

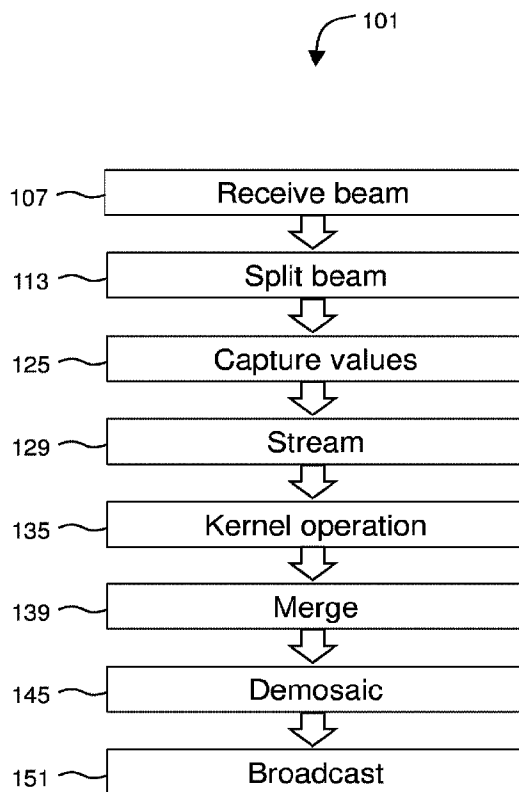


(86) Date de dépôt PCT/PCT Filing Date: 2017/02/10
(87) Date publication PCT/PCT Publication Date: 2017/08/17
(45) Date de délivrance/Issue Date: 2024/01/09
(85) Entrée phase nationale/National Entry: 2018/08/31
(86) N° demande PCT/PCT Application No.: US 2017/017396
(87) N° publication PCT/PCT Publication No.: 2017/139596
(30) Priorités/Priorities: 2016/02/12 (US62/294,820);
2016/05/31 (US15/169,006); 2016/10/17 (US62/409,053)

(51) Cl.Int./Int.Cl. *H04N 5/14* (2006.01),
G02B 27/10 (2006.01), *G06T 5/50* (2006.01)
(72) Inventeurs/Inventors:
KISER, WILLIE C., US;
TOCCI, NORA, US;
TOCCI, MICHAEL D., US
(73) Propriétaire/Owner:
CONTRAST, INC., US
(74) Agent: SMART & BIGGAR LP

(54) Titre : DISPOSITIFS ET PROCEDES POUR UNE VIDEO A PLAGE DYNAMIQUE ELEVEE

(54) Title: DEVICES AND METHODS FOR HIGH DYNAMIC RANGE VIDEO



(57) Abrégé/Abstract:

Systems and methods of the invention merge information from multiple image sensors to provide a high dynamic range (HDR) video. The present invention provides for real-time HDR video production using multiple sensors and pipeline processing techniques. According to the invention, multiple sensors with different exposures each produces an ordered stream of frame-independent pixel values. The pixel values are streamed through a pipeline on a processing device. The pipeline includes a kernel operation that identifies saturated ones of the pixel values. The streams of pixel values are merged to produce an HDR video.

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

(19) World Intellectual Property
Organization
International Bureau



(10) International Publication Number
WO 2017/139596 A1

(43) International Publication Date
17 August 2017 (17.08.2017)

- (51) International Patent Classification:
H04N 5/14 (2006.01) *G02B 27/10* (2006.01)
G06T 5/50 (2006.01)
- (21) International Application Number:
PCT/US2017/017396
- (22) International Filing Date:
10 February 2017 (10.02.2017)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
62/294,820 12 February 2016 (12.02.2016) US
15/169,006 31 May 2016 (31.05.2016) US
62/409,053 17 October 2016 (17.10.2016) US
- (71) Applicant: **CONTRAST OPTICAL DESIGN & ENGINEERING, INC.** [US/US]; 6100 Uptown Boulevard NE, Suite 560, Albuquerque, NM 87110 (US).
- (72) Inventors: **KISER, Willie, C.**; c/o Contrast Optical Design & Engineering, Inc., 6100 Uptown Boulevard NE,

Suite 560, Albuquerque, NM 87110 (US). **TOCCI, Nora**; c/o Contrast Optical Design & Engineering, Inc., 6100 Uptown Boulevard NE, Suite 560, Albuquerque, NM 87110 (US). **TOCCI, Michael, D.**; c/o Contrast Optical Design & Engineering, Inc., 6100 Uptown Boulevard NE, Suite 560, Albuquerque, NM 87110 (US).

(74) Agents: **MEYERS, Thomas C.** et al.; Brown Rudnick LLP, One Financial Center, Boston, MA 02111 (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, 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, 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.

[Continued on next page]

(54) Title: DEVICES AND METHODS FOR HIGH DYNAMIC RANGE VIDEO

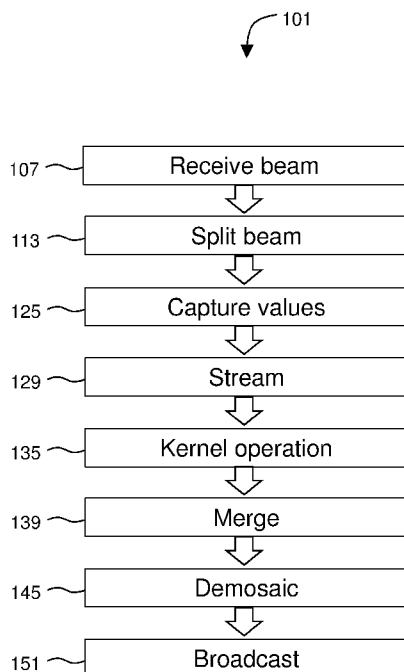


FIG. 1

(57) Abstract: Systems and methods of the invention merge information from multiple image sensors to provide a high dynamic range (HDR) video. The present invention provides for real-time HDR video production using multiple sensors and pipeline processing techniques. According to the invention, multiple sensors with different exposures each produces an ordered stream of frame-independent pixel values. The pixel values are streamed through a pipeline on a processing device. The pipeline includes a kernel operation that identifies saturated ones of the pixel values. The streams of pixel values are merged to produce an HDR video.

WO 2017/139596 A1

WO 2017/139596 A1

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE,

SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

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

DEVICES AND METHODS FOR HIGH DYNAMIC RANGE VIDEO

Cross-Reference to Related Applications

This application claims benefit of U.S. Provisional Application 62/409,053, filed October 17, 2016; and to U.S. Application 15/169,006, filed May 31, 2016, which application claims benefit of U.S. Provisional Application 62/294,820, filed February 12, 2016.

Technical Field

This disclosure relates to real-time capture and production of high dynamic range video.

Background

The human visual system is capable of identifying and processing visual features with high dynamic range. For example, real-world scenes with contrast ratios of 1,000,000:1 or greater can be accurately processed by the human visual cortex. However, most image acquisition devices are only capable of reproducing or capturing low dynamic range, resulting in a loss of image accuracy. The problem is ever more significant in video imaging.

There are examples of creating High Dynamic Range images by post processing images from multiple sensors, each subject to different exposures. The resulting “blended” image is intended to capture a broader dynamic range than would be possible from a single sensor without a post-processing operation.

Conventional approaches to capturing High Dynamic Range images include using sensors to acquire image frames of varying exposures and then processing those frames after acquisition (post-processing). Such approaches present significant challenges when filming video. If the subject, or camera, is moving then no pair of images will include exactly the same scene. In such cases, after the initial image capture, a post-processing step is required before a High Dynamic Range video can be produced.

Summary

Methods and devices of the invention process streams of pixels from multiple sensors in a frame-independent manner, resulting in real-time High Dynamic Range (HDR) video production. The real-time aspect of the invention is accomplished by analyzing streams of pixels from the various sensors without reference to the frame to which those pixels belong. Thus, the frame-independent nature of the invention means that there is no need to wait for an entire frame of data to be read from a sensor before processing pixel data. The result is an HDR video that has a dynamic range greater than the range that can be obtained using a single image sensor, typically 8 bits in depth.

The multiple image sensors used in the invention are exposed to identical scenes with different light levels and each produces an ordered stream of pixel values that are processed in real-time and independent of frame in which they will reside. The video processing pipeline used to produce HDR images includes kernel and merge operations that include identifying saturated pixel values and merging streams of pixel values. The merging operation includes replacing saturated pixel values with corresponding pixel values originating from a different sensor. The resulting product is high dynamic range video output.

Pixel values from multiple image sensors stream through the inventive pipeline and emerge as an HDR video signal without any need to aggregate or store any full images or frames of data as part of the pipeline. Because the kernel and merge operations process pixel values as they stream through the pipeline, no post-processing is required to produce HDR video. As a result, pixel values stream continually through the pipeline without waiting for a complete set of pixels from a sensor before being processed. For example, a CMOS sensor has millions of pixels, but the first pixel values to stream off of that sensor may be processed by the kernel operation or even enter the merge operation before a pixel value from, e.g., the millionth pixel from that same sensor even reaches the processing device. Thus, the frame-independent pixel processing systems and methods of the invention produce true real-time HDR video.

Additionally, a synchronization module in the pipeline synchronizes the streams of pixel values arriving from the sensors. This means that when, for example, the 60th pixel from a first sensor enters the kernel operation, the 60th pixel from each of the other sensors is also

simultaneously entering the kernel operation. As a result, pixel values from corresponding pixels on different sensors flow through the pipeline synchronously. This allows two things. First, the synchronization module can correct small phase discrepancies in data arrival times to the system from multiple sensors.). Second, the synchronization allows the kernel operation to consider—for a given pixel value from a specific pixel on one of the image sensors—values from the neighborhood of surrounding pixels on that sensor and also consider values from a corresponding neighborhood of pixels on another of the image sensors. This allows the kernel operation to create an estimated value for a saturated pixel from one sensor based on a pattern of values from the surrounding neighborhood on the same or another sensor.

The HDR pipeline can also correct for differences in spectral characteristics of each of the multiple sensors. Optical components such as beamsplitters, lenses, or filters—even if purported to be spectrally neutral—may have slight wavelength-dependent differences in the amounts of light transmitted. That is, each image sensor may be said to have its own “color correction space” whereby images from that sensor need to be corrected out of that color correction space to true color. The optical system can be calibrated (e.g., by taking a picture of a calibration card) and a color correction matrix can be determined and stored for each image sensor. The HDR video pipeline can then perform the counter-intuitive step of adjusting the pixel values from one sensor toward the color correction space of another sensor—which may in some cases involve nudging the colors away from true color. This may be accomplished by multiplying a vector of RGB values from the one sensor by the inverse of the color correction matrix of the other sensor. After this color correction to the second sensor, the streams are merged, and the resulting HDR video signal is color corrected to true color (e.g., by multiplying the RGB vectors by the applicable color correction matrix). This operation accounts for spectral differences of each image sensor.

A preferred pipeline includes other processing modules as described in detail in the Detailed Description of the invention below.

In certain aspects, the invention provides methods for producing real-time HDR video. Methods include streaming pixel values from each of multiple sensors in a frame independent-manner through a pipeline on a processing device, such as a field-programmable gate array or an application-specific integrated circuit. The pipeline includes a kernel operation that identifies saturated pixel values. The pixel values are merged to produce an HDR image. The merging step

excludes at least some of the saturated pixel values from the resulting HDR image. The multiple image sensors preferably capture images simultaneously through a single lens. Incoming light is received through the lens and split via at least one beamsplitter onto the multiple image sensors, such that at least about 95% of the light gathered by the lens is captured by the multiple image sensors. The multiple image sensors may include at least a high exposure (HE) sensor and a middle exposure (ME) sensor. Merging pixel streams may include using HE pixel values that are not saturated and ME pixel values corresponding to the saturated pixel values. The multiple sensors may further include a low exposure (LE) sensor, and the method may include identifying saturated pixel values originating from both the HE sensor and the ME sensor. The obtaining, streaming, and merging steps may include streaming the sequences of pixel values through a pipeline on the processing device such that no location on the processing device stores a complete image.

Sequences of pixel values are streamed through the processing device and merged without waiting to receive pixel values from all pixels on the image sensors. Because the pixel values are streamed through a pipeline, at least some of the saturated pixel values may be identified before receiving values from all pixels from the image sensors. Methods of the invention may include beginning to merge portions of the sequences while still streaming later-arriving pixel values through the kernel operation.

In some embodiments, streaming the pixel values through the kernel operation includes examining values from a neighborhood of pixels surrounding a first pixel on the HE sensor, finding saturated values in the neighborhood of pixels, and using information from a corresponding neighborhood on the ME sensor to estimate a value for the first pixel. The pixel values may be streamed through the kernel operation via a path within the processing device that momentarily stores a value from the first pixel proximal to each value originating from the neighborhood of pixels.

Each of the image sensors may include a color filter array, and the method may include demosaicing the HDR imaging after the merging. The multiple image sensors preferably capture images that are optically identical except for light level.

A first pixel value from a first pixel on one of the image sensors may be identified as saturated if it is above some specified level, for example at least about 90% of a maximum possible pixel value.

Aspects of the invention provide an apparatus for HDR video processing. The apparatus includes a processing device (e.g., FPGA or ASIC) and a plurality of image sensors coupled to the processing device. The apparatus is configured to stream pixel values from each of the plurality of image sensors in a frame independent-manner through a pipeline on the processing device. The pipeline includes a kernel operation that identifies saturated pixel values and a merge module to merge the pixel values to produce an HDR image. The apparatus may use or be connected to a single lens and one or more beamsplitters. The plurality of image sensors includes at least a high exposure (HE) sensor and a middle exposure (ME) sensor. The HE sensor, the ME sensor, the lens and the beamsplitters may be arranged to receive an incoming beam of light and split the beam of light into at least a first path that impinges on the HE sensor and a second path that impinges on the ME sensor. Preferably, the beamsplitter directs a majority of the light to the first path and a lesser amount of the light to the second path. In preferred embodiments, the first path and the second path impinge on the HE and the ME sensor, respectively, to generate images that are optically identical but for light level.

The kernel operation may operate on pixel values as they stream from each of the plurality of image sensors by examining, for any given pixel on the HE sensor, values from a neighborhood of pixels surrounding the given pixel, finding saturated values in the neighborhood of pixels, and using information from a corresponding neighborhood on the ME sensor to estimate a value for the given pixel. The pipeline may include—in the order in which the pixel values flow: a sync module to synchronize the pixel values as the pixel values stream onto the processing device from the plurality of image sensors; the kernel operation; the merge module; a demosaicing module; and a tone-mapping operator as well as optionally one or more of a color-conversion module; an HDR conversion module; and an HDR compression module. In certain embodiments, the apparatus also includes a low exposure (LE) sensor. Preferably, the apparatus includes a sync module on the processing device, in the pipeline upstream of the kernel operation, to synchronize the pixel values as the pixel values stream onto the processing device from the plurality of image sensors.

According to one aspect of the present invention, there is provided a method for producing real-time HDR video, the method comprising: receiving incoming light through a lens and splitting the light via at least one beamsplitter onto a first image sensor and a second image sensor, wherein the first sensor receives a higher exposure than the second sensor, and wherein at least 95% of the light gathered by the imaging lens is captured by the multiple image sensors; streaming pixel values from each of multiple sensors in a frame independent-manner through a pipeline on a processing device, wherein the pipeline includes a kernel operation that identifies saturated pixel values; and merging the pixel values to produce an HDR image, wherein the merging includes using first sensor pixel values that are not saturated and second sensor pixel values that correspond to the identified saturated pixel values.

According to another aspect of the present invention, there is provided an apparatus for HDR video processing, the apparatus comprising: a processing device; a first image sensor and a second image sensor coupled to the processing device; and a lens and at least one beamsplitter arranged to receive an incoming beam of light and split the beam of light into at least a first path that impinges on the first image sensor and a second path that impinges on the second image sensor, wherein the first sensor receives a higher exposure than the second sensor, wherein the apparatus is configured to stream pixel values from each of the plurality of image sensors in a frame- independent manner through a pipeline on the processing device, wherein the pipeline includes a kernel operation that identifies saturated pixel values, wherein the kernel operation operates on pixel values as they stream from the first and second image sensors by examining, for a given pixel on the first sensor, values from a neighborhood of pixels surrounding the given pixel, finding saturated values in the neighborhood of pixels, and using information from a corresponding neighborhood on the second sensor to estimate a value for the given pixel and a merge module to merge the pixel values to produce an HDR image.

Brief Description of the Drawings

FIG. 1 shows steps of a method for producing real-time HDR video.

FIG. 2 shows an apparatus for HDR video processing.

FIG. 3 shows an arrangement for multiple sensors.

FIG. 4 shows a processing device on a real-time HDR video apparatus.

FIG. 5 shows operation of a sync module.

FIG. 6 illustrates how pixel values are presented to a kernel operation.

FIG. 7 shows an approach to modeling a pipeline.

FIG. 8 illustrates merging to avoid artifacts.

FIG. 9 shows a camera response curve used to adjust a pixel value.

FIG. 10 shows a color correction processes.

FIG. 11 illustrates a method for combined HDR broadcasting.

FIG. 12 shows image sensors and processor of an apparatus for HDR video processing according to certain embodiments.

Detailed Description

FIG. 1 shows steps of a method 101 for producing real-time HDR video. The method 101 includes receiving 107 light through the lens of an imaging apparatus. One or more beamsplitters split 113 the light into different paths that impinge upon multiple image sensors. Each image sensor then captures 125 a signal in the form of a pixel value for each pixel of the sensor. Where the sensor has, say, 1920×1080 pixels, the pixel values will stream off of the sensor to a connected processing device. The method includes streaming 129 pixel values 501 from each of multiple sensors in a frame independent-manner through a pipeline 231 on a processing device 219. The pipeline 231 includes a kernel operation 135 that identifies saturated pixel values. The pixel values 501 are merged 139. Typically, the merged image will be demosaiced 145 and this produces an HDR image that can be displayed, transmitted, stored, or broadcast 151. In the described method 101, the multiple image sensors all capture 125 images simultaneously through a single lens 311. The pipeline 231 and kernel operation 135 may be provided by an integrated circuit such as a field-programmable gate array or an application-specific integrated circuit. Each of the image sensors may include a color filter array 307. In preferred embodiments, the method 101 includes demosaicing 145 the HDR image after the merging step 139. The multiple image sensors preferably capture images that are optically identical except for light level.

A feature of the invention is that the pixel values 501 are pipeline processed in a frame-independent manner. Sequences of pixel values 501 are streamed 129 through the processing

device 219 and merged 139 without waiting to receive pixel values 501 from all pixels on the image sensors. This means that the obtaining 125, streaming 129, and merging 139 steps may be performed by streaming 129 the sequences of pixel values 501 through the pipeline 231 on the processing device 219 such that no location on the processing device 219 stores a complete image. Because the pixel values are streamed through the pipeline, the final HDR video signal is produced in real-time. An apparatus 201 performing steps of the method 101 thus provides the function of a real-time HDR video camera. Real-time means that HDR video from the camera may be displayed essentially simultaneously as the camera captures the scene (e.g., at the speed that the signal travels from sensor to display minus a latency no greater than a frame of film). There is no requirement for post-processing the image data and no requirement to capture, store, compare, or process entire “frames” of images.

The output is an HDR video signal because the method 101 and the apparatus 201 use multiple sensors at different exposure levels to capture multiple isomorphic images (i.e., identical but for light level) and merge them. Data from a high exposure (HE) sensor are used where portions of an image are dim and data from a mid-exposure (ME) (or lower) sensor(s) are used where portions of an image are more brightly illuminated. The method 101 and apparatus 201 merge the HE and ME (and optionally LE) images to produce an HDR video signal. Specifically, the method 101 and the apparatus 201 identify saturated pixels in the images and replace those saturated pixels with values derived from sensors of a lower exposure. In preferred embodiments, a first pixel value from a first pixel on one of the image sensors is identified as saturated if it is above some specified level, for example at least 90% of a maximum possible pixel value.

FIG. 2 shows an apparatus 201 for HDR video processing. The apparatus 201 includes a processing device 219 such as a field-programmable gate array (FPGA) or an application-specific integrated circuit (ASIC). A plurality of image sensors 265 are coupled to the processing device 219. The apparatus 201 is configured to stream pixel values 501 from each of the plurality of image sensors 265 in a frame independent-manner through a pipeline 231 on the processing device 219. The pipeline 231 includes a kernel operation 413 that identifies saturated pixel values 501 and a merge module to merge the pixel values 501 to produce an HDR image.

The kernel operation 413 operates on pixel values 501 as they stream from each of the plurality of image sensors 265 by examining, for a given pixel on the HE sensor 213, values

from a neighborhood 601 of pixels surrounding the given pixel, finding saturated values in the neighborhood 601 of pixels, and using information from a corresponding neighborhood 601 on the ME sensor 211 to estimate a value for the given pixel.

Various components of the apparatus 201 may be connected via a printed circuit board 205. The apparatus 201 may also include memory 221 and optionally a processor 227 (such as a general-purpose processor like an ARM microcontroller). Apparatus 201 may further include or be connected to one or more of an input-output device 239 or a display 267. Memory can include RAM or ROM and preferably includes at least one tangible, non-transitory medium. A processor may be any suitable processor known in the art, such as the processor sold under the trademark XEON E7 by Intel (Santa Clara, CA) or the processor sold under the trademark OPTERON 6200 by AMD (Sunnyvale, CA). Input/output devices according to the invention may include a video display unit (e.g., a liquid crystal display or LED display), keys, buttons, a signal generation device (e.g., a speaker, chime, or light), a touchscreen, an accelerometer, a microphone, a cellular radio frequency antenna, port for a memory card, and a network interface device, which can be, for example, a network interface card (NIC), Wi-Fi card, or cellular modem. The apparatus 201 may include or be connected to a storage device 241. The plurality of sensors are preferably provided in an arrangement that allows multiple sensors 265 to simultaneously receive images that are identical except for light level.

FIG. 3 shows an arrangement for the multiple sensors 265. The multiple sensors preferably include at least a high exposure (HE) sensor 213 and a middle exposure (ME) sensor 211. Each image sensor may have its own color filter array 307. The color filter arrays 307 may operate as a Bayer filter such that each pixel receives either red, green, or blue light. As is known in the art, a Bayer filter includes a repeating grid of red, green, blue, green filters such that a sequence of pixel values streaming from the sensor corresponds to values for red, green, blue, green, red, green, blue, green, red, green, blue, green,...etc.

As shown in FIG. 3, the apparatus 201 may also include or be optically connected to a lens 311 and at least one beamsplitter 301. The HE sensor 213, the ME sensor 211, the lens 311 and the at least one beamsplitter 301 are arranged to receive an incoming beam of light 305 and split the beam of light 305 into at least a first path that impinges on HE sensor 213 and a second path that impinges on the ME sensor 211. In a preferred embodiment, the apparatus 201 uses a set of partially-reflecting surfaces to split the light from a single photographic lens 311 so that it

is focused onto three imaging sensors simultaneously. In a preferred embodiment, the light is directed back through one of the beamsplitters a second time, and the three sub-images are not split into red, green, and blue but instead are optically identical except for their light levels. This design, shown in FIG. 3, allows the apparatus to capture HDR images using most of the light entering the camera.

In some embodiments, the optical splitting system uses two uncoated, 2-micron thick plastic beamsplitters that rely on Fresnel reflections at air/plastic interfaces so their actual transmittance/reflectance (T/R) values are a function of angle. Glass is also a suitable material option. In one embodiment, the first beamsplitter 301 is at a 45° angle and has an approximate T/R ratio of 92/8, which means that 92% of the light from the camera lens 311 is transmitted through the first beamsplitter 301 and focused directly onto the high-exposure (HE) sensor 213. The beamsplitter 301 reflects 8% of the light from the lens 311 upwards (as shown in FIG. 3), toward the second uncoated beamsplitter 319, which has the same optical properties as the first but is positioned at a 90° angle to the light path and has an approximate T/R ratio of 94/6.

Of the 8% of the total light that is reflected upwards, 94% (or 7.52% of the total light) is transmitted through the second beamsplitter 319 and focused onto the medium-exposure (ME) sensor 211. The other 6% of this upward-reflected light (or 0.48% of the total light) is reflected back down by the second beamsplitter 319 toward the first beamsplitter 301 (which is again at 45°), through which 92% (or 0.44% of the total light) is transmitted and focused onto the low-exposure (LE) sensor 261. With this arrangement, the HE, ME and LE sensors capture images with 92%, 7.52%, and 0.44% of the total light gathered by the camera lens 311, respectively. Thus a total of 99.96% of the total light gathered by the camera lens 311 has been captured by the image sensors. Therefore, the HE and ME exposures are separated by $12.2\times$ (3.61 stops) and the ME and LE are separated by $17.0\times$ (4.09 stops), which means that this configuration is designed to extend the dynamic range of the sensor by 7.7 stops.

This beamsplitter arrangement makes the apparatus 201 light efficient: a negligible 0.04% of the total light gathered by the lens 311 is wasted. It also allows all three sensors to “see” the same scene, so all three images are optically identical except for their light levels. Of course, in the apparatus of the depicted embodiment 201, the ME image has undergone an odd number of reflections and so it is flipped left-right compared to the other images, but this is fixed easily in software. In preferred embodiments, the three sensors independently stream incoming

pixel values directly into a pipeline that includes a synchronization module. This synchronization module can correct small phase discrepancies in data arrival times to the system from multiple sensors.

Thus it can be seen that the beamsplitter 301 directs a majority of the light to the first path and a lesser amount of the light to the second path. Preferably, the first path and the second path impinge on the HE sensor 213 and the ME sensor 211, respectively, to generate images that are optically identical but for light level. In the depicted embodiment, the apparatus 201 includes a low exposure (LE) sensor.

In preferred embodiments, pixel values stream from the HE sensor 213, the ME sensor 211, and the LE sensor 261 in sequences directly to the processing device 219. Those sequences may be not synchronized as they arrive onto the processing device 219.

As shown by FIG. 3, the method 101 may include receiving 107 incoming light through the lens 311 and splitting 113 the light via at least one beamsplitter 301 onto the multiple image sensors, wherein at least 95% of the incoming beam of light 305 is captured by the multiple image sensors.

The apparatus 201 (1) captures optically-aligned, multiple- exposure images simultaneously that do not need image manipulation to account for motion, (2) extends the dynamic range of available image sensors (by over 7 photographic stops in our current prototype), (3) is inexpensive to implement, (4) utilizes a single, standard camera lens 311, and (5) efficiently uses the light from the lens 311.

The method 101 preferably (1) combines images separated by more than 3 stops in exposure, (2) spatially blends pre-demosaiced pixel data to reduce unwanted artifacts, (3) produces HDR images that are radiometrically correct, and (4) uses the highest-fidelity (lowest quantized-noise) pixel data available. The apparatus 201 can work with a variety of different sensor types and uses an optical architecture based on beamsplitters located between the camera lens and the sensors.

FIG. 4 shows the processing device 219 on the apparatus 201. As noted, the processing device 219 may be provided by one or more FPGA, ASIC, or other integrated circuit. Pixel values from the sensors stream through the pipeline 231 on the processing device 219. The pipeline 231 in the processing device 219 includes—in the order in which the pixel values 501 flow: a sync module 405 to synchronize the pixel values 501 as the pixel values 501 stream onto

the processing device 219 from the plurality of image sensors 265; the kernel operation 413; the merge module 421; a demosaicing module 425; and a tone-mapping operator 427. The pipeline 231 may include one or more auxiliary module 431 such as a color-correction module; an HDR conversion module; and an HDR compression module.

FIG. 5 shows operation of the sync module 405 to synchronize the pixel values 501 as the pixel values 501 stream onto the processing device 219 from the plurality of image sensors 265. As depicted in FIG. 5, HE_1 pixel value and ME_1 pixel value are arriving at the sync module 405 approximately simultaneously. However, HE_2 pixel value will arrive late compared to ME_2, and the entire sequence of LE pixel values will arrive late. The sync module 405 can contain small line buffers that circulate the early-arriving pixel values and release them simultaneous with the corresponding later-arriving pixel values. The synchronized pixel values then stream through the pipeline 231 to the kernel operation 413.

FIG. 6 illustrates how the pixel values are presented to the kernel operation 413. The top part of FIG. 6 depicts the HE sensor 213. Each square depicts one pixel of the sensor 213. A heavy black box with a white center is drawn to illustrate a given pixel 615 for consideration and a neighborhood 601 of pixels surrounding the given pixel 615. The heavy black box would not actually appear on a sensor 213 (such as a CMOS cinematic camera sensor)—it is merely drawn to illustrate what the neighborhood 601 includes and to aid understanding how the neighborhood 601 appears when the sequences 621 of pixel values 501 are presented to the kernel operation 413.

The bottom portion of FIG. 6 shows the sequences 621 of pixel values as they stream into the kernel operation 413 after the sync module 405. Pixel values 501 from the neighborhood 601 of pixels on the sensor 213 are still “blackened out” to aid illustration. The given pixel 615 under consideration can be spotted easily because it is surrounded on each side by two black pixels from the row of pixels on the sensor. There are two sequences 621, one of which comes from the depicted HE sensor 213 and one of which originates at the ME sensor 211.

Streaming the pixel values 501 through the kernel operation 413 includes examining values from a neighborhood 601 of pixels surrounding a first pixel 615 on the HE sensor 213, finding saturated values in the neighborhood 601 of pixels, and using information from a corresponding neighborhood 613 from the ME sensor 211 to estimate a value for the first pixel 615. This will be described in greater detail below. To accomplish this, the processing device

must make comparisons between corresponding pixel values from different sensors. It may be useful to stream the pixel values through the kernel operation in a fashion that places the pixel under consideration 615 adjacent to each pixel from the neighborhood 601 as well as adjacent to each pixel from the corresponding neighborhood on another sensor.

FIG. 7 shows an approach to modeling the circuit so that the pipeline places the current pixel 615 adjacent to each of the following pixel values: a pixel value from 1 to the right on the sensor 213, a pixel value from 2 pixels to the right on sensor 213, a pixel value from 1 to the left, and pixel value from two to the left. As shown in FIG. 7, data flows into this portion of the pipeline and is copied four additional times. For each copy, a different and specific amount of delay is added to the main branch. The five copies all continue to flow in parallel. Thus, a simultaneous snapshot across all five copies covers the given current pixel value 615 and the other pixel values from the neighborhood 601. In this way, pixel values on either side of the pixel currently being processed can be used in that processing step, along with the pixel currently being processed. Thus the processing device can simultaneously read and compare the pixel value of the given pixel to the value of neighbors. The approach illustrated in FIG. 7 can be extended for comparisons to upper and lower neighbors, diagonal neighbors, and pixel values from a corresponding neighborhood on another sensor. Thus in some embodiments, streaming 129 the pixel values 501 through the kernel operation 413 includes streaming 129 the pixel values 501 through a path 621 within the processing device 219 that momentarily places a value from the first pixel proximal to each value originating from the neighborhood 601 of pixels.

The neighborhood comparisons may be used in determining whether to use a replacement value for a saturated pixel and what replacement value to use. An approach to using the neighborhood comparisons is discussed further down after a discussion of the merging. A replacement value will be used when the sequences 621 of pixel values 501 are merged 139 by the merge module 421. The merging 139 step excludes at least some of the saturated pixel values 501 from the HDR image.

Previous algorithms for merging HDR images from a set of LDR images with different exposures typically do so after demosaicing the LDR images and merge data pixel-by-pixel without taking neighboring pixel information into account.

To capture the widest dynamic range possible with the smallest number of camera sensors, it is preferable to position the LDR images further apart in exposure than with traditional

HDR acquisition methods. Prior art methods yield undesired artifacts because of quantization and noise effects, and those problems are exacerbated when certain tone mapping operators (TMOs) are applied. Those TMOs amplify small gradient differences in the image to make them visible when the dynamic range is compressed, amplifying merging artifacts as well.

FIG. 8 illustrates an approach to merging that reduces artifacts (e.g., compared to the weighting factor used in a merging algorithm in Debevec and Malik, 1997, Recovering high dynamic range radiance maps from photographs, Proceedings of ACM SIGGRAPH 1997:369–378, incorporated by reference). The “HE sensor”, “ME sensor”, and “LE sensor” bars in FIG. 8 present the range of scene illumination measured by the three sensors

For illustration, the system is simplified with 4-bit sensors (as opposed to the 12-bit sensors as may be used in apparatus 201), which measure only 16 unique brightness values and the sensors are separated by only 1 stop (a factor of 2) in exposure. Since CMOS sensors exhibit an approximately linear relationship between incident exposure and their output value, the values from the three sensors are graphed as a linear function of incident irradiance instead of the traditional logarithmic scale.

Merging images by prior art algorithms that always use data from all three sensors with simple weighting functions, such as that of Debevec and Malik, introduces artifacts. In the prior art, data from each sensor is weighted with a triangle function as shown by the dotted lines, so there are non-zero contributions from the LE sensor at low brightness values (like the sample illumination level indicated), even though the data from the LE sensor is quantized more coarsely than that of the HE sensor.

Methods of the invention, in contrast, use data from the higher-exposure sensor as much as possible and blend in data from the next darker sensor when near saturation.

FIG. 8 shows that the LE sensor measures the scene irradiance more coarsely than the other two sensors. For example, the HE sensor may measure 4 different pixel values in a gradient before the LE sensor records a single increment. In addition, there is always some small amount of noise in the pixel values, and an error of ± 1 in the LE sensor spans a 12 value range in the HE sensor for this example. Although Debevec and Malik’s algorithm blends these values together, the method 101 and apparatus 201 use pixel values from only the longest-exposure sensor (which is less noisy) wherever possible, and blend in the next darker exposure when pixels approach saturation.

In certain embodiments, the method 101 and apparatus 201 not only examine individual pixels when merging the LDR images, but also take into account neighboring pixels 601 (see FIG. 6) that might provide additional information to help in the de-noising process.

One aspect of merging 139 according to the invention is to use pixel data exclusively from the brightest, most well-exposed sensor possible. Therefore, pixels from the HE image are used as much as possible, and pixels in the ME image are only used if the HE pixel is close to saturation. If the corresponding ME pixel is below the saturation level, it is multiplied by a factor that adjusts it in relation to the HE pixel based on the camera's response curve, given that the ME pixel receives $12.2\times$ less irradiance than the HE pixel.

FIG. 9 shows a camera response curve 901 used to obtain a factor for adjusting a pixel value. In a three-sensor embodiment, when the HE sensor is above the saturation level, and if the corresponding ME pixel is above the saturation level, then a similar process is applied to the same pixel in the low-exposure LE image.

It may be found that merging by a "winner take all" approach that exclusively uses the values from the HE sensor until they become saturated and then simply switches to the next sensor results in banding artifacts where transitions occur. To avoid such banding artifacts, the method 101 and apparatus 201 transition from one sensor to the next by spatially blending pixel values between the two sensors. To do this, the method 101 and apparatus 201 scan a neighborhood 601 around the pixel 615 being evaluated (see FIG. 6). If any neighboring pixels in this region are saturated, then the pixel under consideration may be subject to pixel crosstalk or leakage, and the method 101 and apparatus 201 will estimate a value for the pixel based on its neighbors in the neighborhood 601.

The method 101 and apparatus 201 perform merging 139 prior to demosaicing 145 the individual Bayer color filter array images because demosaicing can corrupt colors in saturated regions. For example, a bright orange section of a scene might have red pixels that are saturated while the green and blue pixels are not. If the image is demosaiced before being merged into HDR, the demosaiced orange color will be computed from saturated red-pixel data and non-saturated green/blue-pixel data. As a result, the hue of the orange section will be incorrectly reproduced. To avoid these artifacts, the method 101 and apparatus 201 perform HDR-merging prior to demosaicing.

Since the images are merged prior to the demosaicing step, the method 101 and apparatus 201 work with pixel values instead of irradiance. To produce a radiometrically-correct HDR image, the method 101 and apparatus 201 match the irradiance levels of the HE, ME, and LE sensors using the appropriate beamsplitter transmittance values for each pixel color, since these change slightly as a function of wavelength. Although the method 101 and apparatus 201 use different values to match each of the color channels, for simplicity the process is explained with average values. A pixel value is converted through the camera response curve 901, where the resulting irradiance is adjusted by the exposure level ratio (average of $12.2\times$ for HE/ME), and this new irradiance value is converted back through the camera response curve 901 to a new pixel value.

FIG. 9 shows the 3-step HDR conversion process to match the irradiance levels of the HE, ME, and LE sensors. The HDR conversion process may be done for all HE pixel values (from 1 through 4096, for example), to arrive at a pixel-ratio curve, which gives the scaling factor for converting each ME pixel's value to the corresponding pixel value on the HE sensor for the same irradiance. In practice, separate pixel-ratio curves are calculated for each color (R,G,B) in the Bayer pattern. When comparing pixel values between HE and ME images (or between ME and LE images), a simple multiplier may be used, or the pixel-ratio curves may be used as lookup tables (LUTs), to convert HE pixel values less than 4096 into ME pixel values, or vice versa. When the HE pixel values are saturated, the pixel-ratio curve is extended using the last value obtained there (approximately 8).

The camera response curve 901 can be measured by taking a set of bracketed exposures and solving for a monotonically-increasing function that relates exposure to pixel value (to within a scale constant in the linear domain).

FIG. 9 shows the curve computed from the raw camera data, although a curve computed from a linear best-fit could also be used.

FIG. 9 gives a camera response curve that shows how the camera converts scene irradiance into pixel values. To compute what the ME pixel value should be for a given HE value, the HE pixel value (1) is first converted to a scene irradiance (2), which is next divided by our HE/ME attenuation ratio of 12.2. This new irradiance value (3) is converted through the camera response curve into the expected ME pixel value (4). Although this graph is approximately linear, it is not perfectly so because it is computed from the raw data, without

significant smoothing or applying a linear fit. With the irradiance levels of the three images matched, the merging 139 may be performed.

In an illustrative example of merging 139, two registered LDR images (one high-exposure image IHE and a second medium-exposure image IME) are to be merged 139 into an HDR image IHDR . The merging 139 starts with the information in the high-exposure image IHE and then combines in data from the next darker-exposure image IME, as needed. To reduce the transition artifacts described earlier, the method 101 and apparatus 201 work on each pixel location (x, y) by looking at the information from the surrounding $(2k + 1) \times (2k + 1)$ pixel neighborhood 601, denoted as $N(x,y)$.

In some embodiments as illustrated in FIG. 6, the method 101 and apparatus 201 use a 5×5 pixel neighborhood 601 ($k = 2$), and define a pixel to be saturated if its value is greater than some specific amount, for example 90% of the maximum pixel value (4096 e.g., where sensor 213 is a 12-bit CMOS sensor).

In certain embodiments, the merging 139 includes a specific operation for each of the four cases for the pixel 615 on sensor 213 and its neighborhood 601 (see FIG. 6):

Case 1: The pixel 615 is not saturated and the neighborhood 601 has no saturated pixels, so the pixel value is used as-is.

Case 2: The pixel 615 is not saturated, but the neighborhood 601 has 1 or more saturated pixels, so blend between the pixel value at $IHE(x, y)$ and the one at the next darker-exposure $IME(x, y)$ depending on the amount of saturation present in the neighborhood.

Case 3: The pixel 615 is saturated but the neighborhood 601 has 1 or more non- saturated pixels, which can be used to better estimate a value for $IHE(x,y)$: calculate the ratios of pixel values in the ME image between the unsaturated pixels in the neighborhood and the center pixel, and use this map of ME ratios to estimate the actual value of the saturated pixel under consideration.

Case 4: The pixel 615 is saturated and all pixels in the neighborhood 601 are saturated, so there is no valid information from the high- exposure image, use the ME image and set $IHDR(x, y) = IME(x, y)$.

When there are three LDR images, the process above is simply repeated in a second iteration, substituting IHDR for IHE and ILE for IME. In this manner, data is merged 139 from

the higher exposures while working toward the lowest exposure, and data is only used from lower exposures when the higher-exposure data is at or near saturation.

This produces an HDR image that can be demosaiced 145 and converted from pixel values to irradiance using a camera response curve similar to that of FIG. 9 accounting for all 3 color channels. The final HDR full-color image may then be tone mapped (e.g., with commercial software packages such as FDRTools, HDR Expose, Photomatix, etc.)

The apparatus 201 may be implemented using three Silicon Imaging SI-1920HD high-end cinema CMOS sensors mounted in a camera body. Those sensors have 1920×1080 pixels (5 microns square) with a standard Bayer color filter array, and can measure a dynamic range of around 10 stops (excluding noise). The sensors are aligned by aiming the camera at small pinhole light sources, locking down the HE sensor and then adjusting setscrews to align the ME and LE sensors.

The camera body may include a Hasselblad lens mount to allow the use of high-performance, interchangeable commercial lenses. For beamsplitters, the apparatus may include uncoated pellicle beamsplitters, such as the ones sold by Edmund Optics [part number NT39-482]. The apparatus 201 may perform the steps of the method 101. Preferably, the multiple image sensors include at least a high exposure (HE) sensor 213 and a middle exposure (ME) sensor 211, and the merging includes using HE pixel values 501 that are not saturated and ME pixel values 501 corresponding to the saturated pixel values. The multiple sensors may further include a low exposure (LE) sensor 261, and the method 101 may include identifying saturated pixel values 501 originating from both the HE sensor 213 and the ME sensor 211. Because the pixel values stream through a pipeline, it is possible that at least some of the saturated pixel values 501 are identified before receiving values from all pixels of the multiple image sensors at the processing device 219 and the method 101 may include beginning to merge 139 portions of the sequences while still streaming 129 later-arriving pixel values 501 through the kernel operation 413.

It is understood that optical components such as beamsplitters, lenses, or filters—even if labeled “spectrally neutral”—may have slight wavelength-dependent differences in the amounts of light transmitted. That is, each image sensor may be said to have its own “color correction space” whereby images from that sensor need to be corrected out of that color correction space to true color. The optical system can be calibrated (e.g., by taking a picture of a calibration card)

and a color correction matrix can be stored for each image sensor. The HDR video pipeline can then perform the counter-intuitive step of adjusting the pixel values from one sensor towards the color correction of another sensor—which may in some cases involve nudging the colors away from true color. This may be accomplished by multiplying a vector of RGB values from the one sensor by the inverse color correction matrix of the other sensor. After this color correction to the second sensor, the streams are merged, and the resulting HDR video signal is color corrected to truth (e.g., by multiplying the RGB vectors by the applicable color correction matrix). This color correction process accounts for spectral differences of each image sensor.

FIG. 10 shows a color correction processes 1001 by which the HDR pipeline can correct for differences in spectral characteristics of each of the multiple sensors. To correct for the slight wavelength-dependent differences among the sensors, relationships between electron input and electron output can be measured experimentally using known inputs. By computing a correction factor for each sensor, the information detected by the sensors can be corrected prior to further processing. Thus, in some embodiments, the pipeline 231 includes modules for color correction. The steps of a color correction process may be applied at multiple locations along the pipeline, so the color correction may be implemented via specific modules at different locations on the FPGA. Taken together, those modules may be referred to as a color correction module that implements the color correction process 1001.

The color correction process 1001 converts one sensor's data from its color correction space to the color correction space of another sensor, before merging the images from the two sensors. The merged image data can then be converted to the color correction space of a third sensor, before being combined with the image data from that third sensor. The process may be repeated for as many sensors as desired. After all sensors' images have been combined, the final combined image may be demosaiced 145 and then may be color corrected to truth.

The color correction process 1001 allows images from multiple sensors to be merged, in stages where two images are merged at a time, in a way that preserves color information from one sensor to the next. For example purposes, in FIG. 10, the HE pixel values from the HE sensor are merged with the ME pixel values from the ME sensor. The result of merging is then merged with the LE pixel values from the LE sensor.

The basic principle guiding the color correction process 1001 is to first convert a dark image to the color correction space of the next brightest image, and then to merge the two “non-demosaiced” (or Color Filter Array [CFA] Bayer-patterned) images together.

The color correction process 1001, for an apparatus 201 with an ME sensor, an LE sensor, and an SE sensor, includes three general phases: an SE color correction space (CCS) phase, ME color correction space phase, and LE color correction space phase. The color correction process first begins with the SE color correction space phase, which comprises first demosaicing 1045 the LE pixel values and then transforming 1051 the resulting vectors into the color correction space of the ME image. The demosaicing process 1045 yields a full-color RGB vector value for each pixel.

After it has been demosaiced 1045, the LE image data is next transformed 1045 into the ME color correction space. The purpose is to match the color of the LE pixels (now described by RGB vectors) to the color of the ME array (with all of the ME array’s color imperfections). To perform the transformation 1051, the LE RGB vectors are transformed 1051 by a color correction matrix. For example, Equations 1-3 show how to use the color correction matrices to correct the color values for the HE, ME, and LE sensors, respectively. Equation 1 shows how to use the color correction matrix to correct the color values of the HE sensor, where the 3×3 matrix coefficients, including values A1-A9, represent coefficients selected to strengthen or weaken the pixel value, and an RGB matrix (RLE , GLE , and BLE) represents the demosaiced RGB output signal from the LE sensor. In some cases, the 3×3 matrix coefficients can be derived by comparing the demosaiced output against expected (or so-called “truth”) values. For example, the 3×3 matrix coefficients can be derived by least-squares polynomial modeling between the demosaiced RGB output values and reference values from a reference color chart (e.g., a Macbeth chart). Similarly, Equation 2 shows how to use the color correction matrix to correct the color values of the ME sensor, where the RGB matrix (RME , GME , and BME) represents the demosaiced RGB output signal from the ME sensor, and Equation 3 shows how to use the color correction matrix to correct the color values of the SE sensor, where the RGB matrix (RME, GME, and BME) represents the demosaiced RGB output values from the SE sensor.

Equation 1 – correcting SE pixel values using [A], the Color Correction Matrix for the LE sensor

$$\begin{bmatrix} A_1 & A_2 & A_3 \\ A_4 & A_5 & A_6 \\ A_7 & A_8 & A_9 \end{bmatrix} \begin{bmatrix} R_{LE} \\ G_{LE} \\ B_{LE} \end{bmatrix} = [A] \begin{bmatrix} R_{LE} \\ G_{LE} \\ B_{LE} \end{bmatrix} = \begin{bmatrix} R_{truth} \\ G_{truth} \\ B_{truth} \end{bmatrix}$$

Equation 2 – correcting ME pixel values using [B], the Color Correction Matrix for the ME sensor

$$\begin{bmatrix} B_1 & B_2 & B_3 \\ B_4 & B_5 & B_6 \\ B_7 & B_8 & B_9 \end{bmatrix} \begin{bmatrix} R_{ME} \\ G_{ME} \\ B_{ME} \end{bmatrix} = [B] \begin{bmatrix} R_{ME} \\ G_{ME} \\ B_{ME} \end{bmatrix} = \begin{bmatrix} R_{truth} \\ G_{truth} \\ B_{truth} \end{bmatrix}$$

Equation 3 – correcting SE pixel values using [C], the Color Correction Matrix for the SE sensor

$$\begin{bmatrix} C_1 & C_2 & C_3 \\ C_4 & C_5 & C_6 \\ C_7 & C_8 & C_9 \end{bmatrix} \begin{bmatrix} R_{SE} \\ G_{SE} \\ B_{SE} \end{bmatrix} = [C] \begin{bmatrix} R_{SE} \\ G_{SE} \\ B_{SE} \end{bmatrix} = \begin{bmatrix} R_{truth} \\ G_{truth} \\ B_{truth} \end{bmatrix}$$

To convert an image from a first color correction space (CCS) to a second color correction space, the color correction matrices from one or more sensors can be used. This process may be referred to as converting between color correction spaces or calibrating color correction spaces. Neither the first color correction space nor the second color correction space accurately reflects the true color of the captured image. The first and the second color correction space both have inaccuracies, and those inaccuracies are, in general, different from one another. Thus RGB values from each sensor must be multiplied by a unique color correction matrix for those RGB values to appear as true colors.

The present invention includes a method for converting an image from the LE sensor's color correction space to the ME sensor's color correction space and is illustrated in Equation 4 below:

Equation 4 – converting LE pixel values from LE color correction space to ME color correction space

$$\begin{bmatrix} R_{SE} \\ G_{SE} \\ B_{SE} \end{bmatrix} [C][B]^{-1} = \begin{bmatrix} R_{ME} \\ G_{ME} \\ B_{ME} \end{bmatrix}$$

In Equation 4, the LE sensor's pixel values (R, G, B) are multiplied by the LE sensor's correction matrix, [C], and then multiplied by the inverse of the ME sensor's correction matrix, [B]. The result is a set of pixel values (R, G, B) that are in the ME sensor's color correction space.

Methods of the invention allow matching of the color correction space of the second sensor to the color correction space of the first sensor so that the images from the two sensors may be accurately combined, or merged. The method for applying all the inaccuracies of the second color correction space to the first color correction space, prior to combining images from

the two into an HDR image, is previously unknown. Typical methods for combining data from multiple CFA sensors rely on color-correcting each sensor's data to the "truth" values measured from a calibrated color card, prior to combining the images. This is problematic in an HDR system, where it is known that the brighter sensor's image will have significant portions that are saturated, which saturated portions should actually have been utilized from the darker sensor's image when combining. Color correcting an image that has color information based on saturated pixels will cause colors to be misidentified. Therefore, in an HDR system, color-correcting the brighter image (for example, to "truth" color values), prior to combining images, will lead to colors being misidentified because of the use of saturated pixel data in creating colors from a mosaic-patterned image. For this reason, we specify that (1) the darker image have its color information transformed to match the color space of the brighter image, (2) this transformed darker image be combined with the brighter image, and then (3) the final combined image be color-transformed to "truth" color values.

The solution provided in the present invention avoids this saturated-pixel color misidentification problem by performing the steps of [(a) demosaic 1045, (b) color correct 1051 & (c) mosaic 1057] data from the darker sensor, thereby ensuring all data is accurately returned to its non-demosaiced state prior to the step of merging the darker sensor's data with the brighter sensor's data.

Furthermore, prior to merging the images from two sensors, the present invention matches the color correction spaces of the two sensors. This transformation ensures that the two images (from the first and second color correction space sensors) can be accurately merged, pixel-for-pixel, in non-demosaiced format. It may at first seem counterintuitive to change the color correction space of one sensor to match the color correction space of a second sensor, especially when the second sensor's color correction space is known to differ from the "true" color correction space. However, it is an important feature in ensuring that (1) the brighter sensor's color information not be demosaiced prior to merging, and (2) the color data from both sensors is matched together, prior to merging the images. The color correction process 1001 uses matrices that may themselves be implemented as kernels in the pipeline 231 on the processing device 219. Thus the color correction process 1001 is compatible with an HDR pipeline workflow because the kernels are applied as they receive the pixel values.

After the LE information is transformed 1051 from the LE color correction space to the ME color correction space, the transformed values are mosaiced 1057 (i.e., the demosaicing process is reversed). The transformed scalar pixel values are now comparable with the Bayer-patterned scalar ME pixel values detected by the ME sensor, and the process 1001 includes merging 1061 of ME and HE non-demosaiced (i.e., scalar) sensor data.

The merged non-demosaiced image within the ME color correction space is then demosaiced 1067. This demosaicing 1064 is similar to the demosaicing 1045 described above, except the CFA pixel values undergoing the demosaicing process are now associated with the ME color correction space. The demosaicing 1067 produces RGB vectors in the ME color space. Those RGB vectors are transformed 1071 into the HE color space while also being color corrected ($[B][A]^{-1}[RGB]$). Equation 2 shows how to use the color correction matrix to correct the color values of the ME sensor. The color corrected ME information is transformed 1071 from the ME color correction space to the HE color correction space by multiplying the ME color correction matrix by the inverse of the SE color correction matrix.

After the ME information is transformed 1071 from the ME color correction space to the HE color correction space, the transformed vectors are mosaiced 1075 (i.e., the demosaicing process is reversed). This allows the transformed ME CFA Bayer-patterned pixel values to merge 1079 with the HE pixel values detected by the HE sensor. At this point in the color correction process 1001, the transformed color information detected by the HE and ME sensors is now calibrated to match the color information detected by the HE sensor. This newly merged color value data set now represents color values within the HE color correction space 205.

FIG. 11 illustrates a method 1301 for combined broadcasting of high dynamic range (HDR) video with standard dynamic range (SDR) video. The method 1301 provides for streaming HDR and SDR video. The method 1301 includes detecting 125—using an array of sensors 165—information representing a series of images, processing 1309 the information, and transmitting 1321 the information for HDR and LDR display with less than one frame of delay between detection and transmission.

After the color processing and tone-mapping, the pipeline has produced an HDR video signal. At this stage, the pipe can include a module for subtraction 1315 that, in real-time, subtracts the SDR signal from the HDR signal ($HDR-SDR = \text{residual}$). What flows from the subtraction module is a pair of streams—the SDR video signal and the residual signal.

Preferably, all of the color information is in the SDR signal. At this stage the HDR signal may be subject to HDR compression by a suitable operation (e.g., JPEG or similar). The pair of streams includes the 8-bit SDR signal and the compressed HDR residual signal, which provides for HDR display. This dual signal is broadcast over a communication network and may in-fact be broadcast over television networks, cellular networks, or the Internet. A device that receives the signal displays the video according to the capacity of that device. An SDR display device will “see” the 8-bit SDR signal and display a video at a dynamic range that is standard, which signal has also had a certain TMO applied to it. An HDR display device will decompress the residuals and combine the dual streams into an HDR signal and display HDR video.

Thus, the method 1301 and the apparatus 201 may be used for real-time HDR video capture as well as for the simultaneous delivery of HDR and SDR output in a single transmission. The processing 1309 may include the workflow from the processing device 219 to video (broadcast) output. The method 1301 and the apparatus 201 provide for real-time processing and complementary HDR/SDR display using features described herein such as multiple sensors all obtaining an isomorphic image through a single lens and streaming the resulting pixel values through a pipeline to replace saturated pixels in a merged HDR video signal. The method 101 and the apparatus 201 each captures video information using an array of sensors, processes that video information in real-time, and transmits the video information in real-time in a HDR and LDR compatible format.

FIG. 12 shows image sensors and processor of an apparatus 1201 for HDR video processing according to certain embodiments. The apparatus 1201 includes a processing device 1219 such as a field-programmable gate array (FPGA) or an application-specific integrated circuit (ASIC). A plurality of image sensors 1265 are coupled to the processing device 219 and also to an optical assembly 1259 (e.g., a glass or crystal cube or slab having one or more beamsplitter 1301 therein). The apparatus 1201 is configured to stream pixel values from each of the plurality of image sensors 1265 in a frame independent-manner through a pipeline on the processing device 1219. The pipeline includes a kernel operation that identifies saturated pixel values and a merge module to merge the pixel values to produce an HDR image.

The kernel operation operates on pixel values as they stream from each of the plurality of image sensors 1265 by examining, for a given pixel on the HE sensor 1213, values from a neighborhood of pixels surrounding the given pixel, finding saturated values in the neighborhood

of pixels, and using information from a corresponding neighborhood on the ME sensor 1211 to estimate a value for the given pixel.

Various components of the apparatus 1201 may be connected via a printed circuit board 1205. The apparatus 1201 may also include memory and optionally a processor (such as a general-purpose processor like an ARM microcontroller). Apparatus 1201 may further include or be connected to one or more of an input-output device or a display. Memory can include RAM or ROM and preferably includes at least one tangible, non-transitory medium. A processor may be any suitable processor known in the art, such as the processor sold under the trademark XEON E7 by Intel (Santa Clara, CA) or the processor sold under the trademark OPTERON 6200 by AMD (Sunnyvale, CA). Input/output devices according to the invention may include a video display unit (e.g., a liquid crystal display or LED display), keys, buttons, a signal generation device (e.g., a speaker, chime, or light), a touchscreen, an accelerometer, a microphone, a cellular radio frequency antenna, port for a memory card, and a network interface device, which can be, for example, a network interface card (NIC), Wi-Fi card, or cellular modem. The apparatus 1201 may include or be connected to a storage device. The plurality of sensors are preferably provided in an arrangement that allows multiple sensors 1265 to simultaneously receive images that are identical except for light level.

The apparatus 1201 includes the multiple sensors 265 in the depicted arrangement. The multiple sensors preferably include at least a high exposure (HE) sensor 1213 and a middle exposure (ME) sensor 1211. Each image sensor may have its own color filter array 1307. The color filter arrays 1307 may operate as a Bayer filter such that each pixel receives either red, green, or blue light. As is known in the art, a Bayer filter includes a repeating grid of red, green, blue, green filters such that a sequence of pixel values streaming from the sensor corresponds to values for red, green, blue, green, red, green, blue, green, red, green, blue, green,...etc.

The apparatus 1201 may also include or be optically connected to a lens and at least one beamsplitter 1301. The HE sensor 1213, the ME sensor 1211, the lens 1311 and the at least one beamsplitter 1301 are arranged to receive an incoming beam of light 1305 and split the beam of light 305 into at least a first path that impinges and HE sensor 1213 and a second path that impinges on the ME sensor 1211. In a preferred embodiment, the apparatus 1201 uses a set of partially-reflecting surfaces to split the light from a single photographic lens 1311 so that it is focused onto three imaging sensors simultaneously. In a preferred embodiment, the light is

directed back through one of the beamsplitters a second time, and the three sub-images are not split into red, green, and blue but instead are optically identical except for their light levels.

In preferred embodiments, the first beamsplitter 1301 and the second beamsplitter 1319 are provided as part of an optical assembly 1259. The optical assembly 1259 is preferably a glass or crystal piece, such as a slab, which may be made up of a few wedge-shaped and smaller slab-shaped pieces cemented together. Glass is a suitable material option. In one embodiment, the first beamsplitter 1301 is at a 45° angle and has an approximate T/R ratio of 92/8, which means that 92% of the light from the camera lens is transmitted through the first beamsplitter 1301 and focused directly onto the high-exposure (HE) sensor 213. The beamsplitter 1301 reflects 8% of the light from the lens upwards, toward the second uncoated beamsplitter 1319, which has the same optical properties as the first but is positioned at a 90° angle to the light path and has an approximate T/R ratio of 94/6. Of the 8% of the total light that is reflected upwards, 94% (or 7.52% of the total light) is transmitted through the second beamsplitter 1319 and focused onto the medium-exposure (ME) sensor 211. The other 6% of this upward-reflected light (or 0.48% of the total light) is reflected back down by the second beamsplitter 1319 toward the first beamsplitter 1301 (which is again at 45°), through which 92% (or 0.44% of the total light) is transmitted and focused onto the low-exposure (LE) sensor 261. With this arrangement, the HE, ME and LE sensors capture images with 92%, 7.52%, and 0.44% of the total light gathered by the camera lens, respectively. Thus a total of 99.96% of the total light gathered by the camera lens has been captured by the image sensors. Therefore, the HE and ME exposures are separated by $12.2\times$ (3.61 stops) and the ME and LE are separated by $17.0\times$ (4.09 stops), which means that this configuration is designed to extend the dynamic range of the sensor by 7.7 stops.

A feature and benefit of the depicted embodiment is that the plurality of image sensors 1265 are coupled to the optical assembly 1259 (and also optionally to the processing device 1219). For example, the sensors, which may themselves be a semiconductor material, may be cemented, e.g., with an adhesive to the glass or crystal cube or slab of the assembly 1259. A Bayer filter may be disposed between a surface of the image sensor and the glass of the assembly 1259. This construction in some embodiments uses sturdy solid glass beamsplitter cubes, where the camera sensors are bonded to the glass. A benefit of such a construction is that optical components are unable to come out of alignment. This allows the camera to be used in high-stress environments, such as on fast-moving vehicles, aircraft, and amusement park rides. It also

allows the optical assembly 1259 (with or without the image sensors 1265) to be manufactured as one component and, e.g., shipped for inclusion in a larger device, such as a smart phone. E.g., the optical assembly 1259 along with sensors 1259 and processing device 1219 can be produced for sale to a smartphone manufacturer and when passed from the maker of the optical assembly to the device manufacturer, the optical components will not come out of alignment.

Incorporation by Reference

References and citations to other documents, such as patents, patent applications, patent publications, journals, books, papers, web contents, have been made throughout this disclosure. All such documents are hereby incorporated herein by reference in their entirety for all purposes.

Equivalents

Various modifications of the invention and many further embodiments thereof, in addition to those shown and described herein, will become apparent to those skilled in the art from the full contents of this document, including references to the scientific and patent literature cited herein. The subject matter herein contains important information, exemplification and guidance that can be adapted to the practice of this invention in its various embodiments and equivalents thereof.

CLAIMS:

1. A method for producing real-time HDR video, the method comprising:
receiving incoming light through a lens and splitting the light via at least one beamsplitter onto a first image sensor and a second image sensor, wherein the first sensor receives a higher exposure than the second sensor, and wherein at least 95% of the light gathered by the imaging lens is captured by the multiple image sensors;
streaming pixel values from each of multiple sensors in a frame independent-manner through a pipeline on a processing device, wherein the pipeline includes a kernel operation that identifies saturated pixel values; and
merging the pixel values to produce an HDR image, wherein the merging includes using first sensor pixel values that are not saturated and second sensor pixel values that correspond to the identified saturated pixel values.
2. The method of claim 1, wherein the merging step excludes at least some of the saturated pixel values from the HDR image.
3. The method of claim 2, wherein the multiple image sensors all capture images simultaneously through the lens.
4. The method of claim 3, wherein streaming the pixel values through the kernel operation includes examining values from a neighborhood of pixels surrounding a first pixel on the first sensor, finding saturated values in the neighborhood of pixels, and using information from a corresponding neighborhood on the second sensor to estimate a value for the first pixel.
5. The method of claim 4, wherein at least some of the saturated pixel values are identified before receiving values from all pixels of the multiple image sensors at the processing device.

6. The method of claim 5, further comprising beginning to merge portions of the pixel values while still streaming later-arriving pixel values through the kernel operation.
7. The method of claim 5, wherein sequences of pixel values are streamed through the processing device and merged without waiting to receive pixel values from all pixels on the image sensors.
8. The method of claim 7, wherein the processing device comprises one selected from the group consisting of a field-programmable gate array and an application-specific integrated circuit, and further wherein each of the image sensors comprises a color filter array, and the method further comprises demosaicing the HDR imaging after the merging.
9. The method of claim 8, wherein multiple image sensors capture images that are optically identical except for light level.
10. The method of claim 8, wherein streaming the pixel values through the kernel operation includes streaming the pixel values through a path within the processing device that momentarily places a value from the first pixel proximal to each value originating from the neighborhood of pixels.
11. The method of claim 8, wherein the obtaining, streaming, and merging steps are performed by streaming the sequences of pixel values through a pipeline on the processing device such that no location on the processing device stores a complete image.
12. The method of claim 1, wherein a first pixel value from a first pixel on one of the image sensors is identified as saturated if it is at least 90% of a maximum possible pixel value.
13. An apparatus for HDR video processing, the apparatus comprising:
a processing device;

a first image sensor and a second image sensor coupled to the processing device; and

a lens and at least one beamsplitter arranged to receive an incoming beam of light and split the beam of light into at least a first path that impinges on the first image sensor and a second path that impinges on the second image sensor, wherein the first sensor receives a higher exposure than the second sensor, wherein the apparatus is configured to stream pixel values from each of the plurality of image sensors in a frame-independent manner through a pipeline on the processing device, wherein the pipeline includes

a kernel operation that identifies saturated pixel values, wherein the kernel operation operates on pixel values as they stream from the first and second image sensors by examining, for a given pixel on the first sensor, values from a neighborhood of pixels surrounding the given pixel, finding saturated values in the neighborhood of pixels, and using information from a corresponding neighborhood on the second sensor to estimate a value for the given pixel and

a merge module to merge the pixel values to produce an HDR image.

14. The apparatus of claim 13, wherein the beamsplitter directs a majority of the light to the first path and a lesser amount of the light to the second path.

15. The apparatus of claim 14, wherein the first path and the second path impinge on the first and the second image sensor, respectively, to generate images that are optically identical but for light level.

16. The apparatus of claim 15, wherein the processing device comprises a field-programmable gate array or an application-specific integrated circuit that includes the pipeline.

17. The apparatus of claim 13, wherein the pipeline includes-in the order in which the pixel values flow:

a sync module to synchronize the pixel values as the pixel values stream onto the processing device from the plurality of image sensors;

the kernel operation;

the merge module;

a demosaicing module; and

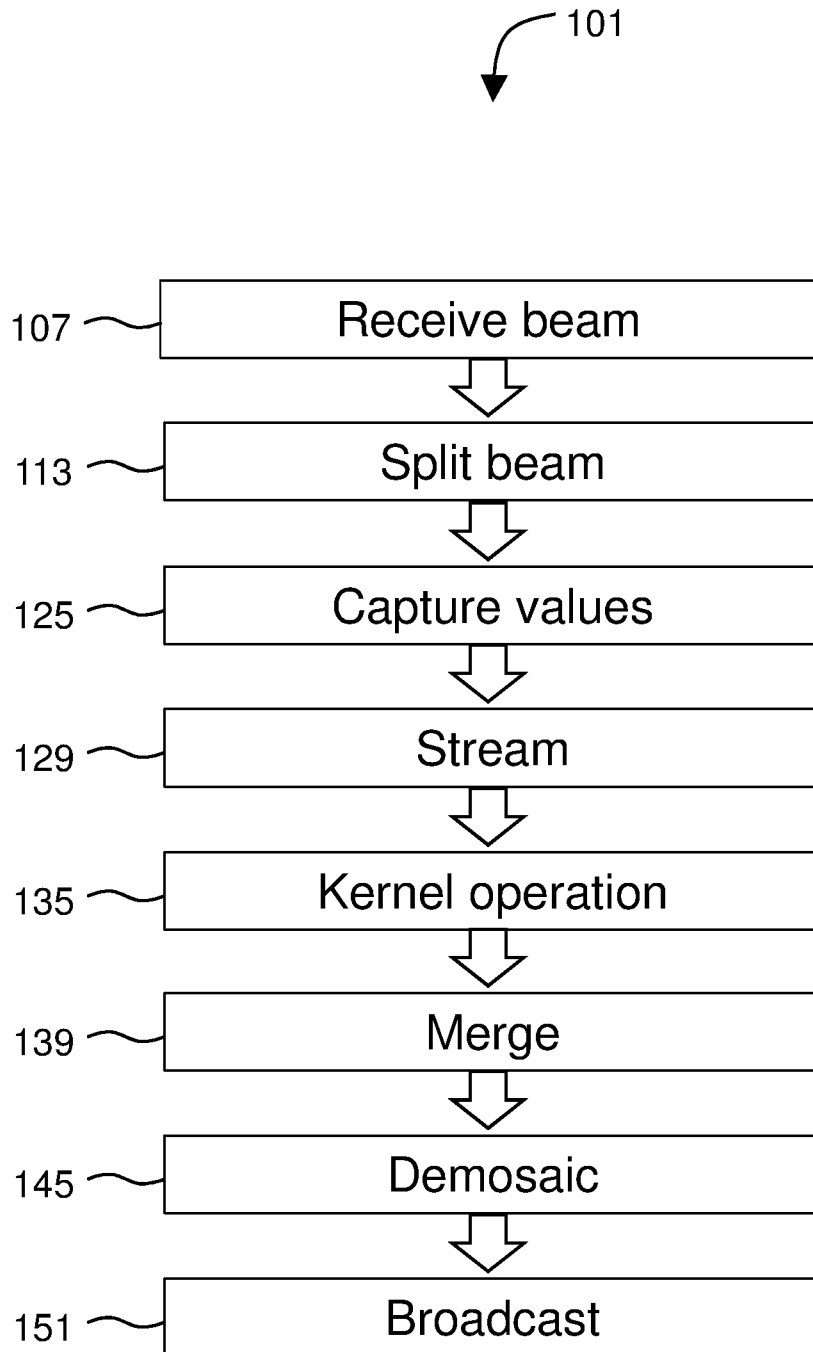
a tone-mapping operator.

18. The apparatus of claim 17, wherein the pipeline further comprises one or more of a color- correction module; an HDR conversion module; and an HDR compression module.

19. The apparatus of claim 16, further comprising:

a sync module on the processing device, in the pipeline upstream of the kernel operation, to synchronize the pixel values as the pixel values stream onto the processing device from the plurality of image sensors.

1/12

**FIG. 1**

2/12

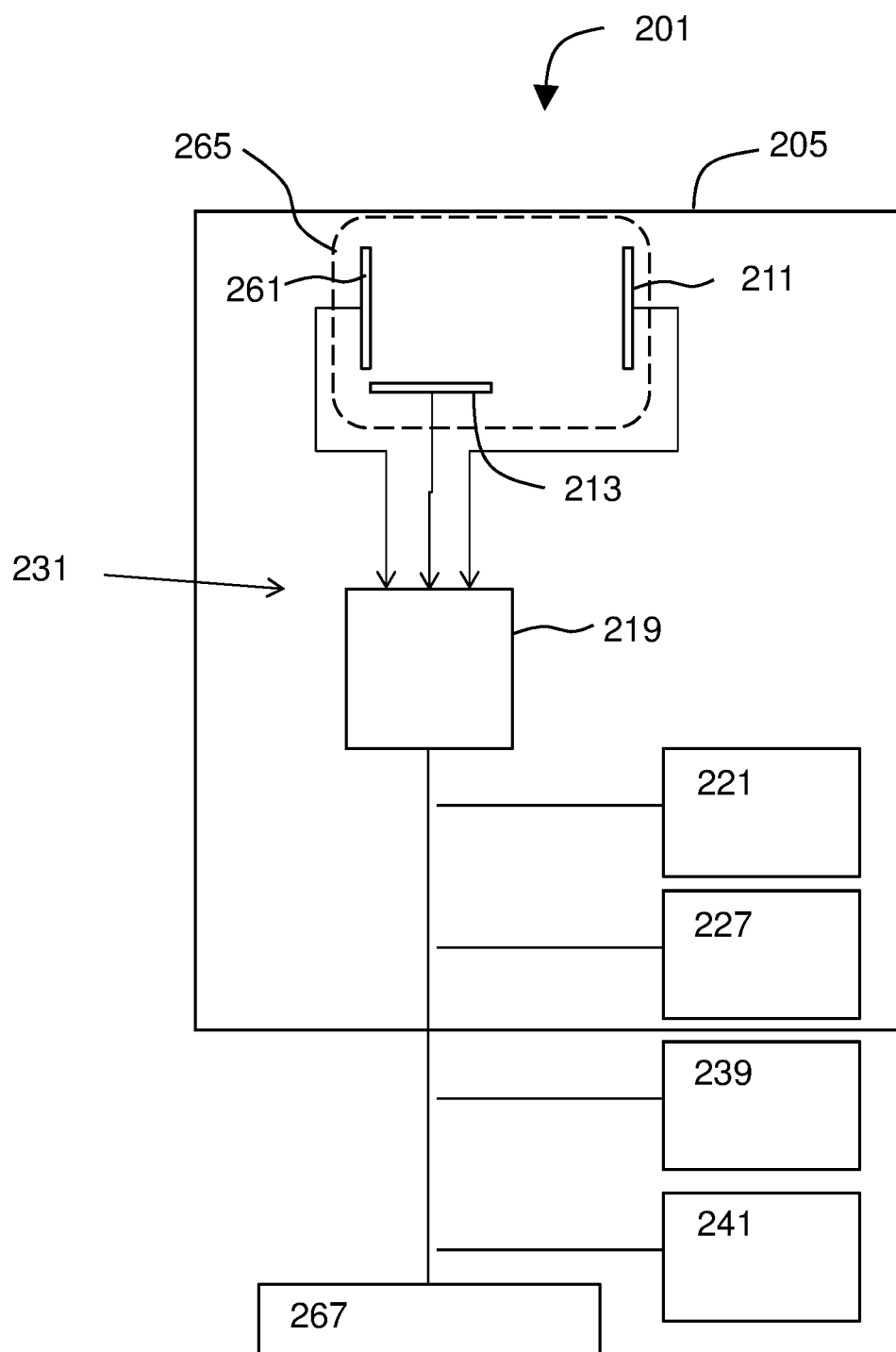


FIG. 2

3/12

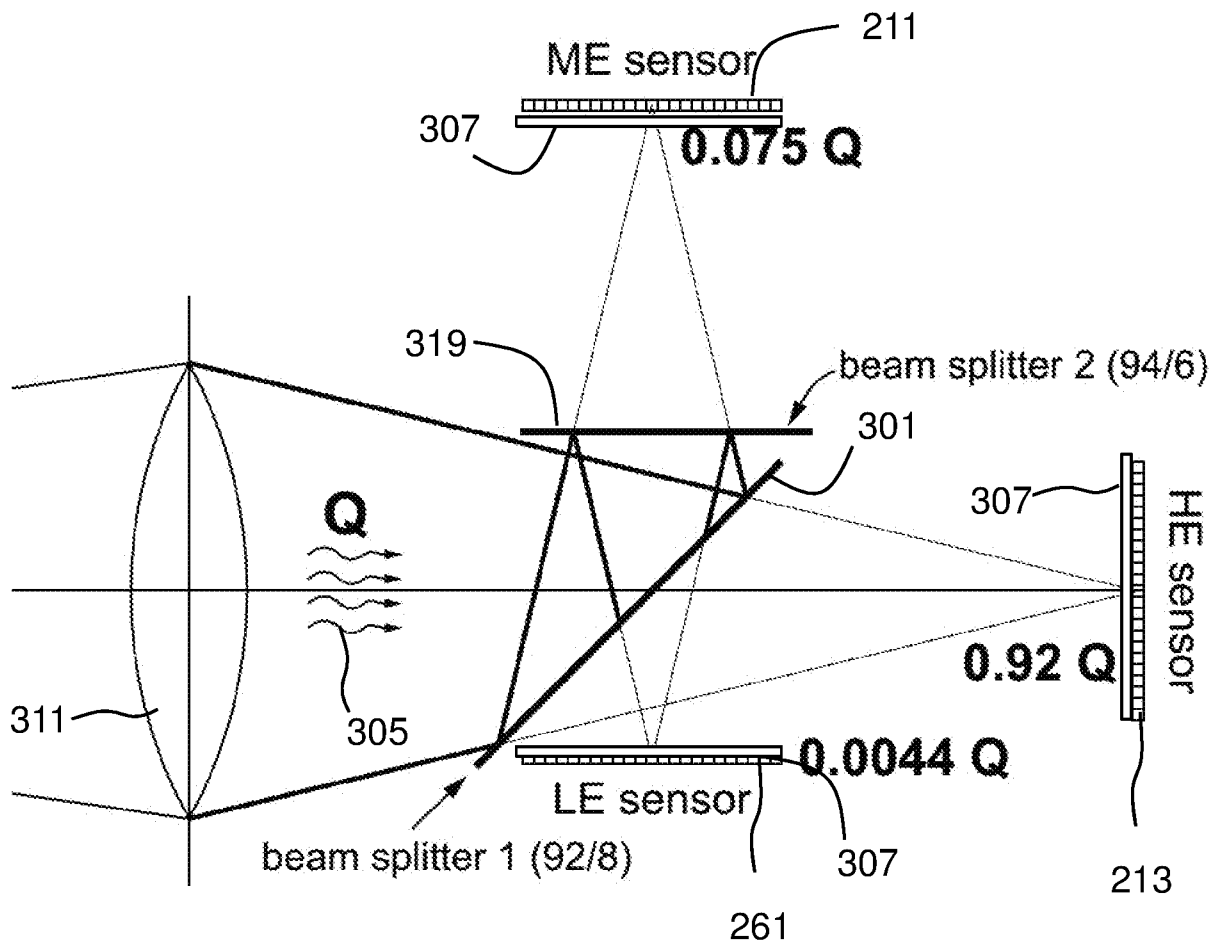


FIG. 3

4/12

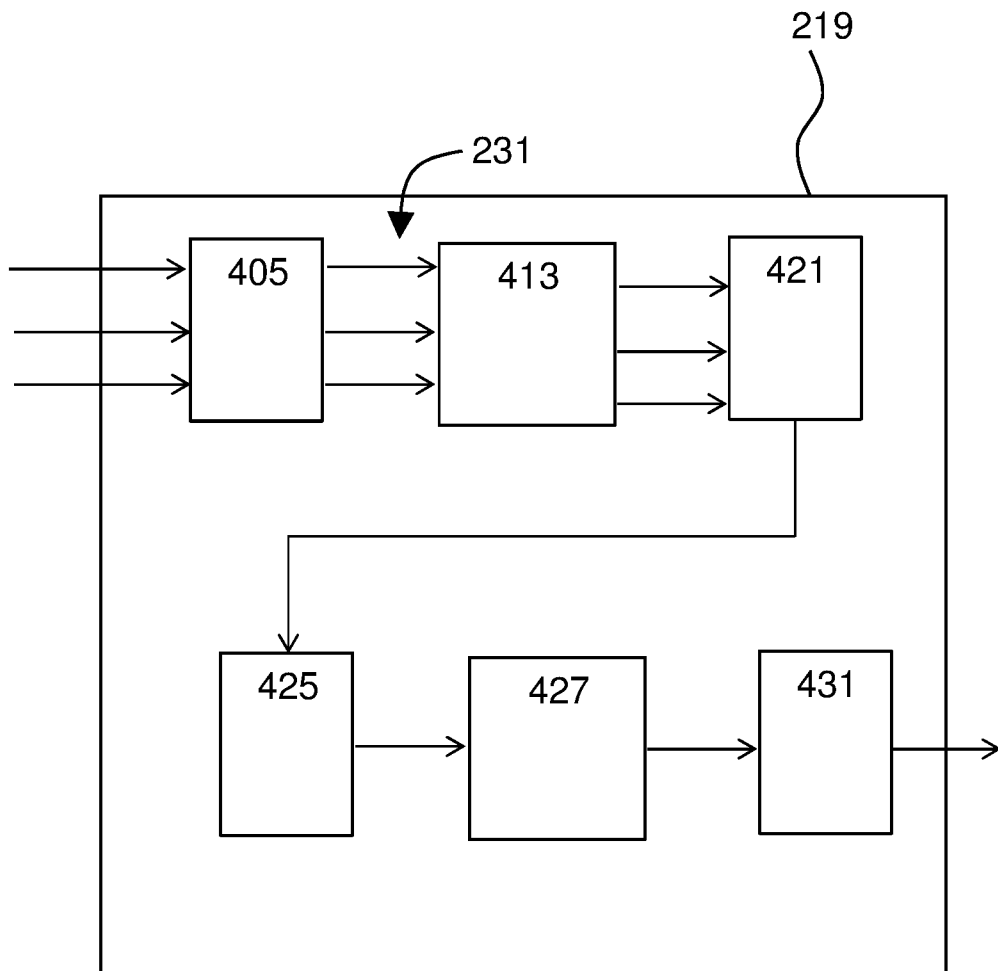


FIG. 4

5/12

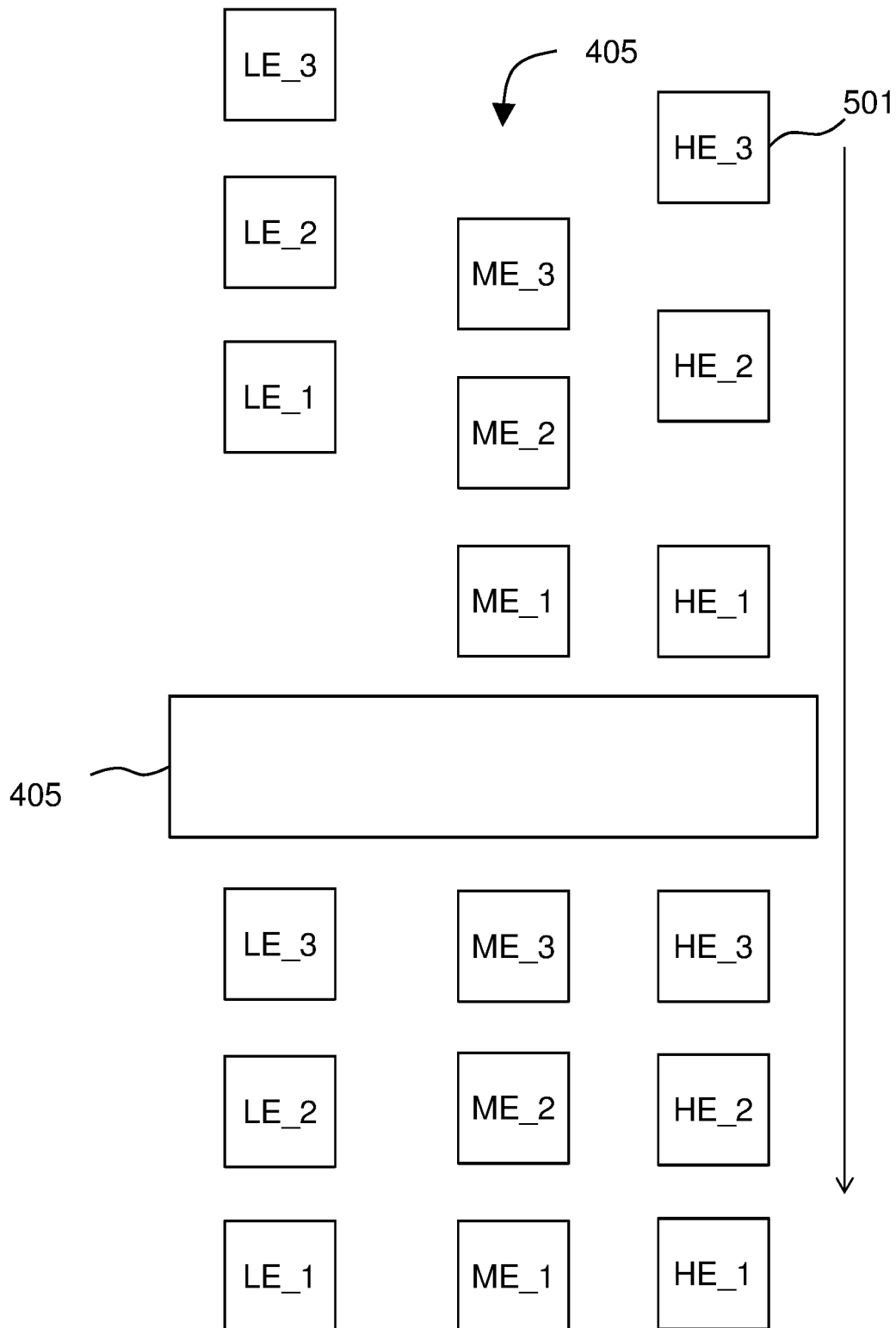


FIG. 5

6/12

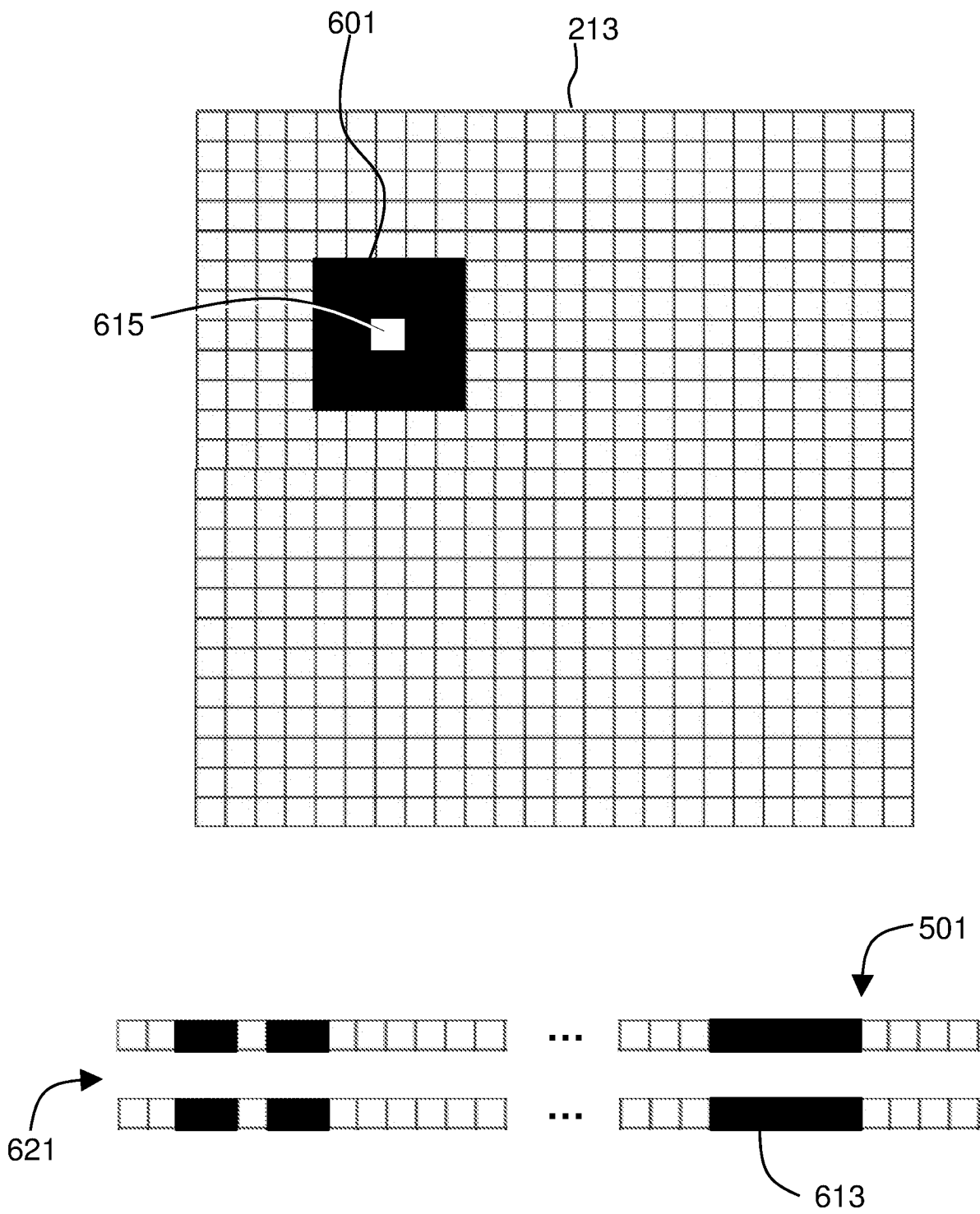


FIG. 6

7/12

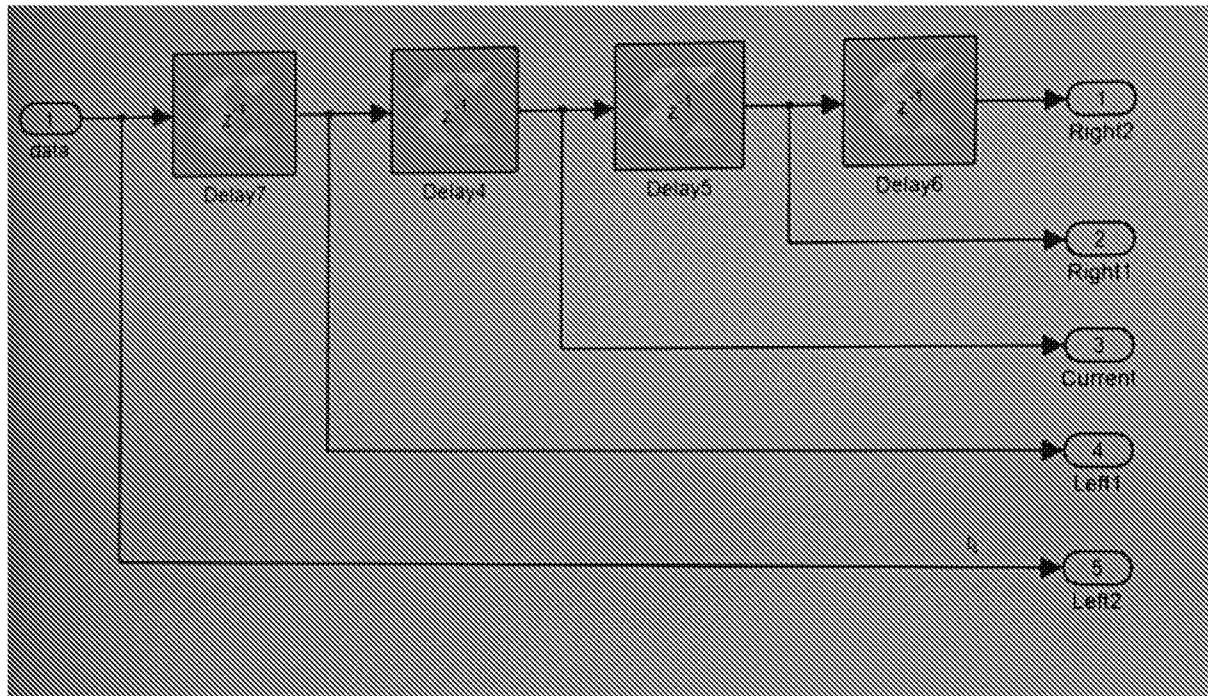


FIG. 7

8/12

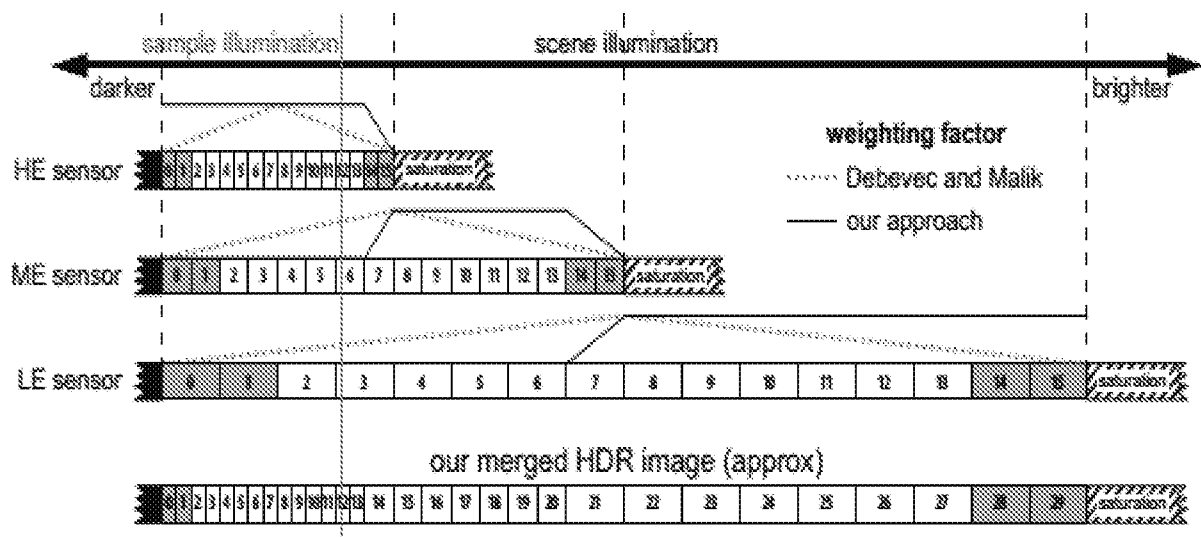


FIG. 8

9/12

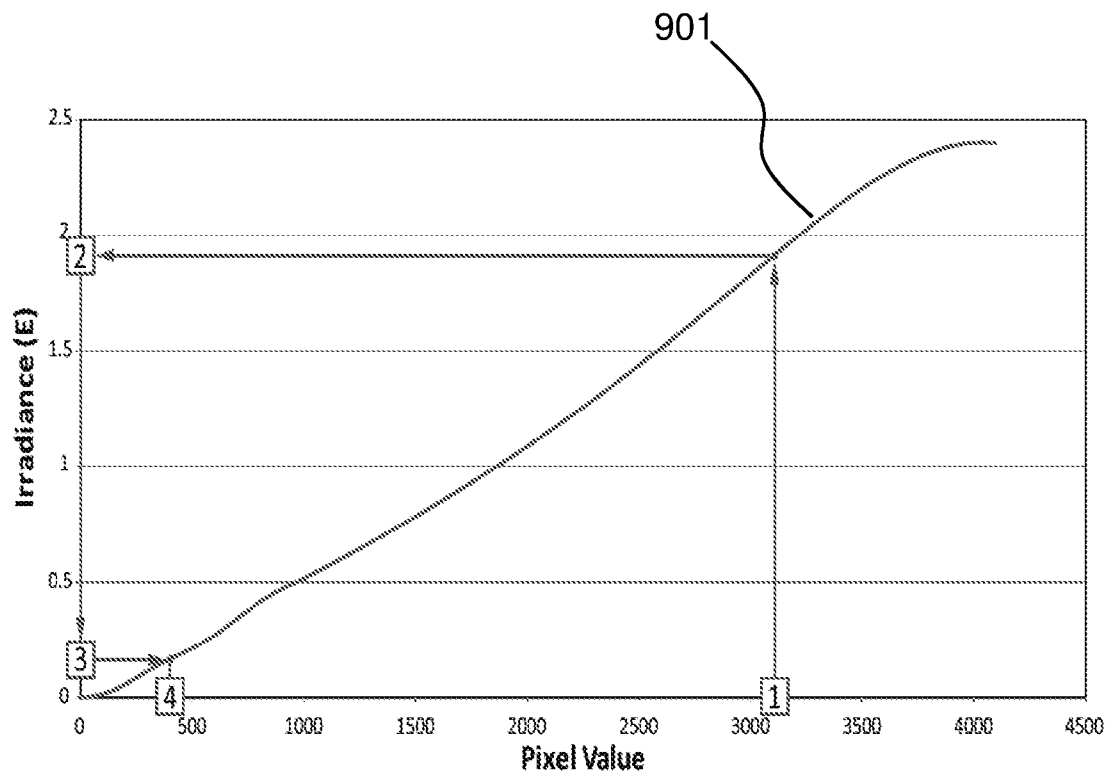


FIG. 9

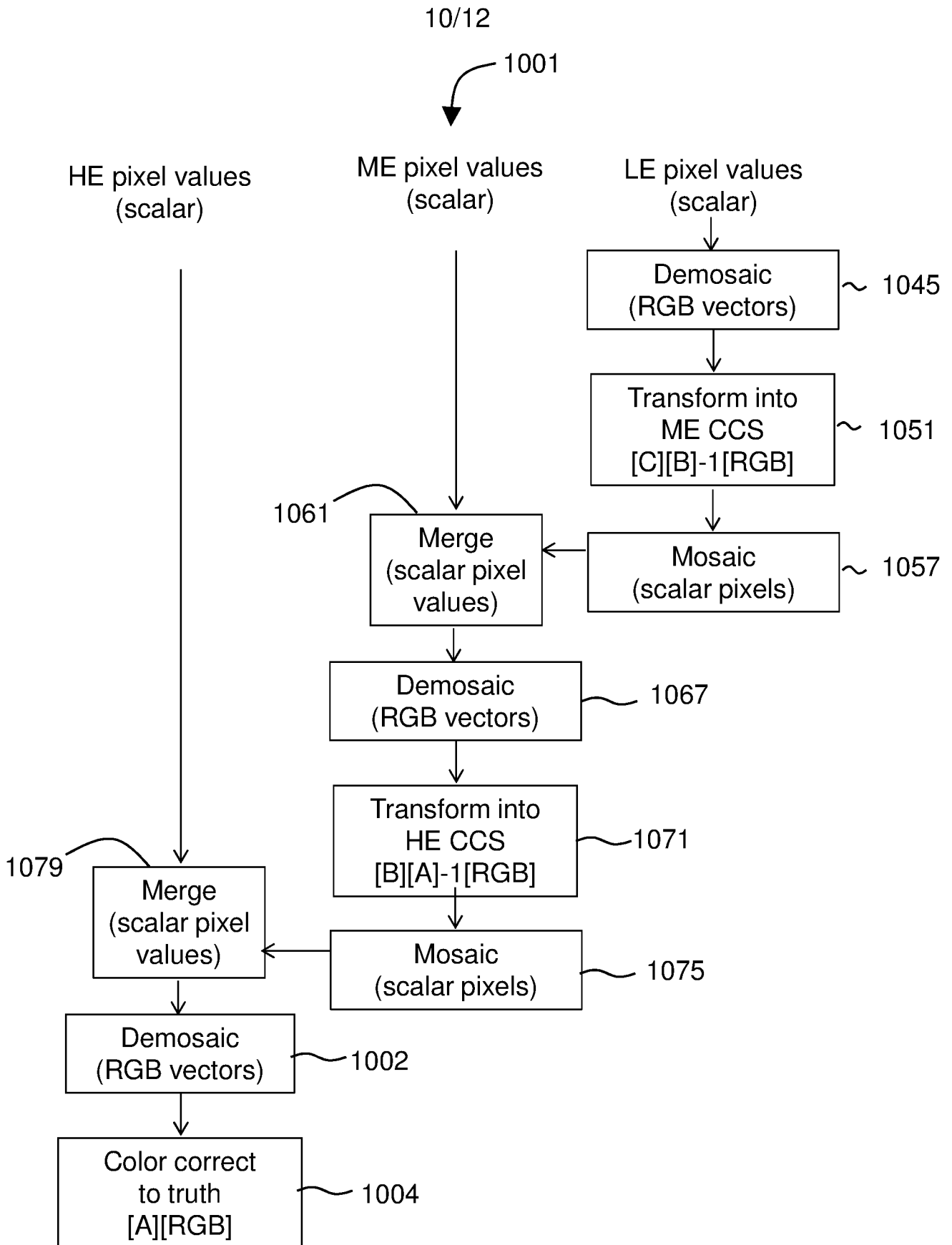


FIG. 10

11/12

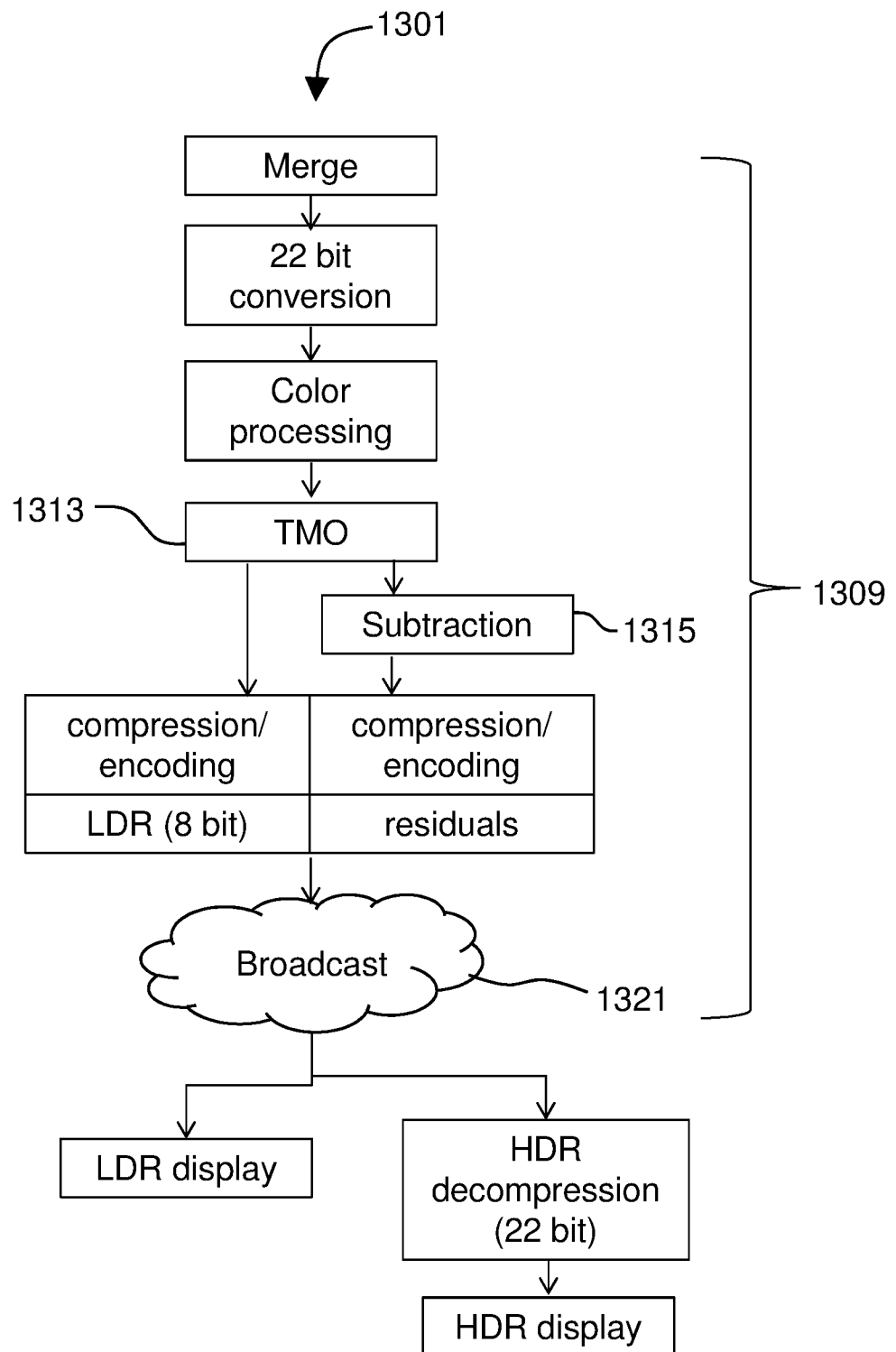


FIG. 11

12/12

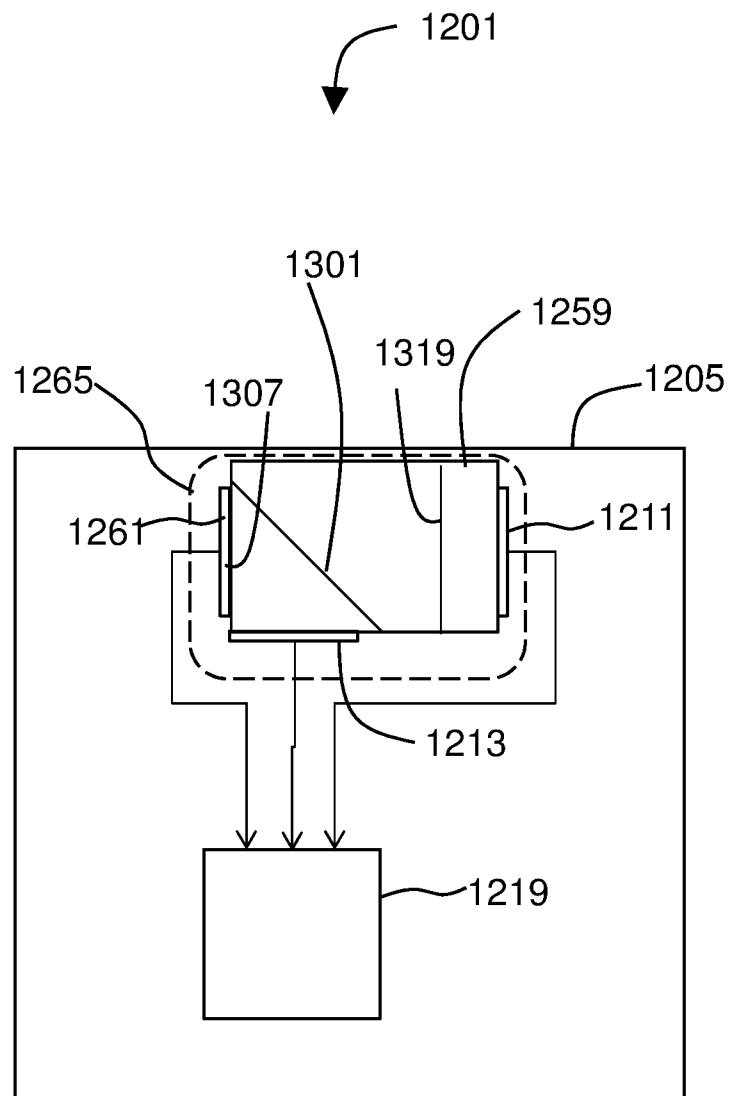


FIG. 12

101

