim 7,
wherein said visual loudness value is extrapolated from said at least two visual loudness components and from said at least two corresponding characteristic luminance values.

15 **9.** A method or an apparatus according to any one of claims 3 to 8,
wherein said **obtaining** (100a) at least two visual loudness components comprises:

- **determining** (100a3) a mask M_h representative of a location of said highlight pixels in said image; and
- **splitting** (100a4) said mask M_h when a density of said highlight pixels in said mask M_h is below a highlight pixels density threshold,

20 said splitting providing at least one sub-mask presenting a higher density of highlight pixels than said mask M_h ;
said highlight visual loudness component V_h being a function of a density of highlight pixels of said at least one sub-mask and of a size of said at least one sub-mask.

25 **10.** A method or an apparatus according to claim 9,
wherein said function is a non-linear function of said size of said at least one sub-mask.

11. Method for adapting a luminance of a multimedia content according to any of the claims 7 to 10,
30 wherein said mid-tone visual loudness component V_m is a function of a barycenter of a luminance of said mid-tone pixels with respect to said characteristic highlight L_h and/or dark L_d luminance values.

12. A method or an apparatus according to any of claims 7 to 11,
wherein said dark visual loudness component V_d is a function of a barycenter of a luminance
of said dark pixels with respect to said characteristic dark luminance value L_d and/or of a zero
luminance reference value.

5

13. A method or an apparatus according to any of claims 3 to 12,
wherein said multimedia content is a video content,
and wherein said **obtaining** (100) a visual loudness value further comprises **adjusting** (100b)
visual loudness components for providing at least two adjusted visual loudness components
as a function of said at least two visual loudness components and of at least two previous
visual loudness components obtained for a previous image in said video content,
said visual loudness value being obtained from said at least two adjusted visual loudness
components.

10

15

14. A computer program product comprising program code instructions for implementing
the method according to at least one of claims 1 or 3 to 13, when said program is executed
on a computer or a processor.

20

15. A non-transitory computer-readable carrier medium storing a computer program
product according to claim 14.

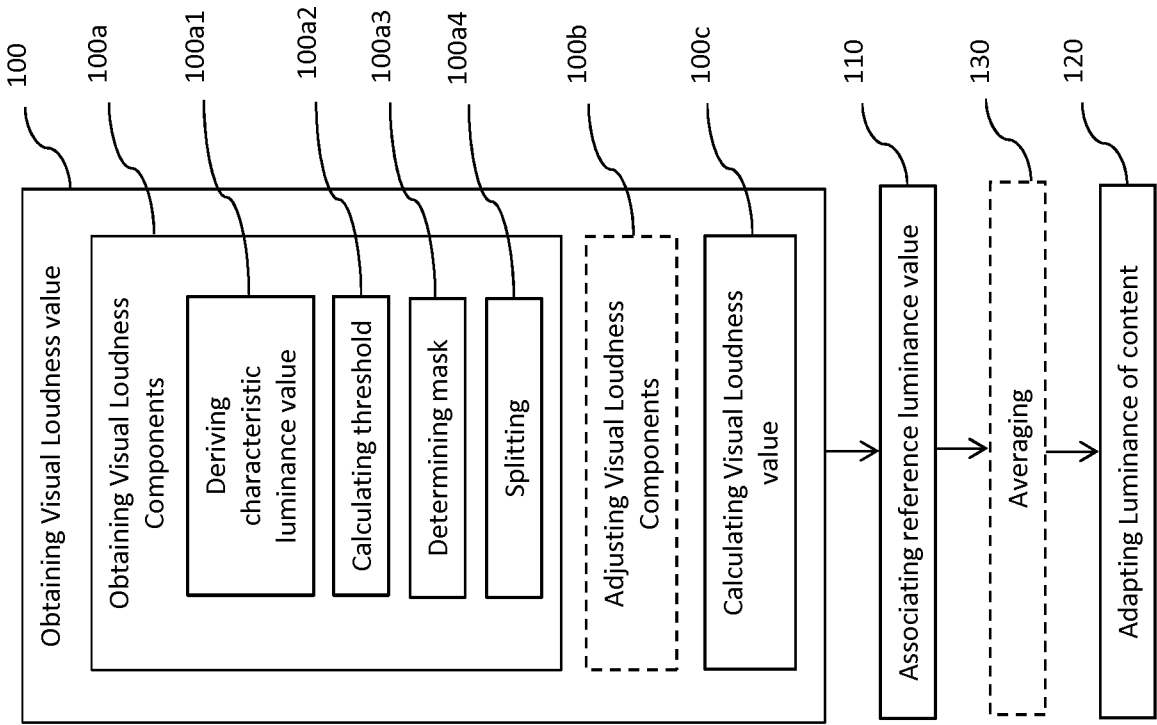


Figure 1b

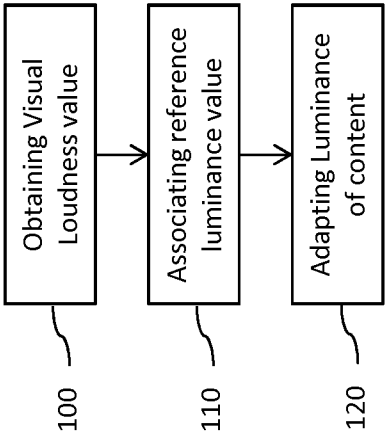


Figure 1a

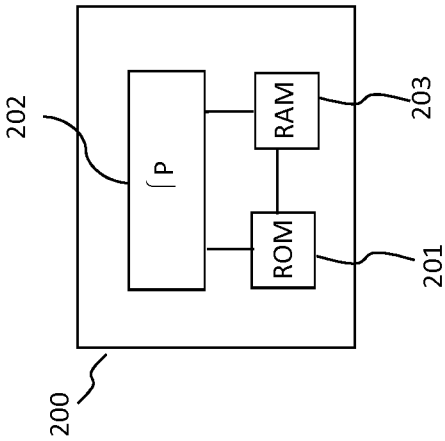


Figure 2

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2017/077161

A. CLASSIFICATION OF SUBJECT MATTER
INV. G09G5/10
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)
G09G 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 A	US 2016/254028 A1 (ATKINS ROBIN [US] ET AL) 1 September 2016 (2016-09-01) figures 1-5 paragraphs [0033] - [0036], [0041], [0050], [0052], [0068], [0075] - [0085] -----	1-8,14, 15 9-13
X A	US 2004/257318 A1 (ITOH HIROSHI [JP]) 23 December 2004 (2004-12-23) figures 1-7 paragraphs [0005], [0024] - [0030], [0057] - [0060], [0064] - [0066], [0082] - [0084], [0123], [0124] -----	1-8,14, 15 9-13
A	US 2015/035868 A1 (LIN MING-TSUN [TW] ET AL) 5 February 2015 (2015-02-05) paragraphs [0012], [0022], [0025] - [0027], [0036], [0037]; figures 1-4 ----- -/-	1-15



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

8 January 2018

Date of mailing of the international search report

17/01/2018

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

Maciu, Emanoil

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2017/077161

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2009/146941 A1 (FUJINE TOSHIYUKI [JP] ET AL) 11 June 2009 (2009-06-11) figures 1-14 paragraphs [0012], [0014], [0015], [0020], [0039], [0076], [0077], [0081] -----	1-15

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2017/077161

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2016254028 A1	01-09-2016	CN 105409203 A	16-03-2016
		EP 3028446 A1	08-06-2016
		HK 1217068 A1	16-12-2016
		JP 6227778 B2	08-11-2017
		JP 2016528826 A	15-09-2016
		KR 20160027975 A	10-03-2016
		KR 20170021384 A	27-02-2017
		RU 2016102884 A	03-08-2017
		TW 201521435 A	01-06-2015
		TW 201630429 A	16-08-2016
		US 2016254028 A1	01-09-2016
		US 2017125063 A1	04-05-2017
		WO 2015017314 A1	05-02-2015
US 2004257318 A1	23-12-2004	CN 1613103 A	04-05-2005
		CN 101241684 A	13-08-2008
		EP 1453030 A1	01-09-2004
		JP 3583122 B2	27-10-2004
		JP WO2003038799 A1	24-02-2005
		US 2004257318 A1	23-12-2004
		US 2006125775 A1	15-06-2006
		WO 03038799 A1	08-05-2003
US 2015035868 A1	05-02-2015	CN 104349098 A	11-02-2015
		TW 201505445 A	01-02-2015
		US 2015035868 A1	05-02-2015
US 2009146941 A1	11-06-2009	CN 101268502 A	17-09-2008
		EP 1939850 A1	02-07-2008
		JP 3953506 B2	08-08-2007
		JP 2007140483 A	07-06-2007
		KR 20080040749 A	08-05-2008
		TW 200721085 A	01-06-2007
		US 2009146941 A1	11-06-2009
		WO 2007046319 A1	26-04-2007