(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 October 2012 (18.10.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/142506 A1

(51) International Patent Classification:
H04N 7/26 (2006.01)

(21) International Application Number:
PCT/US2012/033651

(22) International Filing Date:
13 April 2012 (13.04.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/475,372 14 April 2011 (14.04.2011) US

(71) Applicant *(for all designated States except US)*: **DOLBY LABORATORIES LICENSING CORPORATION** [US/US]; 100 Potrero Avenue, San Francisco, California 94103-4813 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SU, Guan-Ming** [—/US]; c/o Dolby Laboratories, Inc., 100 Potrero Avenue, San Francisco, California 94103 (US). **QU, Sheng** [CN/US]; c/o Dolby Laboratories, Inc., 100 Potrero Avenue, San Francisco, California 94103 (US). **KOEPFER, Hubert** [DE/US]; c/o Dolby Laboratories, Inc., 100 Potrero Avenue, San Francisco, California 94103 (US). **YUAN, Yufei** [CA/US]; c/o Dolby Laboratories, Inc., 100 Potrero Avenue, San Francisco, California 94103 (US). **HULYALKAR, Samir** [US/US]; c/o Dolby Laboratories, Inc., 100 Potrero Avenue, San Francisco, California 94103 (US). **GISH, Walter** [US/US]; c/o Dolby Laboratories, Inc., 3601 West Alameda Avenue, Burbank, California 91505-5300 (US).

(74) Agents: **KONSTANTINIDES, Konstantinos** et al.; Dolby Laboratories, Inc., Intellectual Property Group, 999 Brannan Street, San Francisco, California 94103-4938 (US).

(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, 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, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, 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, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, 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, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to the identity of the inventor (Rule 4.17(i))*

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

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

— *of inventorship (Rule 4.17(iv))*

Published:

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

(54) **Title**: IMAGE PREDICTION BASED ON PRIMARY COLOR GRADING MODEL

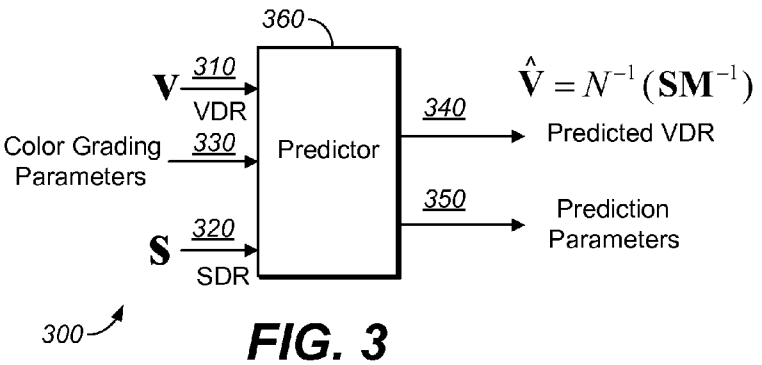


FIG. 3

(57) **Abstract**: Inter-color image prediction is based on color grading modeling. Prediction is applied to the efficient coding of images and video signals of high dynamic range. Prediction models may include a color transformation matrix that models hue and saturation color changes and a non-linear function modeling color correction changes. Under the assumption that the color grading process uses a slope, offset, and power (SOP) operations, an example non linear prediction model is presented.

WO 2012/142506 A1

IMAGE PREDICTION BASED ON PRIMARY COLOR GRADING MODEL

Cross-Reference to Related Applications

[0001] This application claims priority to United States Provisional Patent Application No. 61/475,372 filed 14 April 2011, which is incorporated herein by reference in its entirety.

[0002] This application is also related to co-pending International Patent Application No. PCT/US2012/033605 filed concurrently, 13 April 2012, which is incorporated herein by reference in its entirety.

Technical Field

[0003] The present invention relates generally to images. More particularly, an embodiment of the present invention relates to a multiple-color channel predictor between high dynamic range images and standard dynamic range images based on a primary color grading model.

Background of the Invention

[0004] As used herein, the term 'dynamic range' (DR) may relate to a capability of the human psychovisual system (HVS) to perceive a range of intensity (e.g., luminance, luma) in an image, e.g., from darkest darks to brightest brights. In this sense, DR relates to a 'scene-referred' intensity. DR may also relate to the ability of a display device to adequately or approximately render an intensity range of a particular breadth. In this sense, DR relates to a 'display-referred' intensity. Unless a particular sense is explicitly specified to have particular significance at any point in the description herein, it should be inferred that the term may be used in either sense, e.g. interchangeably.

[0005] As used herein, the term high dynamic range (HDR) relates to a DR breadth that spans the some 14-15 orders of magnitude of the human visual system (HVS). For example, well adapted humans with essentially normal (e.g., in one or more of a statistical, biometric or opthamological sense) have an intensity range that spans about 15 orders of magnitude. Adapted humans may perceive dim light sources of as few as a mere handful of photons. Yet, these same humans may perceive the near painfully brilliant intensity of the noonday sun in desert, sea or snow (or even glance into the sun, however briefly to prevent damage). This span though is available to 'adapted' humans, e.g., those whose HVS has a time period in which to reset and adjust.

[0006] In contrast, the DR over which a human may simultaneously perceive an extensive breadth in intensity range may be somewhat truncated, in relation to HDR. As used herein, the terms 'visual dynamic range' or 'variable dynamic range' (VDR) may individually or interchangeably relate to the DR that is simultaneously perceivable by a HVS. As used

herein, VDR may relate to a DR that spans 5-6 orders of magnitude. Thus while perhaps somewhat narrower in relation to true scene referred HDR, VDR nonetheless represents a wide DR breadth. As used herein, the term 'simultaneous dynamic range' may relate to VDR.

[0007] Until fairly recently, displays have had a significantly narrower DR than HDR or VDR. Television (TV) and computer monitor apparatus that use typical cathode ray tube (CRT), liquid crystal display (LCD) with constant fluorescent white back lighting or plasma screen technology may be constrained in their DR rendering capability to approximately three orders of magnitude. Such conventional displays thus typify a low dynamic range (LDR), also referred to as a standard dynamic range (SDR), in relation to VDR and HDR.

[0008] Advances in their underlying technology however allow more modern display designs to render image and video content with significant improvements in various quality characteristics over the same content, as rendered on less modern displays. For example, more modern display devices may be capable of rendering high definition (HD) content and/or content that may be scaled according to various display capabilities such as an image scaler. Moreover, some more modern displays are capable of rendering content with a DR that is higher than the SDR of conventional displays.

[0009] For example, some modern LCD displays have a backlight unit (BLU) that comprises a light emitting diode (LED) array. The LEDs of the BLU array may be modulated separately from modulation of the polarization states of the active LCD elements. This dual modulation approach is extensible (e.g., to N-modulation layers wherein N comprises an integer greater than two), such as with controllable intervening layers between the BLU array and the LCD screen elements. Their LED array based BLUs and dual (or N-) modulation effectively increases the display referred DR of LCD monitors that have such features.

[00010] Such "HDR displays" as they are often called (although actually, their capabilities may more closely approximate the range of VDR) and the DR extension of which they are capable, in relation to conventional SDR displays represent a significant advance in the ability to display images, video content and other visual information. The color gamut that such an HDR display may render may also significantly exceed the color gamut of more conventional displays, even to the point of capably rendering a wide color gamut (WCG). Scene related HDR or VDR and WCG image content, such as may be generated by "next generation" movie and TV cameras, may now be more faithfully and effectively displayed with the "HDR" displays (hereinafter referred to as 'HDR displays').

[00011] As with the scalable video coding and HDTV technologies, extending image DR typically involves a bifurcate approach. For example, scene referred HDR content that is captured with a modern HDR capable camera may be used to generate an SDR version of the content, which may be displayed on conventional SDR displays. In one approach, generating the SDR version from the captured VDR version may involve applying a global tone mapping operator (TMO) to intensity (e.g., luminance, luma) related pixel values in the HDR content. In a second approach, as described in International Patent Application No. PCT/US2011/048861, herein incorporated by reference for all purposes, generating an SDR image may involve applying an invertible operator (or predictor) on the VDR data. To conserve bandwidth or for other considerations, simultaneous transmission of both the actual captured VDR content and the SDR content may not be a best approach.

[0010] Thus, an inverse tone mapping operator (iTMO), inverted in relation to the original TMO, or an inverse operator in relation to the original predictor, may be applied to the SDR content version that was generated, which allows a version of the VDR content to be predicted. The predicted VDR content version may be compared to originally captured HDR content. For example, subtracting the predicted VDR version from the original VDR version may generate a residual image. An encoder may send the generated SDR content as a base layer (BL), and package the generated SDR content version, any residual image, and the iTMO or other predictors as an enhancement layer (EL) or as metadata.

[0011] Sending the EL and metadata, with its SDR content, residual and predictors, in a bitstream typically consumes less bandwidth than would be consumed in sending both the HDR and SDR contents directly into the bitstream. Compatible decoders that receive the bitstream sent by the encoder may decode and render the SDR on conventional displays. Compatible decoders however may also use the residual image, the iTMO predictors, or the metadata to compute a predicted version of the HDR content therefrom, for use on more capable displays. It is the purpose of this invention to provide novel methods for generating predictors that allow for the efficient coding, transmission, and decoding of VDR data using corresponding SDR data.

[0012] In one example embodiment of this invention, we present a novel VDR predictor that is based on primary color grading. It combines a matrix multiplication operation with scale, offset, and power (SOP) operations. Assuming the SDR content was created using traditional color grading operations, such as: lift, gain, gamma, hue and saturation changes, the disclosed predictor is advantageous over the prior art by yielding smaller prediction error while requiring to estimate only a very small number of prediction coefficients.

[0013] The approaches described in this section are approaches that could be pursued, but not necessarily approaches that have been previously conceived or pursued. Therefore, unless otherwise indicated, it should not be assumed that any of the approaches described in this section qualify as prior art merely by virtue of their inclusion in this section. Similarly, issues identified with respect to one or more approaches should not assume to have been recognized in any prior art on the basis of this section, unless otherwise indicated.

Brief Description of the Drawings

[0014] An embodiment of the present invention is illustrated by way of example, and not in way by limitation, in the figures of the accompanying drawings and in which like reference numerals refer to similar elements and in which:

[0015] FIG. 1 depicts an example data flow for a VDR-SDR system, according to an embodiment of the present invention;

[0016] FIG. 2 depicts an example VDR encoding system according to an embodiment of the present invention;

[0017] FIG. 3 depicts input and output interfaces of a predictor based on a primary color grading model, according to an embodiment of the present invention;

[0018] FIG. 4 depicts an example prediction process that takes into consideration the color grading process according to an embodiment of the present invention;

[0019] FIG. 5 depicts an example prediction process according to an SOP+M color grading model of this invention;

[0020] FIG. 6 depicts an example image decoder with a predictor operating according to the embodiments of this invention.

Description of Example Embodiments

[0021] Inter-color image prediction that is based on color grading modeling is described herein. Given a pair of corresponding VDR and SDR images, that is, images that represent the same scene but at different levels of dynamic range, this section describes methods that allow an encoder to approximate the VDR image in terms of the SDR image and an inter-color predictor based on a primary color grading model. In the following description, for the purposes of explanation, numerous specific details are set forth in order to provide a thorough understanding of the present invention. It will be apparent, however, that the present invention may be practiced without these specific details. In other instances, well-known structures and devices are not described in exhaustive detail, in order to avoid unnecessarily occluding, obscuring, or obfuscating the present invention.

OVERVIEW

[0022] Example embodiments described herein relate to coding images with high dynamic range. An embodiment creates a predictor that allows a VDR image to be expressed in relation to its corresponding SDR representation and parameters of a color grading model.

EXAMPLE VDR-SDR SYSTEM

[0023] FIG. 1 depicts an example data flow in a VDR-SDR system 100, according to an embodiment of the present invention. An HDR image or video sequence is captured using HDR camera 110. Following capture, the captured image or video is processed by a mastering process to create a target VDR image 125. The mastering process may incorporate a variety of processing steps, such as: editing, primary and secondary color correction (or grading), color transformation, and noise filtering. The VDR output 125 of this process represents the director's intend on how the captured image will be displayed on a target VDR display.

[0024] The mastering process may also output a corresponding SDR image 145, representing the director's intend on how the captured image will be displayed on a legacy SDR display. The SDR output 145 may be provided directly from mastering circuit 120 or it may be generated by a separate VDR-to-SDR converter 140.

[0025] In this example embodiment, the VDR 125 and SDR 145 signals are input into an encoder 130. Purpose of encoder 130 is to create a coded bitstream that reduces the bandwidth required to transmit the VDR and SDR signals, but also allows a corresponding decoder 150 to decode and render either the SDR or VDR signals. In an example implementation, encoder 130 may be a layered encoder, such as one of those defined by the MPEG-2 and H.264 coding standards, which represents its output as a base layer, an optional enhancement layer, and metadata. We define as "metadata" any ancillary information that is embedded or transmitted in parallel with the coded bitstream and assists a decoder to render a decoded image. Such metadata may include, but are not limited to, such data as: color space or gamut information, dynamic range information, tone mapping information, or prediction coefficients, such as those described herein.

[0026] On the receiver, a decoder 150, uses the received coded bitstreams and metadata to render either an SDR image or a VDR image, according to the capabilities of the target display. For example, an SDR display may use only the base layer and the metadata to render an SDR image. In contrast, a VDR display may use information from all input layers and the metadata to render a VDR signal.

[0027] FIG. 2 shows in more detail an example implementation of encoder 130 incorporating the methods of this invention. In FIG. 2, SDR' denotes an enhanced SDR signal. SDR video today is 8-bit, 4:2:0, ITU Rec. 709 data. SDR' may have the same color space (primaries and white point) as SDR, but may use high precision, say 12-bits per pixel, with all color components at full spatial resolution (e.g., 4:4:4 RGB). From FIG. 2, SDR can easily be derived from an SDR' signal using a set of forward transforms that may include quantization from say 12 bits per pixel to 8 bits per pixel, color transformation, say from RGB to YUV, and color subsampling, say from 4:4:4 to 4:2:0. The SDR output of converter 210 is applied to compression system 220. Depending on the application, compression system 220 can be either lossy, such as H.264 or MPEG-2, or lossless. The output of the compression system 220 may be transmitted as a base layer 225. To reduce drift between the encoded and decoded signals, encoder 130 may follow compression process 220 with a corresponding decompression process 230 and inverse transforms 240, corresponding to the forward transforms of 210. Thus, predictor 250 may have the following inputs: VDR input 205 and either SDR' signal 245, which corresponds to the SDR' signal as it will be received by a corresponding decoder, or input SDR' 207. Predictor 250, using input VDR and SDR' data will create signal 257 which represents an approximation or estimate of input VDR 205. Adder 260, subtracts the predicted VDR 257 from the original VDR 205 to form output residual signal 265. Subsequently (not shown), residual 265 may also be coded by another lossy or lossless encoder, and may be transmitted to the decoder as an enhancement layer.

[0028] Predictor 250 may also provide the prediction parameters being used in the prediction process as metadata 255. Since prediction parameters may change during the encoding process, for example, on a frame by frame basis, or on a scene by scene basis, these metadata may be transmitted to the decoder as part of the data that also include the base layer and the enhancement layer.

[0029] Since both VDR 205 and SDR' 207 represent the same scene, but are targeting different displays with different characteristics, such as dynamic range and color gamut, it is expected that there is a very close correlation between these two signals. For example, International Patent Application No. PCT/US2012/033605 filed 13 April 2012, describes a multivariate, multi-regression (MMR) predictor 250 that allows the input VDR signal to be predicted using its corresponding SDR' signal and a multivariate MMR operator. Such a predictor can be considered a "blind" predictor in the sense that it does not take into consideration how the SDR' signal was generated out of the original VDR signal. In this invention, we present a novel predictor that minimizes the prediction error under the

assumption that the SDR' signal was generated from the VDR signal using primary color correction operations, such as lift, gain, gamma, and hue and saturation changes.

EXAMPLE COLOR GRADING MODELS

[0030] We define as color grading the pre-production or mastering process where a professional colorist, the director, or the like, may set tones and colors of pixels in the image so that when viewed, the image has a desired appearance that agrees with the director's creative intent. For example, a director may wish the sky to look brighter or a scene to look darker, stand out, or be less prominent. Adjusting tones and colors of pixels in an image may include performing color grading (sometimes referred to with the legacy term 'color timing') on the source video data. Color grading may be performed using a hardware/software system that permits a user to change the video data in various ways to achieve a desired appearance on a target display.

[0031] The main controls in primary color correction are: lift, gain, and gamma (LGG) controls, and hue and saturation (H.S.) controls. Lift changes the darkness, gain changes the brightness, and gamma changes the mid-tones.

[0032] Given input x, and lift (L), gain (G), and *gamma* parameters, the LGG-corrected output y is given by:

$$y = G \cdot (x + L \cdot (1 - x))^{gamma} \quad . \tag{1}$$

[0033] Given input 1 x r vector **x** (say, **x** = [x_0, x_1, …, x_r]), where r represents the number of color components of an input pixel, the hue and saturation operations can be expressed as a matrix-vector multiplication

$$\mathbf{y} = \mathbf{x}\,\mathbf{M}, \tag{2}$$

where **M** is an r x r transformation matrix and **y** is an 1 x r output vector. For example, if the input pixels are defined in the RGB color space, r=3, and **M** will be a 3 x 3 matrix.

[0034] FIG. 3 shows input and output interfaces of predictor 360 according to an example implementation of this invention. From FIG. 3, predictor 360 receives input vectors **v** 310 and **s** 320 representing VDR and SDR image data respectively, and outputs vector $\hat{\mathbf{v}}$ 340, representing the predicted value of input **v**. Optionally, predictor 360 may also receive input metadata 330 related to the color grading process that was used to derive **s** from **v**. Predictor 360 may also output computed prediction parameters 350 which may be passed to a decoder as metadata.

EXAMPLE NOTATION AND NOMENCLATURE

[0035] Assuming images are represented by pixels, where each pixel has n color components, in an example implementation where n=3, denote the three color components of the i-th pixel in the SDR image 320 as

$$\mathbf{s}_i = [s_{i1} \quad s_{i2} \quad s_{i3}]. \tag{3}$$

[0036] Denote the three color components of the i-th pixel in the VDR input 310 as

$$\mathbf{v}_i = [v_{i1} \quad v_{i2} \quad v_{i3}]. \tag{4}$$

[0037] Denote the predicted three color components of the i-th pixel in predicted VDR 340 as

$$\hat{\mathbf{v}}_i = [\hat{v}_{i1} \quad \hat{v}_{i2} \quad \hat{v}_{i3}]. \tag{5}$$

Denote the total number of pixels in one color component as p. In equations (3-5), image pixels may be in RGB, YUV, YCbCr, XYZ, or any other color representation. While equations (3-5) assume a three-color representation for each pixel in an image or video frame, as also shown later, the methods described herein can be easily extended to image and video representations with more than three color components per pixel, or to image representations where one of the inputs may have pixels with a different number of color components than the other input.

[0038] From equations (1) and (2), given a VDR input **v,** assuming the SDR signal is created via a two-stage operation; namely, non-linear operations on each individual color channel for range adjustments, and a linear transform for hue and saturation adjustments, then according to an example implementation of this invention, the VDR to SDR model can be expressed as

$$\mathbf{s} = N(\mathbf{v})\,\mathbf{M}, \tag{6}$$

where $N()$ is a non-linear function and **M** is a matrix operator. Function $N()$ can be any non-linear function being used in tone mapping or color-gamut transformation. It is desirable, but not a requirement, that function $N()$ is invertible; that is, given $y = N(x)$, there is a corresponding function $N^{-1}()$, such that $x = N^{-1}(y)$.

[0039] An example of a sigmoid-like $N()$ function is presented in International Patent Application No. PCT/US2012/029189 filed on 15 March 2012. Alternatively, in another example implementation as described herein, $N()$ may represent a non-linear function in a slope-offset-power (SOP) format, as described in "ASC color decision list (ASC CDL) transfer functions and interchange syntax", by J. Pines and D. Reisner, published by the American Society of Cinematographers, (document: ASC-CDL_Release 1.2, 2009-05-04).

[0040] From equation (6), given input SDR signal **s**, output $\hat{\mathbf{v}}$, an approximation of the original VDR signal **v**, can be derived using

$$\hat{\mathbf{v}} = N^{-1}(\mathbf{s}\mathbf{M}^{-1}), \tag{7}$$

where $N^{-1}()$ is the invert of $N()$ and $\mathbf{M}^{-1}$ is the invert of **M**.

[0041] The traditional lift, gain, and gamma operations, individually or in combination, can be readily translated into slope, offset, and power (SOP) operations as described herein. Given input x and SOP parameters *Slope*, P (power), and O (offset), SOP operations can be expressed as,

$$N(x) = (Slope \cdot x + O)^P, \tag{8}$$

and

$$N^{-1}(y) = \frac{1}{Slope}(y^{1/P} - O). \tag{9}$$

[0042] From equations (6-9), if the output of the VDR to SDR process (say, 145 or 320) provides to the predictor (say, 250 or 360) the color grading parameters used to create the SDR input, then predictor 360 can derive an estimate of the VDR input using equations (7) and (9). In this prediction model, referred herein as the SOP+M model, the color grading parameters being used are a list of *slope*, P, and O parameters (collectively referred to as the ASC Color Decision List (ASC CDL), and the elements of the linear transform matrix **M** which may represent the product of all matrices being used for hue and saturation transforms. In practice however, these **M** and SOP parameters may not be available. In that case, predictor 360 will have to estimate them. An example predictor that estimates these parameters is described next.

EXAMPLE PREDICTION USING SOP+M COLOR GRADING MODEL

[0043] In equation (7), notation is simplified, as the matrix $\mathbf{M}^{-1}$ may be denoted as

$$\mathbf{M}^{-1} = \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \\ \mathbf{m}_3 \end{bmatrix}, \tag{10}$$

wherein $\mathbf{m}_j = [\, m_{j1}\ m_{j2}\ m_{j3}\,]$, for $j = 1,2,3$, denotes a 1 x 3 vector corresponding to the j-th row in $\mathbf{M}^{-1}$ and representing the linear hue and saturation color transformations related to deriving the j-th color component of the predicted output pixel.

[0044] From (7) and (9), $N^{-1}()$ and $\mathbf{M}^{-1}$ may be combined into a single prediction function and apply non-linear regression methods to estimate the free parameters m_{ij}, $Slope_i$, P_i, and O_i (i and $j = 1, 2, 3$) as follows:

$$\hat{v}_{ij} = \frac{1}{Slope_j}((\mathbf{m}_j \mathbf{s}_i^T)^{1/P_j} - O_j), \; i = 1, 2, \ldots, p, \text{ and }, j = 1, 2, 3. \quad (11)$$

[0045] In equation (11), $\mathbf{s}_i^T$ denotes the transpose of vector $\mathbf{s}_i$, and the j-th color component of pixel $\hat{v}_i$ depends on all three color components of the corresponding s_i pixel. Thus, unlike other single-color channel predictors, the SOP+M predictor is an inter-color predictor and thus can better predict hue and saturation transformations during the color grading process.

[0046] From equation (11), each predicted VDR color channel j can be expressed as a non-linear function dependent on six free parameters (m_{j1}, m_{j2}, m_{j3}, $Slope_j$, P_j, and O_j). Thus, for VDR and SDR images with 3-color components, the SOP+M prediction model requires a total of 18 prediction parameters. Taking into consideration all pixels in an image, equation (11) can be also be expressed as

$$\hat{\mathbf{V}} = NS^{-1}(\mathbf{c}, \mathbf{S}), \quad (12)$$

[0047] where $NS^{-1}()$ represents a non-linear function with inputs

$$\mathbf{S} = \begin{bmatrix} \mathbf{s}_0 \\ \mathbf{s}_1 \\ \vdots \\ \mathbf{s}_{p-1} \end{bmatrix}, \quad (13)$$

[0048] and a 1x18 "SOP+M" parameter vector $\mathbf{c} = [\mathbf{c}_1 \; \mathbf{c}_2 \; \mathbf{c}_3]$, where $\mathbf{c}_j = [m_{j1}, m_{j2}, m_{j3}, Slope_j, P_j, \text{and } O_j]$, ($j$=1, 2, 3), and

$$\hat{\mathbf{V}} = \begin{bmatrix} \hat{\mathbf{v}}_0 \\ \hat{\mathbf{v}}_1 \\ \vdots \\ \hat{\mathbf{v}}_{p-1} \end{bmatrix}. \quad (14)$$

[0049] Given the system of non-linear equations (12), one approach to solve for parameter vector $\mathbf{c}$ is by solving the optimization problem

$$\min_{\mathbf{c}} \varepsilon = \left\| \mathbf{V} - \hat{\mathbf{V}} \right\|^2, \quad (15)$$

[0050] where ε denotes the prediction error.

[0051] An example embodiment computes this minimization problem with application of numerical, non-linear regression processes. An embodiment may implement the numerical, non-linear regression processes, for example, with the Levenberg-Marquard method described in Chapter 10 of the book *"Numerical Optimization,"* by J. Nocedal and S. J.

Wright, Springer, 1999. While the Levenberg-Marquard method is a descriptive example, embodiments of are not limited to computing the numerical, non-linear regression processes with this particular method. On the contrary; embodiments of the present invention are well suited to compute the numerical, non-linear regression processes with other methods, as well.

[0052] FIG. 4 depicts an example implementation of a prediction process 400 that takes into consideration the color grading process that was applied to derive input SDR from input VDR. In step 410, process 400 receives SDR and VDR inputs plus optional metadata characterizing the color grading process used to derive the SDR signal from the VDR signal. For example, such metadata may include the color mapping function *N()*, a color transformation matrix **M**, or color decision list (CDL) data representing a sequence of LGG or SOP steps applied to the original VDR image. In step 420, process 400 applies an inverse color grading model to compute an estimate of the VDR input. If any of the color grading model parameters were missing, they will be estimated in step 420 as well. Finally, the predicted VDR and any associated prediction parameters are output in step 430.

[0053] Prediction process 400 can be repeated at a variety of time intervals, as deemed necessary to maintain coding efficiency while processing requirements are within the available computing resources. For example, when coding video signals, process 400 can be computed for each frame, a part of a frame, or a group of frames, whenever the prediction residual error exceeds a particular threshold, or using a combination of criteria.

[0054] Prediction process 400 can also use all available input pixels or a sub-sample of those pixels. In one example implementation, one may use pixels only from every *k*-th pixel row and every *k*-th pixel column of the input data, where *k* is an integer equal or higher than two. In another example implementation one may decide to skip input pixels that are below a certain clipping threshold (for example, very close to zero) or pixels that are above a certain saturation threshold (for example, for *n*-bit data, pixel values that are very close to 2^n -1.) In yet another implementation, one may use a combination of such subsampling and thresholding techniques in order to reduce the pixel sample size and accommodate the computational constrains of a particular implementation.

[0055] FIG. 5 depicts in more detail an example implementation of step 420 when the color grading process is based on a SOP+M model. If the predictor has full access to the hue/saturation transformation matrix **M** and the color decision list (CDL), then, in step 540, the predictor can apply equation (7). Otherwise, in step 520, the predictor can use the system of equations (12) to formulate an optimization problem and solve for the prediction parameters vector **c**. In an example implementation, this system of equations can be solved

my minimizing the prediction error using non-linear regression techniques as discussed earlier.

APPLYING THE SOP+M MODEL TO IMAGES WITH MORE THAN THREE PRIMARY COLORS

[0056] All of the proposed color-grading-based prediction models can easily be extended to images that are expressed with more than three primary colors. As an example, we consider the case where an SDR signal has three primary colors, say RGB, but the VDR signal is defined in the P6 color space, with six primary colors. In this case, equations (3-5) can be re-written as

$$\mathbf{s}_i = [s_{i1} \quad s_{i2} \quad s_{i3}], \tag{16}$$

$$\mathbf{v}_i = [v_{i1} \quad v_{i2} \quad v_{i3} \quad v_{i4} \quad v_{i5} \quad v_{i6}], \tag{17}$$

and

$$\hat{\mathbf{v}}_i = [\hat{v}_{i1} \quad \hat{v}_{i2} \quad \hat{v}_{i3} \quad \hat{v}_{i4} \quad \hat{v}_{i5} \quad \hat{v}_{i6}]. \tag{18}$$

[0057] As before, denote the number of pixels in one color component as p. Given equations (16-18), the SOP+M predictor of equation (11), can be expressed as

$$\hat{v}_{ij} = \frac{1}{Slope_j}((\mathbf{m}_j \mathbf{s}_i^T)^{1/P_j} - O_j), \; j = 1, 2, \ldots, 6, \text{ and } i = 1, 2, \ldots, p, \tag{19}$$

where $\mathbf{m}_j$ denotes the j-th row vector of coefficients of a 6 x 3 matrix $\mathbf{M}^{-1}$, as described earlier. Under this example implementation, in the SOP+M model of equation (12), the parameter vector $\mathbf{c}$ will consist of 36 total unknown parameters per set of p image pixels, which can be computed using the methods described earlier. Note also that $\mathbf{M}^{-1}$ as defined herein is only an approximate representation of the inverse color transformation process under the SOP+M model. For non square $\mathbf{M}$ matrices (resulting when the VDR and SDR images don't have the same number of color components), $\mathbf{M}^{-1}$ is more closely related to the pseudo-inverse of $\mathbf{M}$.

IMAGE DECODING

[0058] Embodiments of the present invention may be implemented either on an image encoder or an image decoder. FIG. 6 shows an example implementation of decoder 150 according to an embodiment of this invention.

[0059] Decoding system 600 receives a coded bitstream that may combine a base layer 690, an optional enhancement layer (or residual) 665, and metadata 645, which are extracted following decompression 630 and miscellaneous inverse transforms 640. For example, in a VDR-SDR system, the base layer 690 may represent the SDR representation of the coded signal and the metadata 645 may include information about the prediction model that was used in the encoder predictor 250 and the corresponding prediction parameters. In one example implementation, when the encoder uses a color-grading-based prediction model according to the methods of this invention, metadata may include the color transformation matrix **M,** the CDL data, information about the $N()$ color transform function, or the values of the color grading parameter vector **c**. Given base layer 690 **s** and the color grading parameters extracted from the metadata 645, predictor 650 can compute predicted $\hat{\mathbf{v}}$ 680 using either equation (7) or equations (11-12). If there is no residual, or the residual is negligible, the predicted value 680 can be output directly as the final VDR image. Otherwise, in adder 660, the output of the predictor (680) is added to the residual 665 to output VDR signal 670.

EXAMPLE COMPUTER SYSTEM IMPLEMENTATION

[0060] Embodiments of the present invention may be implemented with a computer system, systems configured in electronic circuitry and components, an integrated circuit (IC) device such as a microcontroller, a field programmable gate array (FPGA), or another configurable or programmable logic device (PLD), a discrete time or digital signal processor (DSP), an application specific IC (ASIC), and/or apparatus that includes one or more of such systems, devices or components. The computer and/or IC may perform, control or execute instructions relating to color-grading-based prediction, such as those described herein. The computer and/or IC may compute, any of a variety of parameters or values that relate to the color-grading-based prediction as described herein. The image and video dynamic range extension embodiments may be implemented in hardware, software, firmware and various combinations thereof.

[0061] Certain implementations of the invention comprise computer processors which execute software instructions which cause the processors to perform a method of the

invention. For example, one or more processors in a display, an encoder, a set top box, a transcoder or the like may implement color-grading-based prediction methods as described above by executing software instructions in a program memory accessible to the processors. The invention may also be provided in the form of a program product. The program product may comprise any medium which carries a set of computer-readable signals comprising instructions which, when executed by a data processor, cause the data processor to execute a method of the invention. Program products according to the invention may be in any of a wide variety of forms. The program product may comprise, for example, physical media such as magnetic data storage media including floppy diskettes, hard disk drives, optical data storage media including CD ROMs, DVDs, electronic data storage media including ROMs, flash RAM, or the like. The computer-readable signals on the program product may optionally be compressed or encrypted.

[0062] Where a component (e.g. a software module, processor, assembly, device, circuit, etc.) is referred to above, unless otherwise indicated, reference to that component (including a reference to a "means") should be interpreted as including as equivalents of that component any component which performs the function of the described component (e.g., that is functionally equivalent), including components which are not structurally equivalent to the disclosed structure which performs the function in the illustrated example embodiments of the invention.

EQUIVALENTS, EXTENSIONS, ALTERNATIVES AND MISCELLANEOUS

[0063] Example embodiments that relate to applying color grading-based prediction in coding VDR and SDR images are thus described. In the foregoing specification, embodiments of the present invention have been described with reference to numerous specific details that may vary from implementation to implementation. Thus, the sole and exclusive indicator of what is the invention, and is intended by the applicants to be the invention, is the set of claims that issue from this application, in the specific form in which such claims issue, including any subsequent correction. Any definitions expressly set forth herein for terms contained in such claims shall govern the meaning of such terms as used in the claims. Hence, no limitation, element, property, feature, advantage or attribute that is not expressly recited in a claim should limit the scope of such claim in any way. The specification and drawings are, accordingly, to be regarded in an illustrative rather than a restrictive sense.

CLAIMS

What is claimed is:

1. An image prediction method comprising:

receiving a first image and a second image, wherein the second image has different dynamic range than the first image;

receiving metadata related to a color grading process, wherein the color grading process is used to derive the second image from the first image;

applying said first image and second said image into an inverse color grading prediction process;

computing an output image that represents the predicted value of the first image; and

outputting the prediction parameters of the inverse color grading process and the output image.

2. An image prediction method comprising:

receiving a first image and a second image, wherein the second image has different dynamic range than the first image;

applying said first image and second said image into an SOP+M prediction process;

computing an output image that represents the predicted value of the first image; and

outputting the prediction parameters of the SOP+M process and the output image.

3. The method of claim 1 or claim 2 wherein the first image comprises a VDR image and the second image comprises an SDR image.

4. The method of claim 1, wherein the metadata that relates to the color grading process comprise one or more of a color transformation matrix and lift, gain, and gamma or slope, offset, and power parameters.

5. The method of claim 4 wherein one or more of the slope, offset, or power parameters are received as part of a color decision list.

6. The method of claim 1 wherein the inverse color grading prediction process comprises the SOP+M prediction process.

7. The method of claim 2, wherein the SOP+M prediction process comprises:

computing a first result, wherein computing the first result comprises:

raising the result of a vector multiplication between a first vector and a second vector to the power of one over a j-th power parameter;

wherein the first vector represents the values of the color components of the i-th pixel from the second image, and

wherein the second vector represents the j-th row of a color transformation

matrix; and

expressing each j-th color component of the i-th output pixel as the ratio of a numerator divided by the j-th slope parameter;

wherein the numerator comprises the j-th offset parameter subtracted from the first result.

8. The method of claim 7 wherein the output has n total color parameters, wherein n comprises a positive integer number.

9. The method of claim 8 wherein n is equal to three (3).

10. The method of claim 2, wherein the SOP+M prediction process further comprises solving for the prediction parameters.

11. The method of claim 10, wherein solving for the prediction parameters comprises minimizing a least squares error between the output image and the first input image.

12. The method of claim 1, wherein the inverse color grading prediction process further comprises multiplying a color transformation matrix with the output of a color grading function.

13. The method of claim 1 wherein the color grading process further comprises multiplying a color transformation matrix with the output of a color grading function.

14. The method of claim 13 wherein the color grading function comprises an SOP function.

15. The method of claim 14, wherein the output of the SOP function is expressed in terms of slope, offset, and power parameters, and an input.

16. The method of claim 15 wherein the output of the SOP function comprises an intermediate sum raised to the power parameter, wherein the intermediate sum comprises the offset parameter added to a result of multiplying the input of the function with the slope parameter.

17. An image decoding method comprising:

receiving a first image with a first dynamic range;

receiving metadata related to a color grading process, wherein the color grading process is used to derive the first image from a second image that has a second dynamic range that is different than the first dynamic range;

applying the first image and the metadata into an inverse color grading prediction process to compute an output image that represents the predicted value of the second image.

18. The method of claim 17 wherein the first image comprises an SDR image and the second image comprises a VDR image.

19. The method of claim 16 wherein the inverse color grading process comprises the SOP+M prediction process.

20. The method of claim 17, wherein the metadata that relate to the color grading process comprise one or more of a color transformation matrix and lift, gain, and gamma or slope, offset, and power parameters

21. The method of claim 17, wherein the inverse color prediction process comprises:

computing a first result, wherein computing the first result comprises:

raising the result of a vector multiplication between a first vector and a second vector to the power of one over a j-th power parameter;

wherein the first vector represents the values of the color components of the i-th pixel from the second image, and

wherein the second vector represents the j-th row of a color transformation matrix; and

expressing each j-th color component of the i-th output pixel as the ratio of a numerator divided by the j-th slope parameter;

wherein the numerator comprises the j-th offset parameter subtracted from the first result.

22. An apparatus comprising a processor and configured to perform any one of the methods recited in Claims 1-21.

23. A computer-readable storage medium having stored thereon computer-executable instruction for executing a method in accordance with any of the claims 1-21.

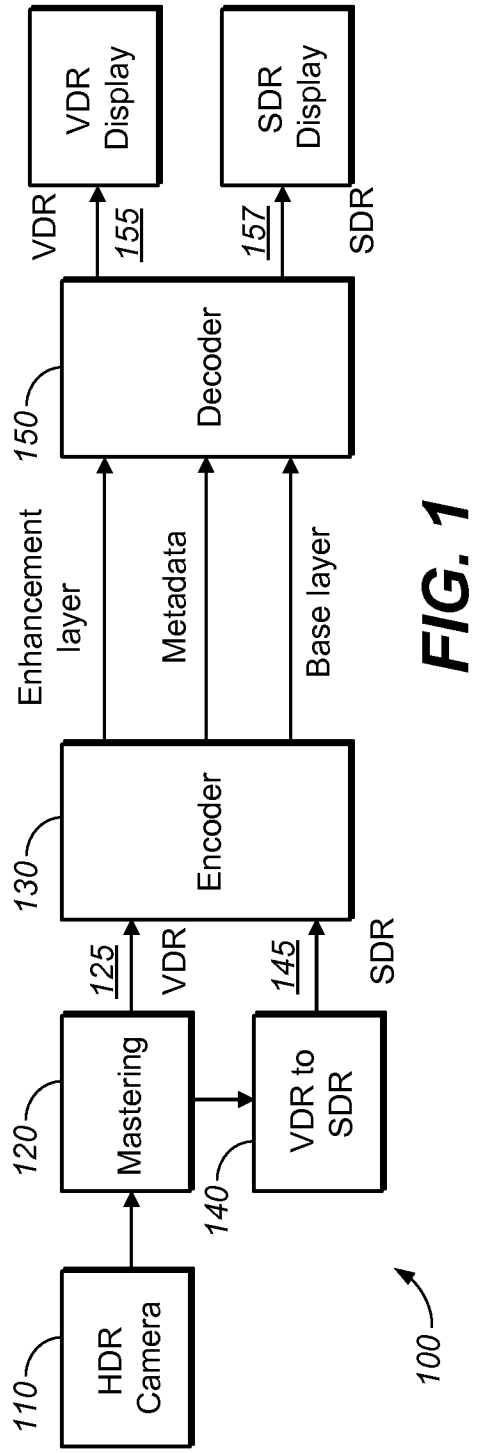

HDR Camera
110
Mastering
120
VDR to SDR
140
125
VDR
145
SDR
Encoder
130
Enhancement layer
Metadata
Base layer
Decoder
150
VDR
155
157
SDR
VDR Display
SDR Display
100

FIG. 1

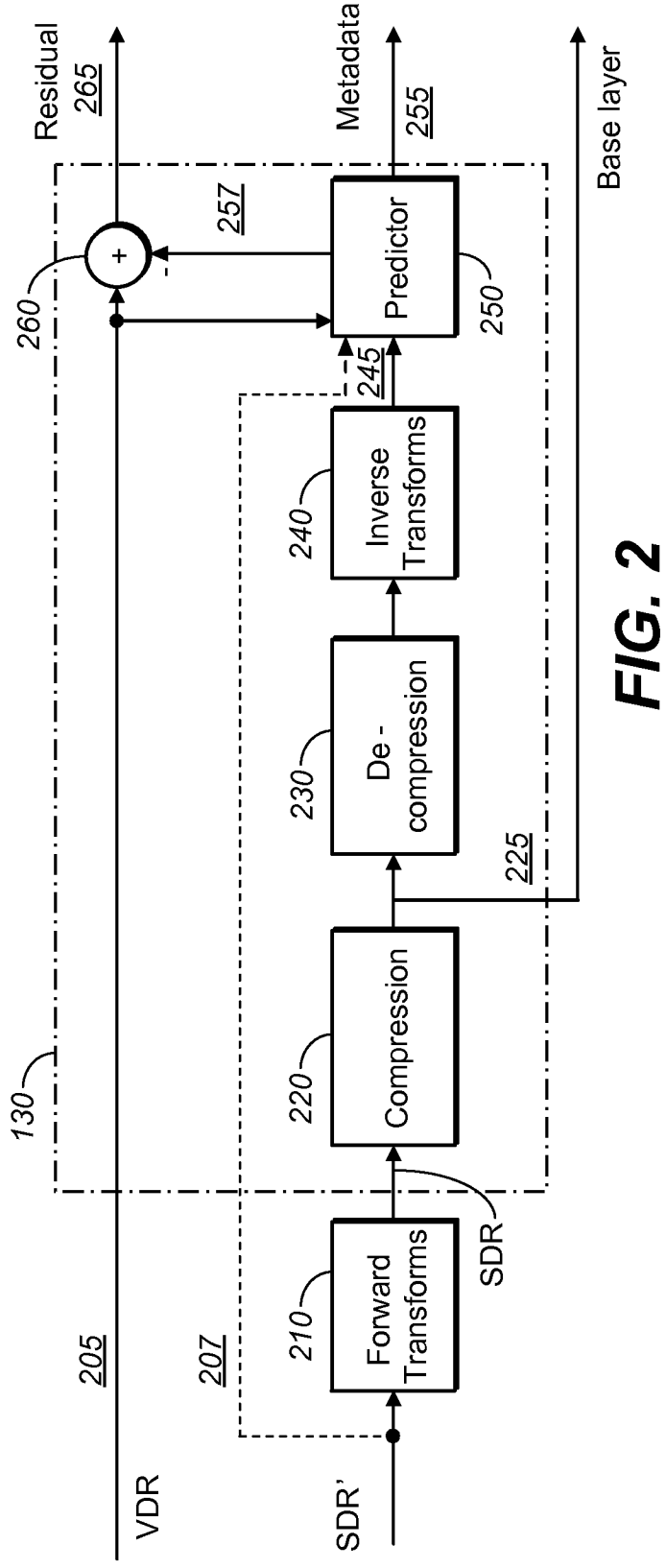

Residual
265
Metadata
255
Base layer
257
260
+
-
Predictor
250
245
Inverse Transforms
240
De - compression
230
225
Compression
220
Forward Transforms
210
SDR
130
205
207
VDR
SDR'

FIG. 2

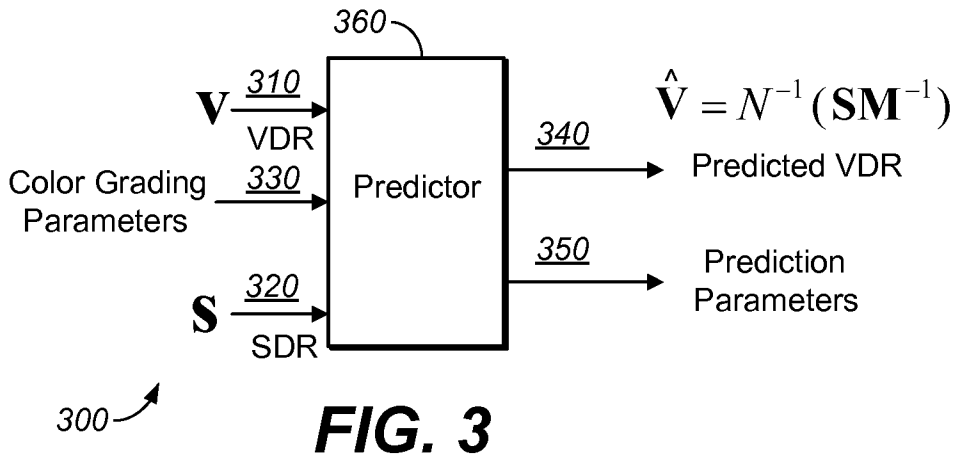

360
310
V
VDR
Color Grading Parameters
330
320
S
SDR
Predictor
340
$\hat{\mathbf{V}} = N^{-1}(\mathbf{SM}^{-1})$
Predicted VDR
350
Prediction Parameters
300

FIG. 3

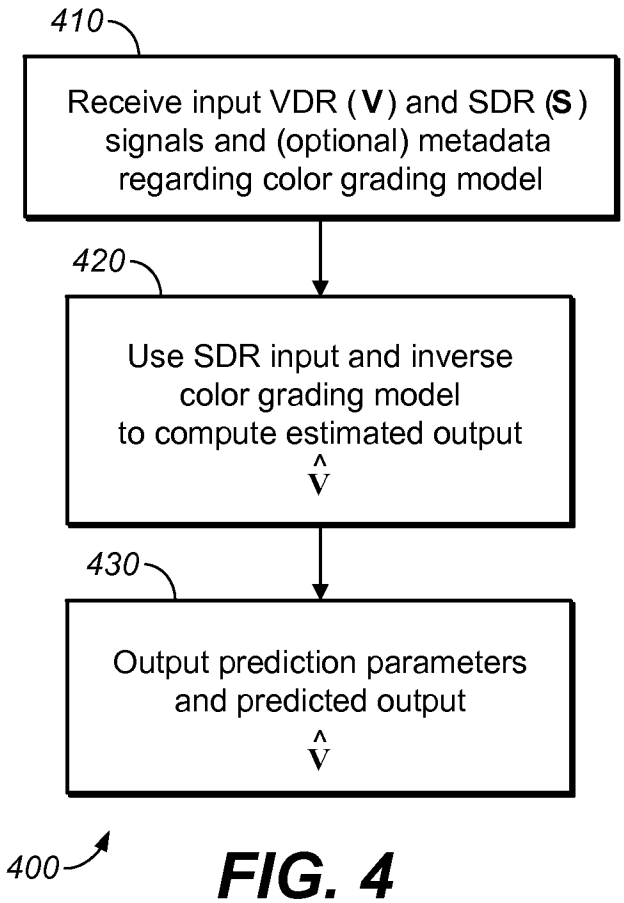

410
Receive input VDR (V) and SDR (S) signals and (optional) metadata regarding color grading model
420
Use SDR input and inverse color grading model to compute estimated output
$\hat{\mathbf{V}}$
430
Output prediction parameters and predicted output
$\hat{\mathbf{V}}$
400

FIG. 4

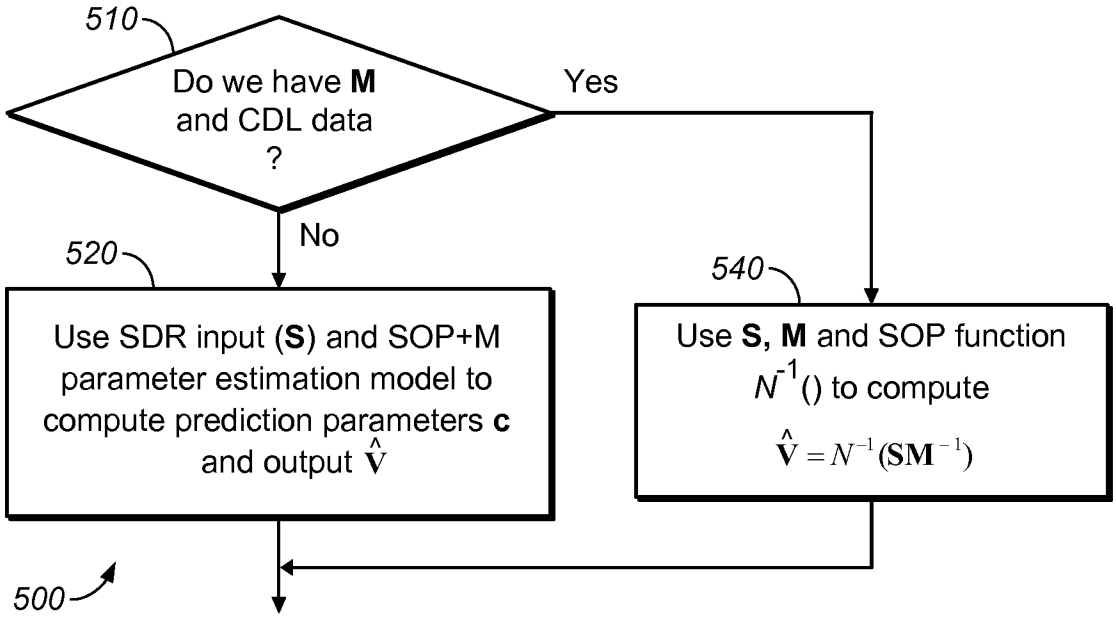

510
Do we have **M** and CDL data ?
Yes
No
520
Use SDR input (**S**) and SOP+M parameter estimation model to compute prediction parameters **c** and output $\hat{\mathbf{V}}$
540
Use **S**, **M** and SOP function $N^{-1}()$ to compute $\hat{\mathbf{V}} = N^{-1}(\mathbf{SM}^{-1})$
500

FIG. 5

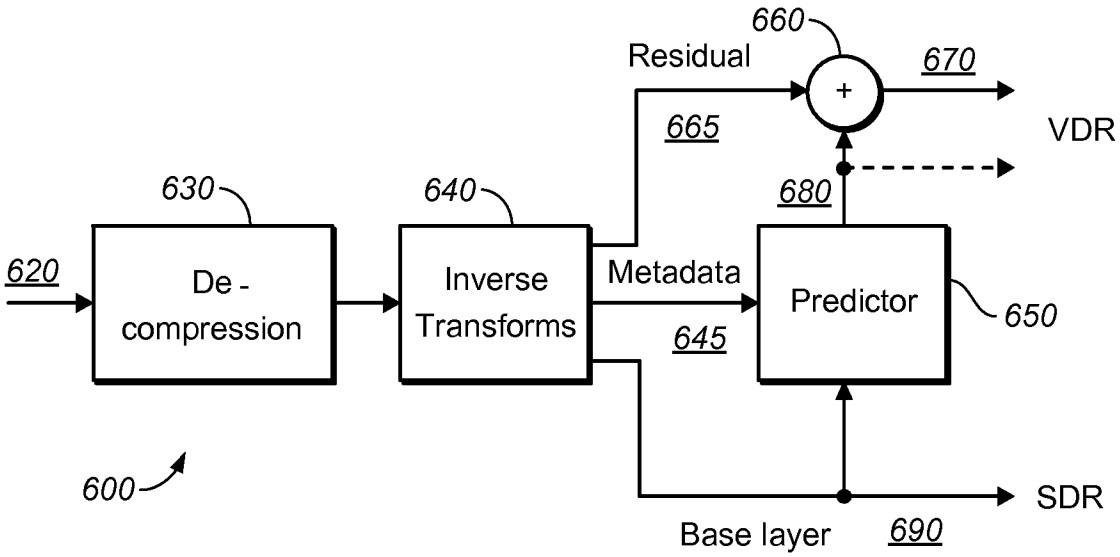

660
Residual
670
665
VDR
680
630
640
620
De - compression
Inverse Transforms
Metadata
Predictor
650
645
600
SDR
Base layer
690

FIG. 6

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2012/033651

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04N7/26
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	MARTIN WINKEN ET AL: "Bit-Depth Scalable Video Coding", IMAGE PROCESSING, 2007. ICIP 2007. IEEE INTERNATIONAL CONFERENCE ON, IEEE, PI, 1 September 2007 (2007-09-01), pages I-5, XP031157664, ISBN: 978-1-4244-1436-9 pages I-6; figure 1 -----	1-23
X	EP 1 827 024 A1 (SHARP KK [JP]) 29 August 2007 (2007-08-29) paragraph [0028] - paragraph [0036]; figure 2 ----- -/--	1-23

[X] Further documents are listed in the continuation of Box C.

[X] 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	Date of mailing of the international search report
13 July 2012	20/07/2012

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Oelbaum, Tobias

1

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2012/033651

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2010/105036 A1 (DOLBY LAB LICENSING CORP [US]; GISH WALTER C [US]; WEBB RICHARD W [US]) 16 September 2010 (2010-09-16) paragraph [0064] - paragraph [0084]; figure 7	1-23
X	EP 2 290 983 A2 (MAX PLANCK GESELLSCHAFT [DE]) 2 March 2011 (2011-03-02) abstract; figure 1 paragraph [0038] - paragraph [0052]	1-23
A	WO 2007/139534 A1 (THOMSON LICENSING [FR]; DOSER INGO TOBIAS [US]) 6 December 2007 (2007-12-06) paragraph [0024] - paragraph [0027]; figure 2	1-23
A	WO 2005/101309 A1 (UNIV BRITISH COLUMBIA [CA]; WHITEHEAD LORNE A [CA]; SEETZEN HELGE [CA]) 27 October 2005 (2005-10-27) abstract; figures 1A, 1B	1-23

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2012/033651

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 1827024 A1	29-08-2007	EP 1827024 A1	29-08-2007
		JP 4607136 B2	05-01-2011
		JP 2007243942 A	20-09-2007
		JP 2010213360 A	24-09-2010
		US 2007201560 A1	30-08-2007
WO 2010105036 A1	16-09-2010	CN 102388612 A	21-03-2012
		EP 2406959 A1	18-01-2012
		KR 20110115169 A	20-10-2011
		US 2011194618 A1	11-08-2011
		WO 2010105036 A1	16-09-2010
EP 2290983 A2	02-03-2011	CN 101371583 A	18-02-2009
		CN 101742306 A	16-06-2010
		EP 1989882 A2	12-11-2008
		EP 2290983 A2	02-03-2011
		EP 2320653 A2	11-05-2011
		JP 2009524371 A	25-06-2009
		KR 20080107389 A	10-12-2008
		US 2010172411 A1	08-07-2010
		WO 2007082562 A2	26-07-2007
WO 2007139534 A1	06-12-2007	CN 101449586 A	03-06-2009
		EP 2020146 A1	04-02-2009
		JP 2009538560 A	05-11-2009
		US 2009067506 A1	12-03-2009
		WO 2007139534 A1	06-12-2007
WO 2005101309 A1	27-10-2005	CA 2563520 A1	27-10-2005
		CN 1954339 A	25-04-2007
		CN 101901480 A	01-12-2010
		EP 1745436 A1	24-01-2007
		HK 1098563 A1	30-09-2010
		JP 4814217 B2	16-11-2011
		JP 2007534060 A	22-11-2007
		KR 20070004853 A	09-01-2007
		US 2007269104 A1	22-11-2007
		US 2008031517 A1	07-02-2008
		US 2012176399 A1	12-07-2012
		WO 2005101309 A1	27-10-2005