

CAPTURING AND PROCESSING OF IMAGES USING MONOLITHIC CAMERA ARRAY WITH HETEROGENEOUS IMAGERS

Inventors:

Kartik Venkataraman

Amandeep S. Jabbi

Robert H. Mullis, Jr.

RELATED APPLICATION

[0001] This application claims priority to U.S. patent application serial number 61/054,694 entitled "Monolithic Integrated Array of Heterogeneous Image Sensors," filed on May 20, 2008, which is incorporated by reference herein in its entirety.

FIELD OF THE INVENTION

[0002] The present invention is related to an image sensor including a plurality of heterogeneous imagers, more specifically to an image sensor with a plurality of wafer-level imagers having custom filters, sensors and optics of varying configurations.

BACKGROUND

[0003] Image sensors are used in cameras and other imaging devices to capture images. In a typical imaging device, light enters through an opening (aperture) at one end of the imaging device and is directed to an image sensor by an optical element such as a lens. In most imaging devices, one or more layers of optical elements are placed between the aperture and the image sensor to focus light onto the image sensor. The image sensor consists of pixels that generate signals upon receiving light via the optical element. Commonly used image sensors include CCD (charge-coupled device) image sensors and CMOS (complementary metal-oxide-semiconductor) sensors.

[0004] Filters are often employed in the image sensor to selectively transmit lights of certain wavelengths onto pixels. A Bayer filter mosaic is often formed on the image sensor. The Bayer filter is a color filter array that arranges one of the RGB color filters on each of the color pixels. The Bayer filter pattern includes 50% green filters, 25% red filters and 25% blue filters. Since each pixel generates a signal representing strength of a color component in the light and not the full range of colors, demosaicing is performed to interpolate a set of red, green and blue values for each image pixel.

[0005] The image sensors are subject to various performance constraints. The performance constraints for the image sensors include, among others, dynamic range, signal

to noise (SNR) ratio and low light sensitivity. The dynamic range is defined as the ratio of the maximum possible signal that can be captured by a pixel to the total noise signal.

Typically, the well capacity of an image sensor limits the maximum possible signal that can be captured by the image sensor. The maximum possible signal in turn is dependent on the strength of the incident illumination and the duration of exposure (e.g., integration time, and shutter width). The dynamic range can be expressed as a dimensionless quantity in decibels (dB) as:

$$DR = \frac{\text{full well capacity}}{\text{RMS noise}} \dots\dots\dots \text{equation (1)}$$

Typically, the noise level in the captured image influences the floor of the dynamic range. Thus, for an 8 bit image, the best case would be 48dB assuming the RMS noise level is 1 bit. In reality, however, the RMS noise levels are higher than 1 bit, and this further reduces the dynamic range.

[0006] The signal to noise ratio (SNR) of a captured image is, to a great extent, a measure of image quality. In general, as more light is captured by the pixel, the higher the SNR. The SNR of a captured image is usually related to the light gathering capability of the pixel.

[0007] Generally, Bayer filter sensors have low light sensitivity. At low light levels, each pixel's light gathering capability is constrained by the low signal levels incident upon each pixel. In addition, the color filters over the pixel further constrain the signal reaching the pixel. IR (Infrared) filters also reduce the photo-response from near-IR signals, which can carry valuable information.

[0008] These performance constraints of image sensors are greatly magnified in cameras designed for mobile systems due to the nature of design constraints. Pixels for mobile cameras are typically much smaller than the pixels of digital still cameras (DSC). Due to limits in light gathering ability, reduced SNR, limits in the dynamic range, and reduced sensitivity to low light scenes, the cameras in mobile cameras show poor performance.

SUMMARY

[0009] Embodiments provide a camera array, an imaging device including a camera array and/or a method for capturing image that employ a plurality of imagers fabricated on a substrate where each imager includes a plurality of sensor elements. The plurality of imagers include at least a first imager formed on a first location of the substrate and a second imager

formed on a second location of the substrate. The first imager and the second imager may have the same imaging characteristics or different imaging characteristics.

[0010] In one embodiment, the first imaging characteristics and the second imager have different imaging characteristics. The imaging characteristics may include, among others, the size of the imager, the type of pixels included in the imager, the shape of the imager, filters associated with the imager, exposure time of the imager, aperture size associated with the imager, the configuration of the optical element associated with the imager, gain of the imager, the resolution of the imager, and operational timing of the imager.

[0011] In one embodiment, the first imager includes a filter for transmitting a light spectrum. The second imager also includes the same type of filter for transmitting the same light spectrum as the first imager but captures an image that is sub-pixel phase shifted from an image captured by the first imager. The images from the first imager and the second imager are combined using a super-resolution process to obtain images of higher resolution.

[0012] In one embodiment, the first imager includes a first filter for transmitting a first light spectrum and the second imager includes a second filter for transmitting a second light spectrum. The images from the first and second imagers are then processed to obtain a higher quality image.

[0013] In one embodiment, lens elements are provided to direct and focus light onto the imagers. Each lens element focuses light onto one imager. Because each lens element is associated with one imager, each lens element may be designed and configured for a narrow light spectrum. Further, the thickness of the lens element may be reduced, decreasing the overall thickness of the camera array. The lens elements are fabricated using wafer level optics (WLO) technology.

[0014] In one embodiment, the plurality of imagers include at least one near-IR imager dedicated to receiving near-IR (Infrared) spectrum. An image generated from the near-IR imager may be fused with images generated from other imagers with color filters to reduce noise and increase the quality of the images.

[0015] In one embodiment, the plurality of imagers may be associated with lens elements that provide a zooming capability. Different imagers may be associated with lens of different focal lengths to have different fields-of-views and provide different levels of zooming capability. A mechanism may be provided to provide smooth transition from one zoom level to another zoom level.

[0016] In one or more embodiments, the plurality of imagers is coordinated and operated to obtain at least one of a high dynamic range image, a panoramic image, a hyper-spectral image, distance to an object and a high frame rate video.

[0017] The features and advantages described in the specification are not all inclusive and, in particular, many additional features and advantages will be apparent to one of ordinary skill in the art in view of the drawings, specification, and claims. Moreover, it should be noted that the language used in the specification has been principally selected for readability and instructional purposes, and may not have been selected to delineate or circumscribe the inventive subject matter.

BRIEF DESCRIPTION OF DRAWINGS

[0018] Figure 1 is a plan view of a camera array with a plurality of imagers, according to one embodiment.

[0019] Figure 2A is a perspective view of a camera array with lens elements, according to one embodiment.

[0020] Figure 2B is a cross-sectional view of a camera array, according to one embodiment.

[0021] Figures 3A and 3B are sectional diagrams illustrating changes in the heights of lens elements depending on changes in the dimensions of imagers, according to one embodiment.

[0022] Figure 3C is a diagram illustrating chief ray angles varying depending on differing dimensions of the lens elements.

[0023] Figure 4 is a functional block diagram for an imaging device, according to one embodiment.

[0024] Figure 5 is a functional block diagram of an image processing pipeline module, according to one embodiment.

[0025] Figure 6A through 6E are plan views of camera arrays having different layouts of heterogeneous imagers, according to embodiments.

[0026] Figure 7 is a flowchart illustrating a process of generating an enhanced image from lower resolution images captured by a plurality of imagers, according to one embodiment.

DETAILED DESCRIPTION

[0027] A preferred embodiment of the present invention is now described with reference to the figures where like reference numbers indicate identical or functionally

similar elements. Also in the figures, the left most digits of each reference number corresponds to the figure in which the reference number is first used.

[0028] Embodiments relate to using a distributed approach to capturing images using a plurality of imagers of different imaging characteristics. Each imager may be spatially shifted from another imager in such a manner that an imager captures an image that is shifted by a sub-pixel amount with respect to another imager captured by another imager. Each imager may also include separate optics with different filters and operate with different operating parameters (e.g., exposure time). Distinct images generated by the imagers are processed to obtain an enhanced image. Each imager may be associated with an optical element fabricated using wafer level optics (WLO) technology.

[0029] A sensor element or pixel refers to an individual light sensing element in a camera array. The sensor element or pixel includes, among others, traditional CIS (CMOS Image Sensor), CCD (charge-coupled device), high dynamic range pixel, multispectral pixel and various alternatives thereof.

[0030] An imager refers to a two dimensional array of pixels. The sensor elements of each imager have similar physical properties and receive light through the same optical component. Further, the sensor elements in the each imager may be associated with the same color filter.

[0031] A camera array refers to a collection of imagers designed to function as a unitary component. The camera array may be fabricated on a single chip for mounting or installing in various devices.

[0032] An array of camera array refers to an aggregation of two or more camera arrays. Two or more camera arrays may operate in conjunction to provide extended functionality over a single camera array.

[0033] Image characteristics of an imager refer to any characteristics or parameters of the imager associated with capturing of images. The imaging characteristics may include, among others, the size of the imager, the type of pixels included in the imager, the shape of the imager, filters associated with the imager, the exposure time of the imager, aperture size associated with the imager, the configuration of the optical element associated with the imager, gain of the imager, the resolution of the imager, and operational timing of the imager.

Structure of Camera array

[0034] Figure 1 is a plan view of a camera array 100 with imagers 1A through NM, according to one embodiment. The camera array 100 is fabricated on a semiconductor chip to include a plurality of imagers 1A through NM. Each of the imagers 1A through NM may

include a plurality of pixels (e.g., 0.32 Mega pixels). In one embodiment, the imagers 1A through NM are arranged into a grid format as illustrated in Figure 1. In other embodiments, the imagers are arranged in a non-grid format. For example, the imagers may be arranged in a circular pattern, zigzagged pattern or scattered pattern.

[0035] The camera array may include two or more types of heterogeneous imagers, each imager including two or more sensor elements or pixels. Each one of the imagers may have different imaging characteristics. Alternatively, there may be two or more different types of imagers where the same type of imagers shares the same imaging characteristics.

[0036] In one embodiment, each imager 1A through NM has its own filter and/or optical element (e.g., lens). Specifically, each of the imagers 1A through NM or a group of imagers may be associated with spectral color filters to receive certain wavelengths of light. Example filters include a traditional filter used in the Bayer pattern (R, G, B or their complements C, M, Y), an IR-cut filter, a near-IR filter, a polarizing filter, and a custom filter to suit the needs of hyper-spectral imaging. Some imagers may have no filter to allow reception of both the entire visible spectra and near-IR, which increases the imager's signal-to-noise ratio. The number of distinct filters may be as large as the number of imagers in the camera array. Further, each of the imagers 1A through NM or a group of imagers may receive light through lens having different optical characteristics (e.g., focal lengths) or apertures of different sizes.

[0037] In one embodiment, the camera array includes other related circuitry. The other circuitry may include, among others, circuitry to control imaging parameters and sensors to sense physical parameters. The control circuitry may control imaging parameters such as exposure times, gain, and black level offset. The sensor may include dark pixels to estimate dark current at the operating temperature. The dark current may be measured for on-the-fly compensation for any thermal creep that the substrate may suffer from.

[0038] In one embodiment, the circuit for controlling imaging parameters may trigger each imager independently or in a synchronized manner. The start of the exposure periods for the various imagers in the camera array (analogous to opening a shutter) may be staggered in an overlapping manner so that the scenes are sampled sequentially while having several imagers being exposed to light at the same time. In a conventional video camera sampling a scene at N exposures per second, the exposure time per sample is limited to 1/N seconds. With a plurality of imagers, there is no such limit to the exposure time per sample because multiple imagers may be operated to capture images in a staggered manner.

[0039] Each imager can be operated independently. Entire or most operations associated with each individual imager may be individualized. In one embodiment, a master setting is programmed and deviation (i.e., offset or gain) from such master setting is configured for each imager. The deviations may reflect functions such as high dynamic range, gain settings, integration time settings, digital processing settings or combinations thereof. These deviations can be specified at a low level (e.g., deviation in the gain) or at a higher level (e.g., difference in the ISO number, which is then automatically translated to deltas for gain, integration time, or otherwise as specified by context/master control registers) for the particular camera array. By setting the master values and deviations from the master values, higher levels of control abstraction can be achieved to facilitate simpler programming model for many operations. In one embodiment, the parameters for the imagers are arbitrarily fixed for a target application. In another embodiment, the parameters are configured to allow a high degree of flexibility and programmability.

[0040] In one embodiment, the camera array is designed as a drop-in replacement for existing camera image sensors used in cell phones and other mobile devices. For this purpose, the camera array may be designed to be physically compatible with conventional image sensors of approximately the same resolution although the achieved resolution of the camera array may exceed conventional image sensors in many photographic situations. Taking advantage of the increased performance, the camera array of the embodiment may include fewer pixels to obtain equal or better quality images compared to conventional image sensors. Alternatively, the size of the pixels in the imager may be reduced compared to pixels in conventional image sensors while achieving comparable results.

[0041] In order to match the raw pixel count of a conventional image sensor without increasing silicon area, the logic overhead for the individual imagers is preferably constrained in the silicon area. In one embodiment, much of the pixel control logic is a single collection of functions common to all or most of the imagers with a smaller set of functions applicable each imager. In this embodiment, the conventional external interface for the imager may be used because the data output does not increase significantly for the imagers.

[0042] In one embodiment, the camera array including the imagers replaces a conventional image sensor of M megapixels. The camera array includes $N \times N$ imagers, each sensor including pixels of $\frac{M}{N^2}$. Each imager in the camera array also has the same aspect ratio as the conventional image sensor being replaced. Table 1 lists example

configurations of camera arrays according to the present invention replacing conventional image sensor.

Table 1

Conventional Image Sensor		Camera array Including Imagers					
Total Mpixels	Effective Resolution	Total Mpixels	No. of Horizontal Imagers	No. of Vertical Imagers	Imager Mpixels	Super-Resolution Factor	Effective Resolution
8	3.2	8	5	5	0.32	3.2	3.2
		8	4	4	0.50	2.6	3.2
		8	3	3	0.89	1.9	3.2
5	2.0	5	5	5	0.20	3.2	2.0
		5	4	4	0.31	2.6	2.0
		5	3	3	0.56	1.9	2.0
3	1.2	3	5	5	0.12	3.2	1.2
		3	4	4	0.19	2.6	1.2
		3	3	3	0.33	1.9	1.2

[0043] The Super-Resolution Factors in Table 1 are estimates and the Effective Resolution values may differ based on the actual Super-Resolution factors achieved by processing.

[0044] The number of imagers in the camera array may be determined based on, among other factors, (i) resolution, (ii) parallax, (iii) sensitivity, and (iv) dynamic range. A first factor for the size of imager is the resolution. From a resolution point of view, the preferred number of the imagers ranges from 2 x 2 to 6 x 6 because an array size of larger than 6 x 6 is likely to destroy frequency information that cannot be recreated by the super-resolution process. For example, 8 Megapixel resolution with 2 x 2 imager will require each imager to have 2 Megapixels. Similarly, 8 Megapixel resolution with a 5 x 5 array will require each imager to have 0.32 Megapixels.

[0045] A second factor that may constrain the number of imagers is the issue of parallax and occlusion. With respect to an object captured in an image, the portion of the background scene that is occluded from the view of the imager is called as “occlusion set.” When two imagers capture the object from two different locations, the occlusion set of each imager is different. Hence, there may be scene pixels captured by one imager but not the other. To resolve this issue of occlusion, it is desirable to include a certain minimal set of imagers for a given type of imager.

[0046] A third factor that may put a lower bound on the number of imagers is the issue of sensitivity in low light conditions. To improve low light sensitivity, imagers for

detecting near-IR spectrum may be needed. The number of imagers in the camera array may need to be increased to accommodate such near-IR imagers.

[0047] A fourth factor in determining the size of the imager is dynamic range. To provide dynamic range in the camera array, it is advantageous to provide several imagers of the same filter type (chroma or luma). Each imager of the same filter type may then be operated with different exposures simultaneously. The images captured with different exposures may be processed to generate a high dynamic range image.

[0048] Based on these factors, the preferred number of imagers is 2 x 2 to 6 x 6. 4 x 4 and 5 x 5 configurations are more preferable than 2 x 2 and 3 x 3 configurations because the former are likely to provide sufficient number of imagers to resolve occlusion issues, increase sensitivity and increase the dynamic range. At the same time, the computational load required to recover resolution from these array sizes will be modest in comparison to that required in the 6x6 array. Arrays larger than 6 x 6 may, however, be used to provide additional features such as optical zooming and multispectral imaging.

[0049] Another consideration is the number of imagers dedicated to luma sampling. By ensuring that the imagers in the array dedicated to near-IR sampling do not reduce the achieved resolution, the information from the near-IR images is added to the resolution captured by the luma imagers. For this purpose, at least 50% of the imagers may be used for sampling the luma and/or near-IR spectra. In one embodiment with 4 x 4 imagers, 4 imagers samples luma, 4 imagers samples near-IR, and the remaining 8 imagers samples two chroma (Red and Blue). In another embodiment with 5 x 5 imagers, 9 imagers samples luma, 8 imagers samples near-IR, and the remaining 8 imagers samples two chroma (Red and Blue). Further, the imagers with these filters may be arranged symmetrically within the camera array to address occlusion due to parallax.

[0050] In one embodiment, the imagers in the camera array are spatially separated from each other by a predetermined distance. By increasing the spatial separation, the parallax between the images captured by the imagers may be increased. The increased parallax is advantageous where more accurate distance information is important. Separation between two imagers may also be increased to approximate the separation of a pair of human eyes. By approximating the separation of human eyes, a realistic stereoscopic 3D image may be provided to present the resulting image on an appropriate 3D display device.

[0051] In one embodiment, multiple camera arrays are provided at different locations on a device to overcome space constraints. One camera array may be designed to fit within a restricted space while another camera array may be placed in another restricted space of the

device. For example, if a total of 20 imagers are required but the available space allows only a camera array of 1 x 10 imagers to be provided on either side of a device, two camera arrays each including 10 imagers may be placed on available space at both sides of the device. Each camera array may be fabricated on a substrate and be secured to a motherboard or other parts of a device. The images collected from multiple camera arrays may be processed to generate images of desired resolution and performance.

[0052] A design for a single imager may be applied to different camera arrays each including other types of imagers. Other variables in the camera array such as spatial distances, color filters and combination with the same or other sensors may be modified to produce a camera array with differing imaging characteristics. In this way, a diverse mix of camera arrays may be produced while maintaining the benefits from economies of scale.

Wafer Level Optics Integration

[0053] In one embodiment, the camera array employs wafer level optics (WLO) technology. WLO is a technology that molds optics on glass wafers followed by packaging of the optics directly with the imager into a monolithic integrated module. The WLO procedure may involve, among other procedures, using a diamond-turned mold to create each plastic lens element on a glass substrate.

[0054] Figure 2A is a perspective view of a camera array assembly 200 with wafer level optics 210 and a camera array 230, according to one embodiment. The wafer level optics 210 includes a plurality of lens elements 220, each lens element 220 covering one of twenty-five imagers 240 in the camera array 230. Note that the camera array assembly 200 has an array of smaller lens elements occupy much less space compared to a single large lens covering the entire camera array 230.

[0055] Figure 2B is a sectional view of a camera array assembly 250, according to one embodiment. The camera assembly 250 includes a top lens wafer 262, a bottom lens wafer 268, a substrate 278 with multiple imagers formed thereon and spacers 258, 264 270. The camera array assembly 250 is packaged within an encapsulation 254. A top spacer 258 is placed between the encapsulation 254 and the top lens wafer 262. Multiple optical elements 288 are formed on the top lens wafer 262. A middle spacer 264 is placed between the top lens wafer 262 and a bottom lens wafer 268. Another set of optical elements 286 is formed on the bottom lens wafer 268. A bottom spacer 270 is placed between the bottom lens wafer 268 and the substrate 278. Through-silicon vias 274 are also provided to paths for transmitting signal from the imagers. The top lens wafer 262 may be partially coated with light blocking materials 284 (e.g., chromium) to block of light. The portions of the top lens

wafer 262 not coated with the blocking materials 284 serve as apertures through which light passes to the bottom lens wafer 268 and the imagers. In the embodiment of Figure 2B, filters 282 are formed on the bottom lens wafer 268. Light blocking materials 280 (e.g., chromium) may also be coated on the bottom lens 268 and the substrate 278 to function as an optical isolator. The bottom surface of the surface is covered with a backside redistribution layer ("RDL") and solder balls 276.

[0056] In one embodiment, the camera array assembly 250 includes 5 x 5 array of imagers. The camera array 250 has a width W of 7.2mm, and a length of 8.6mm. Each imager in the camera array may have a width S of 1.4mm. The total height t_1 of the optical components is approximately 1.26mm and the total height t_2 the camera array assembly is less than 2mm.

[0057] Figures 3A and 3B are diagrams illustrating changes in the height t of a lens element pursuant to changes in dimensions in an x-y plane. A lens element 320 in Figure 3B is scaled by $1/n$ compared to a lens element 310 in Figure 3A. As the diameter L/n of the lens element 320 is smaller than the diameter L by a factor of n , the height t/n of the lens element 320 is also smaller than the height t of the lens element 310 by a factor of n . Hence, by using an array of smaller lens elements, the height of the camera array assembly can be reduced significantly. The reduced height of the camera array assembly may be used to design less aggressive lenses having better optical properties such as improved chief ray angle, reduced distortion, and improved color aberration.

[0058] Figure 3C illustrates improving a chief ray angle (CRA) by reducing the thickness of the camera array assembly. CRA1 is the chief ray angle for a single lens covering an entire camera array. Although the chief ray angle can be reduced by increasing the distance between the camera array and the lens, the thickness constraints imposes constraints on increasing the distance. Hence, the CRA1 for camera array having a single lens element is large, resulting in reduced optical performance. CRA2 is the chief ray angle for an imager in the camera array that is scaled in thickness as well as other dimensions. The CRA2 remains the same as the CRA1 of the conventional camera array and results in no improvement in the chief ray angle. By modifying the distance between the imager and the lens element as illustrated in Figure 3C, however, the chief ray angle CRA3 in the camera array assembly may be reduced compared to CRA1 or CRA2, resulting in better optical performance. As described above, the camera arrays according to the present invention has reduced thickness requirements, and therefore, the distance of the lens element and the camera array may be increased to improve the chief ray angle.

[0059] In addition, the lens elements are subject to less rigorous design constraints yet produces better or equivalent performance compared to conventional lens element covering a wide light spectrum because each lens element may be designed to direct a narrow band of light. For example, an imager receiving visible or near-IR spectrum may have a lens element specifically optimized for this spectral band of light. For imagers detecting other light spectrum, the lens element may have differing focal lengths so that the focal plane is the same for different spectral bands of light. The matching of the focal plane across different wavelengths of light increases the sharpness of image captured at the imager and reduces longitudinal chromatic aberration.

[0060] Other advantages of smaller lens element include, among others, reduced cost, reduced amount of materials, and the reduction in the manufacturing steps. By providing n^2 lenses that are $1/n$ the size in x and y dimension (and thus $1/n$ thickness), the wafer size for producing the lens element may also be reduced. This reduces the cost and the amount of materials considerably. Further, the number of lens substrate is reduced, which results in reduced number of manufacturing steps and reduced attendant yield costs. The placement accuracy required to register the lens array to the imagers is typically no more stringent than in the case of a conventional imager because the pixel size for the camera array according to the present invention may be substantially same as a conventional image sensor

[0061] In one embodiment, the WLO fabrication process includes: (i) incorporating lens element stops by plating the lens element stops onto the substrate before lens molding, and (ii) etching holes in the substrate and performing two-sided molding of lenses through the substrate. The etching of holes in the substrate is advantageous because index mismatch is not caused between plastic and substrate. In this way, light absorbing substrate that forms natural stops for all lens elements (similar to painting lens edges black) may be used.

[0062] In one embodiment, filters are part of the imager. In another embodiment, filters are part of a WLO subsystem.

Imaging System and Processing Pipeline

[0063] Figure 4 is a functional block diagram illustrating an imaging system 400, according to one embodiment. The imaging system 400 may include, among other components, the camera array 410, an image processing pipeline module 420 and a controller 440. The camera array 410 includes two or more imagers, as described above in detail with reference to Figures 1 and 2. Images 412 are captured by the two or more imagers in the camera array 410.

[0064] The controller 440 is hardware, software, firmware or a combination thereof for controlling various operation parameters of the camera array 410. The controller 440 receives inputs 446 from a user or other external components and sends operation signals 442 to control the camera array 410. The controller 440 may also send information 444 to the image processing pipeline module 420 to assist processing of the images 412.

[0065] The image processing pipeline module 420 is hardware, firmware, software or a combination for processing the images received from the camera array 410. The image processing pipeline module 420 processes multiple images 412, for example, as described below in detail with reference to Figure 5. The processed image 422 is then sent for display, storage, transmittal or further processing.

[0066] Figure 5 is a functional block diagram illustrating the image processing pipeline module 420, according to one embodiment. The image processing pipeline module 420 may include, among other components, an upstream pipeline processing module 510, an image pixel correlation module 514, a parallax confirmation and measurement module 518, a parallax compensation module 522, a super-resolution module 526, an address conversion module 530, an address and phase offset calibration module 554, and a downstream color processing module 564.

[0067] The address and phase offset calibration module 554 is a storage device for storing calibration data produced during camera array characterization in the manufacturing process or a subsequent recalibration process. The calibration data indicates mapping between the addresses of physical pixels 572 in the imagers and the logical addresses 546, 548 of an image.

[0068] The address conversion module 530 performs normalization based on the calibration data stored in the address and phase offset calibration module 554. Specifically, the address conversion module 530 converts “physical” addresses of the individual pixels in the image to “logical” addresses 548 of the individual pixels in the imagers or vice versa. In order for super-resolution processing to produce an image of enhanced resolution, the phase difference between corresponding pixels in the individual imagers needs to be resolved. The super-resolution process may assume that for each pixel in the resulting image the set of input pixels from each of the imager is consistently mapped and that the phase offset for each imager is already known with respect to the position of the pixel in the resulting image. The address conversion module 530 resolves such phase differences by converting the physical addresses in the images 412 into logical addresses 548 of the resulting image for subsequent processing.

[0069] The images 412 captured by the imagers 540 are provided to the upstream pipeline processing module 510. The upstream pipe processing module 510 may perform one or more of Black Level calculation and adjustments, fixed noise compensation, optical PSF (point spread function) deconvolution, noise reduction, and crosstalk reduction. After the image is processed by the upstream pipeline processing module 510, an image pixel correlation module 514 performs calculation to account for parallax that becomes more apparent as objects being captured approaches to the camera array. Specifically, the image pixel correlation module 514 aligns portions of images captured by different imagers to compensate for the parallax. In one embodiment, the image pixel correlation module 514 compares the difference between the average values of neighboring pixels with a threshold and flags the potential presence of parallax when the difference exceeds the threshold. The threshold may change dynamically as a function of the operating conditions of the camera array. Further, the neighborhood calculations may also be adaptive and reflect the particular operating conditions of the selected imagers.

[0070] The image is then processed by the parallax confirmation and measurement module 518 to detect and meter the parallax. In one embodiment, parallax detection is accomplished by a running pixel correlation monitor. This operation takes place in logical pixel space across the imagers with similar integration time conditions. When the scene is at practical infinity, the data from the imagers is highly correlated and subject only to noise-based variations. When an object is close enough to the camera, however, a parallax effect is introduced that changes the correlation between the imagers. Due to the spatial layout of the imagers, the nature of the parallax-induced change is consistent across all imagers. Within the limits of the measurement accuracy, the correlation difference between any pair of imagers dictates the difference between any other pair of imagers and the differences across the other imagers. This redundancy of information enables highly accurate parallax confirmation and measurement by performing the same or similar calculations on other pairs of imagers. If parallax is present in the other pairs, the parallax should occur at roughly the same physical location of the scene taking into account the positions of the imagers. The measurement of the parallax may be accomplished at the same time by keeping track of the various pair-wise measurements and calculating an “actual” parallax difference as a least squares (or similar statistic) fit to the sample data. Other methods for detecting the parallax may include detecting and tracking vertical and horizontal high-frequency image elements from frame-to-frame.

[0071] The parallax compensation module 522 processes images including objects close enough to the camera array to induce parallax differences larger than the accuracy of the phase offset information required by super resolution process. The parallax compensation module 522 uses the scan-line based parallax information generated in the parallax detection and measurement module 518 to further adjust mapping between physical pixel addresses and logical pixel addresses before the super-resolution process. There are two cases that occur during this processing. In a more common case, addressing and offsetting adjustment are required when the input pixels have shifted positions relative to the image-wise-corresponding pixels in other imagers. In this case, no further processing with respect to parallax is required before performing super-resolution. In a less common case, a pixel or group of pixels are shifted in such a way that exposes the occlusion set. In this case, the parallax compensation process generates tagged pixel data indicating that the pixels of the occlusion set should not be considered in the super-resolution process.

[0072] After the parallax change has been accurately determined for a particular imager, the parallax information 524 is sent to the address conversion module 530. The address conversion module 530 uses the parallax information 524 along with the calibration data 558 from the address and phase offset calibration module 554 to determine the appropriate X and Y offsets to be applied to logical pixel address calculations. The address conversion module 530 also determines the associated sub-pixel offset for a particular imager pixel with respect to pixels in the resulting image 428 produced by the super-resolution process. The address conversion module 530 takes into account the parallax information 524 and provides logical addresses 546 accounting for the parallax.

[0073] After performing the parallax compensation, the image is processed by the super-resolution module 526 to obtain a high resolution synthesized image 422 from low resolution images, as described below in detail. The synthesized image 422 may then be fed to the downstream color processing module 564 to perform one or more of the following operations: focus recover, white balance, color correction, gamma correction, RGB to YUV correction, edge-aware sharpening, contrast enhancement and compression.

[0074] The image processing pipeline module 420 may include components for additional processing of the image. For example, the image processing pipeline module 420 may include a correction module for correcting abnormalities in images caused by a single pixel defect or a cluster of pixel defects. The correction module may be embodied on the same chip as the camera array, as a component separate from the camera array or as a part of the super-resolution module 526.

Super-Resolution Processing

[0075] In one embodiment, the super-resolution module 526 generates a higher resolution synthesized image by processing low resolution images captured by the imagers 540. The overall image quality of the synthesized image is higher than images captured from any one of the imagers individually. In other words, the individual imagers operate synergistically, each contributing to higher quality images using their ability to capture a narrow part of the spectrum without sub-sampling. The image formation associated with the super-resolution techniques may be expressed as follows:

$$y_k = W_k \cdot x + n_k, \forall k = 1 \dots p \quad \dots \dots \dots \text{equation (2)}$$

where W_k represents the contribution of the HR scene (x) (via blurring, motion, and sub-sampling) to each of the LR images (y_k) captured on each of the k imagers and n_k is the noise contribution.

[0076] Figures 6A through 6E illustrate various configurations of imagers for obtaining a high resolution image through a super-resolution process, according to embodiments of the present invention. In Figures 6A through 4E, “R” represents an imager having a red filter, “G” represents a imager having a green filter, “B” represents an imager having a blue filter, “P” represents a polychromatic imager having sensitivity across the entire visible spectra and near-IR spectrum, and “I” represents an imager having a near-IR filter. The polychromatic imager may sample image from all parts of the visible spectra and the near-IR region (i.e., from 650nm to 800nm). In the embodiment of Figure 6A, the center columns and rows of the imagers include polychromatic imagers. The remaining areas of the camera array are filled with imagers having green filters, blue filters, and red filters. The embodiment of Figure 6A does not include any imagers for detecting near-IR spectrum alone.

[0077] The embodiment of Figure 6B has a configuration similar to conventional Bayer filter mapping. This embodiment does not include any polychromatic imagers or near-IR imagers. As described above in detail with reference to Figure 1, the embodiment of Figure 6B is different from conventional Bayer filter configuration in that each color filter is mapped to each imager instead of being mapped to an individual pixel.

[0078] Figure 6C illustrates an embodiment where the polychromatic imagers form a symmetric checkerboard pattern. Figure 6D illustrates an embodiment where four near-IR imagers are provided. Figure 6E illustrates an embodiment with irregular mapping of imagers. The embodiments of Figures 6A through 6E are merely illustrative and various other layouts of imagers can also be used.

[0079] The use of polychromatic imagers and near-IR imagers is advantageous because these sensors may capture high quality images in low lighting conditions. The images captured by the polychromatic imager or the near-IR imager are used to denoise the images obtained from regular color imagers.

[0080] The premise of increasing resolution by aggregating multiple low resolution images is based on the fact that the different low resolution images represent slightly different viewpoints of the same scene. If the LR images are all shifted by integer units of a pixel, then each image contains essentially the same information. Therefore, there is no new information in LR images that can be used to create the HR image. In the imagers according to embodiments, the layout of the imagers may be preset and controlled so that each imager in a row or a column is a fixed sub-pixel distance from its neighboring imagers. The wafer level manufacturing and packaging process allows accurate formation of imagers to attain the sub-pixel precisions required for the super-resolution processing.

[0081] An issue of separating the spectral sensing elements into different imagers is parallax caused by the physical separation of the imagers. By ensuring that the imagers are symmetrically placed, at least two imagers can capture the pixels around the edge of a foreground object. In this way, the pixels around the edge of a foreground object may be aggregated to increase resolution as well as avoiding any occlusions. Another issue related to parallax is the sampling of color. The issue of sampling the color may be reduced by using parallax information in the polychromatic imagers to improve the accuracy of the sampling of color from the color filtered imagers.

[0082] In one embodiment, near-IR imagers are used to determine relative luminance differences compared to a visible spectra imager. Objects have differing material reflectivity results in differences in the images captured by the visible spectra and the near-IR spectra. At low lighting conditions, the near-IR imager exhibits a higher signal to noise ratios. Therefore, the signals from the near-IR sensor may be used to enhance the luminance image. The transferring of details from the near-IR image to the luminance image may be performed before aggregating spectral images from different imagers through the super-resolution process. In this way, edge information about the scene may be improved to construct edge-preserving images that can be used effectively in the super-resolution process. The advantage of using near-IR imagers is apparent from equation (2) where any improvement in the estimate for the noise (i.e., n) leads to a better estimate of the original HR scene (x).

[0083] Figure 7 is a flowchart illustrating a process of generating an HR image from LR images captured by a plurality of imagers, according to one embodiment. First, luma

images, near-IR images and chroma images are captured 710 by imagers in the camera array. Then normalization is performed 714 on the captured images to map physical addresses of the imagers to logical addresses in the enhanced image. Parallax compensation is then performed 720 to resolve any differences in the field-of-views of the imagers due to spatial separations between the imagers. Super-resolution processing is then performed 724 to obtain super-resolved luma images, super-resolved near-IR images, and super-resolved chroma images.

[0084] Then it is determined 728 if the lighting condition is better than a preset parameter. If the lighting condition is better than the parameter, the process proceeds to normalize 730 a super-resolved near-IR image with respect to a super-resolved luma image. A focus recovery is then performed 742. In one embodiment, the focus recovery is performed 742 using PSF (point spread function) deblurring per each channel. Then the super-resolution is processed 746 based on near-IR images and the luma images. A synthesized image is then constructed 750.

[0085] If it is determined 728 that the lighting condition is not better than the preset parameter, the super-resolved near-IR images and luma images are aligned 734. Then the super-resolved luma images are denoised 738 using the near-IR super-resolved images. Then the process proceeds to performing focus recovery 742 and repeats the same process as when the lighting condition is better than the preset parameter. Then the process terminates.

Image Fusion of Color Images with Near-IR Images

[0086] The spectral response of CMOS imagers is typically very good in the near-IR regions covering 650nm to 800nm and reasonably good between 800nm and 1000nm. Although near-IR images having no chroma information, information in this spectral region is useful in low lighting conditions because the near-IR images are relatively free of noise. Hence, the near-IR images may be used to denoise color images under the low lighting conditions.

[0087] In one embodiment, an image from a near-IR imager is fused with another image from a visible light imager. Before proceeding with the fusion, a registration is performed between the near-IR image and the visible light image to resolve differences in viewpoints. The registration process may be performed in an offline, one-time, processing step. After the registration is performed, the luminance information on the near-IR image is interpolated to grid points that correspond to each grid point on the visible light image.

[0088] After the pixel correspondence between the near-IR image and the visible light image is established, denoising and detail transfer process may be performed. The denoising

process allows transfer of signal information from the near-IR image to the visible light image to improve the overall SNR of the fusion image. The detail transfer ensures that edges in the near-IR image and the visible light image are preserved and accentuated to improve the overall visibility of objects in the fused image.

[0089] In one embodiment, a near-IR flash may serve as a near-IR light source during capturing of an image by the near-IR imagers. Using the near-IR flash is advantageous, among other reasons, because (i) the harsh lighting on objects of interest may be prevented, (ii) ambient color of the object may be preserved, and (iii) red-eye effect may be prevented.

[0090] In one embodiment, a visible light filter that allows only near-IR rays to pass through is used to further optimize the optics for near-IR imaging. The visible light filter improves the near-IR optics transfer function because the light filter results in sharper details in the near-IR image. The details may then be transferred to the visible light images using a dual bilateral filter as described, for example, in Eric P. Bennett et al., "Multispectral Video Fusion," Computer Graphics (ACM SIGGRAPH Proceedings) (July 25, 2006), which is incorporated by reference herein in its entirety.

Dynamic Range Determination by Differing Exposures at Imagers

[0091] An auto-exposure (AE) algorithm is important to obtaining an appropriate exposure for the scene to be captured. The design of the AE algorithm affects the dynamic range of captured images. The AE algorithm determines an exposure value that allows the acquired image to fall in the linear region of the camera array's sensitivity range. The linear region is preferred because a good signal-to-noise ratio is obtained in this region. If the exposure is too low, the picture becomes under-saturated while if the exposure is too high the picture becomes over-saturated. In conventional cameras, an iterative process is taken to reduce the difference between measured picture brightness and previously defined brightness below a threshold. This iterative process requires a large amount of time for convergence, and sometimes results in an unacceptable shutter delay.

[0092] In one embodiment, the picture brightness of images captured by a plurality of imagers is independently measured. Specifically, a plurality of imagers are set to capturing images with different exposures to reduce the time for computing the adequate exposure. For example, in a camera array with 5 x 5 imagers where 8 luma imagers and 9 near-IR imagers are provided, each of the imagers may be set with different exposures. The near-IR imagers are used to capture low-light aspects of the scene and the luma imagers are used to capture the high illumination aspects of the scene. This results in a total of 17 possible exposures. If exposure for each imager is offset from an adjacent imager by a factor of 2, for example, a

maximum dynamic range of 2^{17} or 102dB can be captured. This maximum dynamic range is considerably higher than the typical 48dB attainable in a conventional camera with 8bit image outputs.

[0093] At each time instant, the responses (under-exposed, over-exposed or optimal) from each of the multiple imagers are analyzed based on how many exposures are needed at the subsequent time instant. The ability to query multiple exposures simultaneously in the range of possible exposures accelerates the search compared to the case where only one exposure is tested at once. By reducing the processing time for determining the adequate exposure, shutter delays and shot-to-shot lags may be reduced.

[0094] In one embodiment, the HDR image is synthesized from multiple exposures by combining the images after linearizing the imager response for each exposure. The images from the imagers may be registered before combining to account for the difference in the viewpoints of the imagers.

[0095] In one embodiment, at least one imager includes HDR pixels to generate HDR images. HDR pixels are specialized pixels that capture high dynamic range scenes. Although HDR pixels show superior performances compared to other pixels, HDR pixels show poor performance at low lighting conditions in comparison with near-IR imagers. To improve performance at low lighting conditions, signals from the near-IR imagers may be used in conjunction with the signal from the HDR imager to attain better quality images across different lighting conditions.

[0096] In one embodiment, an HDR image is obtained by processing images captured by multiple imagers by processing, as disclosed, for example, in Paul Debevec et al., "Recovering High Dynamic Range Radiance Maps from Photographs," Computer Graphics (ACM SIGGRAPH Proceedings), (August 16, 1997), which is incorporated by reference herein in its entirety. The ability to capture multiple exposures simultaneously using the imager is advantageous because artifacts caused by motion of objects in the scene can be mitigated or eliminated.

Hyperspectral Imaging by Multiple Imagers

[0097] In one embodiment, a multi-spectral image is rendered by multiple imagers to facilitate the segmentation or recognition of objects in a scene. Because the spectral reflectance coefficients vary smoothly in most real world objects, the spectral reflectance coefficients may be estimated by capturing the scene in multiple spectral dimensions using imagers with different color filters and analyzing the captured images using Principal Components Analysis (PCA).

[0098] In one embodiment, half of the imagers in the camera array are devoted to sampling in the basic spectral dimensions (R, G, and B) and the other half of the imagers are devoted to sampling in a shifted basic spectral dimensions (R', G', and B'). The shifted basic spectral dimensions are shifted from the basic spectral dimensions by a certain wavelength (e.g., 10nm).

[0099] In one embodiment, pixel correspondence and non-linear interpolation is performed to account for the sub-pixel shifted views of the scene. Then the spectral reflectance coefficients of the scene are synthesized using a set of orthogonal spectral basis functions as disclosed, for example, in J.P.S. Parkkinen, J. Hallikainen and T. Jaaskelainen, "Characteristic Spectra of Munsell Colors," J. Opt. Soc. Am., A 6:318 (August 1989), which is incorporated by reference herein in its entirety. The basis functions are eigenvectors derived by PCA of a correlation matrix and the correlation matrix is derived from a database storing spectral reflectance coefficients measured by, for example, Munsell color chips (a total of 1257) representing the spectral distribution of a wide range of real world materials to reconstruct the spectrum at each point in the scene.

[0100] At first glance, capturing different spectral images of the scene through different imagers in the camera array appears to trade resolution for higher dimensional spectral sampling. However, some of the lost resolution may be recovered. The multiple imagers sample the scene over different spectral dimensions where each sampling grid of each imager is offset by a sub-pixel shift from the others. In one embodiment, no two sampling grid of the imager overlap. That is, the superposition of all the sampling grids from all the imagers forms a dense, possibly non-uniform, montage of points. Scattered data interpolation methods may be used to determine the spectral density at each sample point in this non-uniform montage for each spectral image, as described, for example, in Shiaofen Fang et al., "Volume Morphing Methods for Landmark Based 3D Image Deformation" by SPIE vol. 2710, proc. 1996 SPIE Intl Symposium on Medical Imaging, page 404-415, Newport Beach, CA (February 1996), which is incorporated by reference herein in its entirety. In this way, a certain amount of resolution lost in the process of sampling the scene using different spectral filters may be recovered.

[0101] As described above, image segmentation and object recognition are facilitated by determining the spectral reflectance coefficients of the object. The situation often arises in security applications wherein a network of cameras is used to track an object as it moves from the operational zone of one camera to another. Each zone may have its own unique lighting conditions (fluorescent, incandescent, D65, etc.) that may cause the object to have a

different appearance in each image captured by different cameras. If these cameras capture the images in a hyper-spectral mode, all images may be converted to the same illuminant to enhance object recognition performance.

[0102] In one embodiment, camera arrays with multiple imagers are used for providing medical diagnostic images. Full spectral digitized images of diagnostic samples contribute to accurate diagnosis because doctors and medical personnel can place higher confidence in the resulting diagnosis. The imagers in the camera arrays may be provided with color filters to provide full spectral data. Such camera array may be installed on cell phones to capture and transmit diagnostic information to remote locations as described, for example, in Andres W. Martinez et al., "Simple Telemedicine for Developing Regions: Camera Phones and Paper-Based Microfluidic Devices for Real-Time, Off-Site Diagnosis," Analytical Chemistry (American Chemical Society) (April 11, 2008), which is incorporated by reference herein in its entirety. Further, the camera arrays including multiple imagers may provide images with a large depth of field to enhance the reliability of image capture of wounds, rashes, and other symptoms.

[0103] In one embodiment, a small imager (including, for example, 20-500 pixels) with a narrow spectral bandpass filters is used to produce a signature of the ambient and local light sources in a scene. By using the small imager, the exposure and white balance characteristics may be determined more accurately at a faster speed. The spectral bandpass filters may be ordinary color filters or diffractive elements of a bandpass width adequate to allow the number of camera arrays to cover the visible spectrum of about 400nm. These imagers may run at a much higher frame rate and obtain data (which may or may not be used for its pictorial content) for processing into information to control the exposure and white balance of other larger imagers in the same camera array. The small imagers may also be interspersed within the camera array.

Optical Zoom Implemented Using Multiple Imagers

[0104] In one embodiment, a subset of imagers in the camera array includes telephoto lenses. The subset of imagers may have other imaging characteristics same as imagers with non-telephoto lenses. Images from this subset of imagers are combined and super-resolution processed to form a super-resolution telephoto image. In another embodiment, the camera array includes two or more subsets of imagers equipped with lenses of more than two magnifications to provide differing zoom magnifications.

[0105] Embodiments of the camera arrays may achieve its final resolution by aggregating images through super-resolution. Taking an example of providing 5 x 5 imagers

with a 3X optical zoom feature, if 17 imagers are used to sample the luma (G) and 8 imagers are used to sample the chroma (R and B), 17 luma imagers allow a resolution that is four times higher than what is achieved by any single imager in the set of 17 imagers. If the number of the imager is increased from 5 x 5 to 6 x 6, an addition of 11 extra imagers becomes available. In comparison with the 8 Megapixel conventional image sensor fitted with a 3X zoom lens, a resolution that is 60% of the conventional image sensor is achieved when 8 of the additional 11 imagers are dedicated to sampling luma (G) and the remaining 3 imagers are dedicated to chroma (R and B) and near-IR sampling at 3X zoom. This considerably reduces the chroma sampling (or near-IR sampling) to luma sampling ratio. The reduced chroma to luma sampling ratio is somewhat offset by using the super-resolved luma image at 3X zoom as a recognition prior on the chroma (and near-IR) image to resample the chroma image at a higher resolution.

[0106] With 6 x 6 imagers, a resolution equivalent to the resolution of conventional image sensor is achieved at 1X zoom. At 3X zoom, a resolution equivalent to about 60% of conventional image sensor outfitted with a 3X zoom lens is obtained by the same imagers. Also, there is a decrease in luma resolution at 3X zoom compared with conventional image sensors with resolution at 3X zoom. The decreased luma resolution, however, is offset by the fact that the optics of conventional image sensor has reduced efficiency at 3X zoom due to crosstalk and optical aberrations.

[0107] The zoom operation achieved by multiple imagers has the following advantages. First, the quality of the achieved zoom is considerably higher than what is achieved in the conventional image sensor due to the fact that the lens elements may be tailored for each change in focal length. In conventional image sensors, optical aberrations and field curvature must be corrected across the whole operating range of the lens, which is considerably harder in a zoom lens with moving elements than in a fixed lens element where only aberrations for a fixed focal length need to be corrected. Additionally, the fixed lens in the imagers has a fixed chief ray angle for a given height, which is not the case with conventional image sensor with a moving zoom lens. Second, the imagers allow simulation of zoom lenses without significantly increasing the optical track height. The reduced height allows implementation of thin modules even for camera arrays with zooming capability.

[0108] The overhead required to support a certain level of optical zoom in camera arrays according to some embodiments is tabulated in Table 2.

Table 2

No. of Imagers in Camera array	No. of Luma Imagers at different Zoom levels			No. of Chroma Imagers at different Zoom Levels		
	1X	2X	3X	1X	2X	3X
25	17	0	0	8	0	0
36	16	0	8	8	0	4

[0109] In one embodiment, the pixels in the images are mapped onto an output image with a size and resolution corresponding to the amount of zoom desired in order to provide a smooth zoom capability from the widest-angle view to the greatest-magnification view. Assuming that the higher magnification lenses have the same center of view as the lower magnification lenses, the image information available is such that a center area of the image has a higher resolution available than the outer area. In the case of three or more distinct magnifications, nested regions of different resolution may be provided with resolution increasing toward the center.

[0110] An image with the most telephoto effect has a resolution determined by the super-resolution ability of the imagers equipped with the telephoto lenses. An image with the widest field of view can be formatted in at least one of two following ways. First, the wide field image may be formatted as an image with a uniform resolution where the resolution is determined by the super-resolution capability of the set of imagers having the wider-angle lenses. Second, the wide field image is formatted as a higher resolution image where the resolution of the central part of the image is determined by the super-resolution capability of the set of imagers equipped with telephoto lenses. In the lower resolution regions, information from the reduced number of pixels per image area is interpolated smoothly across the larger number of “digital” pixels. In such an image, the pixel information may be processed and interpolated so that the transition from higher to lower resolution regions occurs smoothly.

[0111] In one embodiment, zooming is achieved by inducing a barrel-like distortion into some, or all, of the array lens so that a disproportionate number of the pixels are dedicated to the central part of each image. In this embodiment, every image has to be processed to remove the barrel distortion. To generate a wide angle image, pixels closer to the center are sub-sampled relative to outer pixels are super-sampled. As zooming is

performed, the pixels at the periphery of the imagers are progressively discarded and the sampling of the pixels nearer the center of the imager is increased.

[0112] In one embodiment, mipmap filters are built to allow images to be rendered at a zoom scale that is between the specific zoom range of the optical elements (e.g., 1X and 3X zoom scales of the camera array). Mipmaps are a precalculated optimized set of images that accompany a baseline image. A set of images associated with the 3X zoom luma image can be created from a baseline scale at 3X down to 1X. Each image in this set is a version of the baseline 3X zoom image but at a reduced level of detail. Rendering an image at a desired zoom level is achieved using the mipmap by (i) taking the image at 1X zoom, and computing the coverage of the scene for the desired zoom level (i.e., what pixels in the baseline image needs to be rendered at the requested scale to produce the output image), (ii) for each pixel in the coverage set, determine if the pixel is in the image covered by the 3X zoom luma image, (iii) if the pixel is available in the 3X zoom luma image, then choose the two closest mipmap images and interpolate (using smoothing filter) the corresponding pixels from the two mipmap images to produce the output image, and (iv) if the pixel is unavailable in the 3X zoom luma image, then choose the pixel from the baseline 1X luma image and scale up to the desired scale to produce the output pixel. By using mipmaps, smooth optical zoom may be simulated at any point between two given discrete levels (i.e., 1X zoom and 3X zoom).

Capturing Video Images

[0113] In one embodiment, the camera array generates high frame image sequences. The imagers in the camera array can operate independently to capture images. Compared to conventional image sensors, the camera array may capture images at the frame rate up to N time (where N is the number of imagers). Further, the frame period for each imager may overlap to improve operations under low-light conditions. To increase the resolution, a subset of imagers may operate in a synchronized manner to produce images of higher resolution. In this case, the maximum frame rate is reduced by the number of imagers operating in a synchronized manner. The high-speed video frame rates can enable slow-motion video playback at a normal video rate.

[0114] In one example, two luma imagers (green imagers or near-IR imagers), two blue imagers and two green imagers are used to obtain high-definition 1080p images. Using permutations of four luma imagers (two green imagers and two near-IR imagers or three green imagers and one near-IR imager) together with one blue imager and one red imager, the chroma imagers can be upsampled to achieve 120 frames/sec for 1080p video. For higher frame rate imaging devices, the number of frame rates can be scaled up linearly. For

Standard-Definition (480p) operation, a frame rate of 240 frames/sec may be achieved using the same camera array.

[0115] Conventional imaging devices with a high-resolution image sensor (e.g., 8 Megapixels) use binning or skipping to capture lower resolution images (e.g., 1080p30, 720p30 and 480p30). In binning, rows and columns in the captured images are interpolated in the charge, voltage or pixel domains in order to achieve the target video resolutions while reducing the noise. In skipping, rows and columns are skipped in order to reduce the power consumption of the sensor. Both of these techniques result in reduced image quality.

[0116] In one embodiment, the imagers in the camera arrays are selectively activated to capture a video image. For example, 9 imagers (including one near-IR imager) may be used to obtain 1080p (1920 x 1080 pixels) images while 6 imagers (including one near-IR imager) may be used to obtain 720p (1280 x 720 pixels) images or 4 imagers (including one near-IR imager) may be used to obtain 480p (720 x 480 pixels) images. Because there is an accurate one-to-one pixel correspondence between the imager and the target video images, the resolution achieved is higher than traditional approaches. Further, since only a subset of the imagers is activated to capture the images, significant power savings can also be achieved. For example, 60% reduction in power consumption is achieved in 1080p and 80% of power consumption is achieved in 480p.

[0117] Using the near-IR imager to capture video images is advantageous because the information from the near-IR imager may be used to denoise each video image. In this way, the camera arrays of embodiments exhibit excellent low-light sensitivity and can operate in extremely low-light conditions. In one embodiment, super-resolution processing is performed on images from multiple imagers to obtain higher resolution video imagers. The noise-reduction characteristics of the super-resolution process along with fusion of images from the near-IR imager results in a very low-noise images.

[0118] In one embodiment, high-dynamic-range (HDR) video capture is enabled by activating more imagers. For example, in a 5 x 5 camera array operating in 1080p video capture mode, there are only 9 cameras active. A subset of the 16 cameras may be overexposed and underexposed by a stop in sets of two or four to achieve a video output with a very high dynamic range.

Other Applications for Multiple Imagers

[0119] In one embodiment, the multiple imagers are used for estimating distance to an object in a scene. Since information regarding the distance to each point in an image is available in the camera array along with the extent in x and y coordinates of an image

element, the size of an image element may be determined. Further, the absolute size and shape of physical items may be measured without other reference information. For example, a picture of a foot can be taken and the resulting information may be used to accurately estimate the size of an appropriate shoe.

[0120] In one embodiment, reduction in depth of field is simulated in images captured by the camera array using distance information. The camera arrays according to the present invention produce images with greatly increased depth of field. The long depth of field, however, may not be desirable in some applications. In such case, a particular distance or several distances may be selected as the “in best focus” distance(s) for the image and based on the distance (z) information from parallax information, the image can be blurred pixel-by-pixel using, for example, a simple Gaussian blur. In one embodiment, the depth map obtained from the camera array is utilized to enable a tone mapping algorithm to perform the mapping using the depth information to guide the level, thereby emphasizing or exaggerating the 3D effect.

[0121] In one embodiment, apertures of different sizes are provided to obtain aperture diversity. The aperture size has a direct relationship with the depth of field. In miniature cameras, however, the aperture is generally made as large as possible to allow as much light to reach the camera array. Different imagers may receive light through apertures of different sizes. For imagers to produce a large depth of field, the aperture may be reduced whereas other imagers may have large apertures to maximize the light received. By fusing the images from sensor images of different aperture sizes, images with large depth of field may be obtained without sacrificing the quality of the image.

[0122] In one embodiment, the camera array according to the present invention refocuses based on images captured from offsets in viewpoints. Unlike a conventional plenoptic camera, the images obtained from the camera array of the present invention do not suffer from the extreme loss of resolution. The camera array according to the present invention, however, produces sparse data points for refocusing compared to the plenoptic camera. In order to overcome the sparse data points, interpolation may be performed to refocus data from the sparse data points.

[0123] In one embodiment, each imager in the camera array has a different centroid. That is, the optics of each imager are designed and arranged so that the fields of view for each imager slightly overlap but for the most part constitute distinct tiles of a larger field of view. The images from each of the tiles are panoramically stitched together to render a single high-resolution image.

[0124] In one embodiment, camera arrays may be formed on separate substrates and mounted on the same motherboard with spatial separation. The lens elements on each imager may be arranged so that the corner of the field of view slightly encompasses a line perpendicular to the substrate. Thus, if four imagers are mounted on the motherboard with each imager rotated 90 degrees with respect to another imager, the fields of view will be four slightly overlapping tiles. This allows a single design of WLO lens array and imager chip to be used to capture different tiles of a panoramic image.

[0125] In one embodiment, one or more sets of imagers are arranged to capture images that are stitched to produce panoramic images with overlapping fields of view while another imager or sets of imagers have a field of view that encompasses the tiled image generated. This embodiment provides different effective resolution for imagers with different characteristics. For example, it may be desirable to have more luminance resolution than chrominance resolution. Hence, several sets of imagers may detect luminance with their fields of view panoramically stitched. Fewer imagers may be used to detect chrominance with the field of view encompassing the stitched field of view of the luminance imagers.

[0126] In one embodiment, the camera array with multiple imagers is mounted on a flexible motherboard such that the motherboard can be manually bent to change the aspect ratio of the image. For example, a set of imagers can be mounted in a horizontal line on a flexible motherboard so that in the quiescent state of the motherboard, the fields of view of all of the imagers are approximately the same. If there are four imagers, an image with double the resolution of each individual imager is obtained so that details in the subject image that are half the dimension of details that can be resolved by an individual imager. If the motherboard is bent so that it forms part of a vertical cylinder, the imagers point outward. With a partial bend, the width of the subject image is doubled while the detail that can be resolved is reduced because each point in the subject image is in the field of view of two rather than four imagers. At the maximum bend, the subject image is four times wider while the detail that can be resolved in the subject is further reduced.

Offline Reconstruction and Processing

[0127] The images processed by the imaging system 400 may be previewed before or concurrently with saving of the image data on a storage medium such as a flash device or a hard disk. In one embodiment, the images or video data includes rich light field data sets and other useful image information that were originally captured by the camera array. Other traditional file formats could also be used. The stored images or video may be played back or transmitted to other devices over various wired or wireless communication methods.

[0128] In one embodiment, tools are provided for users by a remote server. The remote server may function both as a repository and an offline processing engine for the images or video. Additionally, applets mashed as part of popular photo-sharing communities such as Flickr, Picasaweb, Facebook etc. may allow images to be manipulated interactively, either individually or collaboratively. Further, software plug-ins into image editing programs may be provided to process images generated by the imaging device 400 on computing devices such as desktops and laptops.

[0129] Various modules described herein may comprise a general-purpose computer selectively activated or reconfigured by a computer program stored in the computer. Such a computer program may be stored in a computer readable storage medium, such as, but is not limited to, any type of disk including floppy disks, optical disks, CD-ROMs, magnetic-optical disks, read-only memories (ROMs), random access memories (RAMs), EPROMs, EEPROMs, magnetic or optical cards, application specific integrated circuits (ASICs), or any type of media suitable for storing electronic instructions, and each coupled to a computer system bus. Furthermore, the computers referred to in the specification may include a single processor or may be architectures employing multiple processor designs for increased computing capability.

[0130] While particular embodiments and applications of the present invention have been illustrated and described herein, it is to be understood that the invention is not limited to the precise construction and components disclosed herein and that various modifications, changes, and variations may be made in the arrangement, operation, and details of the methods and apparatuses of the present invention without departing from the spirit and scope of the invention as it is defined in the appended claims.

CLAIMS

What is claimed is:

1. A camera array fabricated on a semiconductor substrate to include a plurality of sensor elements, comprising:
 - a first imager fabricated at a first location of the semiconductor substrate, the first imager including at least two first image sensor elements; and
 - a second imager fabricated on a second location of the semiconductor substrate, the second imager including at least two second image sensor elements not overlapping with first image sensor elements.
2. The camera array of claim 1, wherein the first imager has first imaging characteristics and the second imager has second imaging characteristics different from the first imaging characteristics.
3. The camera array of claim 2, wherein the first imager comprises a first optical filter transmitting a first light spectrum, and the second imager comprises a second optical filter transmitting a second light spectrum.
4. The camera array of claim 3, further comprising an optical element comprising:
 - a first lens element having first configuration for focusing the first light spectrum onto the first imager; and
 - a second lens element having second configuration for focusing the second light spectrum onto the second imager.
5. The camera array of claim 1, further comprising an optical element fabricated on the semiconductor substrate using wafer level optics (WLO) technology.
6. The camera array of claim 1, wherein the first imager generates a first image covering a first field of view and the second imager generates a second image covering a second field of view, the second field of view shifted by a distance less than a size of a sensor element with respect to the first field of view.
7. The camera array of claim 2, wherein the first imaging characteristics comprise at least one of a size of the first imager, a type of sensor elements included in the

first imager, a shape of the first imager, a filter associated with the first imager, exposure time of the first imager, aperture size associated with the first imager, a configuration of the optical element associated with the first imager, gain of the first imager, resolution of the first imager, and operational timing of the first imager.

8. The camera array of claim 1, wherein the first imager generates an image for a color component, and the second imager generates an image for near-IR (Infrared) component.

9. The camera array of claim 1, wherein the first imager is triggered to capture a first image at a first time, and the second imager is triggered to capture a second image at a second time.

10. An imaging device for generating an image, comprising:
a camera array fabricated on a semiconductor substrate to include a plurality of sensor elements, the camera array comprising a first group of first imagers on the semiconductor substrate and a second group of second imagers on the semiconductor substrate, each of the first and second imagers including a plurality of image sensor elements, each of the first imager having first imaging characteristics, each of the second imager having second imaging characteristics; and
an image processing module coupled to the camera array for receiving first images captured by first imagers and second images captured by second imagers, the image processing module generating a synthesized image based on the first images and the second images.

11. The imaging device of claim 10, wherein at least one of the first imagers captures a first field of view shifted with respect to a second field of view by a distance less than a size of a sensor element, the second field of view captured by another of the first imagers.

12. The imaging device of claim 10, wherein the image processing module performs super-resolution on the first images and the second images to obtain the synthesized image having a resolution higher than a resolution of the first images or the second images.

13. The imaging device of claim 10, wherein the first group of imagers comprise first optical filters transmitting a first light spectrum, and the second group of imagers comprise second optical filters transmitting a second light spectrum.

14. The imaging device of claim 10, wherein the camera array comprises a plurality of optical elements fabricated using wafer level optics (WLO) technology, each optical element focusing light onto each imager of the first imagers and the second imagers.

15. The imaging device of claim 10, wherein the camera array comprises:
first filters to transmit a first light spectrum;
a first set of optical elements configured to focus the first light spectrum onto the first imagers;
second filters to transmit a second light spectrum; and
a second set of optical elements configured to focus the second light spectrum onto the second imagers, the second set of optical element having configurations different from the first set of optical elements.

16. The imaging device of claim 10, wherein the first imagers generate first images for a color component, and the second imagers generate second images for near-IR (Infrared) component.

17. The imaging device of claim 10, wherein an image captured by an imager in the camera array is spatially shifted from the image captured by another imager in the camera array by a sub-sensor element distance.

18. The imaging device of claim 10, wherein a total number of imagers in the camera array is not less than 16 and not more than 36.

19. The imaging device of claim 10, wherein the image processing module comprises a module for detecting and compensating parallax caused by difference in the fields of view of imagers.

20. The imaging device of claim 19, wherein the image processing module further comprises an address conversion module mapping physical addresses of the imagers to logical addresses based on at least calibration data.

21. The imaging device of claim 10, wherein the first imagers comprise filters and the second imagers do not comprise filters.

22. The imaging device of claim 10, wherein the camera array comprises 4 imagers with green filters, 4 imagers with near-IR filters or without filters, 4 imagers with red filters, and 4 imagers with blue filters.

23. The imaging device of claim 10, wherein the camera array comprises 9 imagers with green filters, 8 imagers with near-IR filters or without filters, 4 imagers with red filters, and 4 imagers with blue filters.

24. The imaging device of claim 10, wherein the first imagers include near-IR filters and the second imagers include color filters, first images generated by the first imagers fused with second images generated by the second imagers to generate a color image with reduced noise.

25. The imaging device of claim 10, wherein the first imagers include filters for receiving one of red, green and blue light spectrum, and the second group of imagers include filters for receiving light spectrum shifted from one of red, green or blue light spectrum by a predetermined amount of wavelength.

26. A method of generating an image, comprising:
capturing a first image at a first imager formed on a first location of a semiconductor substrate, the first imager including a plurality of sensor elements;
capturing a second image at a second imager formed on a second location of the semiconductor substrate, the second imager including a plurality of sensor elements; and
processing the first image and the second image to obtain a third image.

27. The method of claim 26, wherein the first imager has a first imaging characteristics and the second imager has a second imaging characteristics different from the first imaging characteristics.

28. The method of claim 27, wherein the first image is a monochromatic image of a first color, and the second image is a monochromatic image of a second color.

29. The method of claim 27, further comprising:
focusing light onto the first imager by a first lens element configured to transmit a first light spectrum corresponding to the first color; and

focusing light onto the second imager by a second lens element configured to transmit a second spectrum of light corresponding to the second color.

30. The method of claim 26, wherein the first image is captured at a first exposure level and the second image is captured at a second exposure level higher than the first exposure level.

31. The method of claim 26, wherein processing the first image and the second image comprises super-resolution processing the first image and the second image.

32. The method of claim 26, wherein processing the first image and the second image comprises fusing the first image and the second image.

33. The method of claim 26, wherein processing the first image and the second image comprises stitching the first image and the second image to produce a panoramic image.