



(86) Date de dépôt PCT/PCT Filing Date: 2016/02/26  
(87) Date publication PCT/PCT Publication Date: 2016/09/09  
(85) Entrée phase nationale/National Entry: 2017/08/28  
(86) N° demande PCT/PCT Application No.: EP 2016/054041  
(87) N° publication PCT/PCT Publication No.: 2016/139134  
(30) Priorité/Priority: 2015/03/05 (EP15305340.0)

(51) Cl.Int./Int.Cl. *H04N 5/232* (2006.01),  
*H04N 13/00* (2006.01), *H04N 13/02* (2006.01)

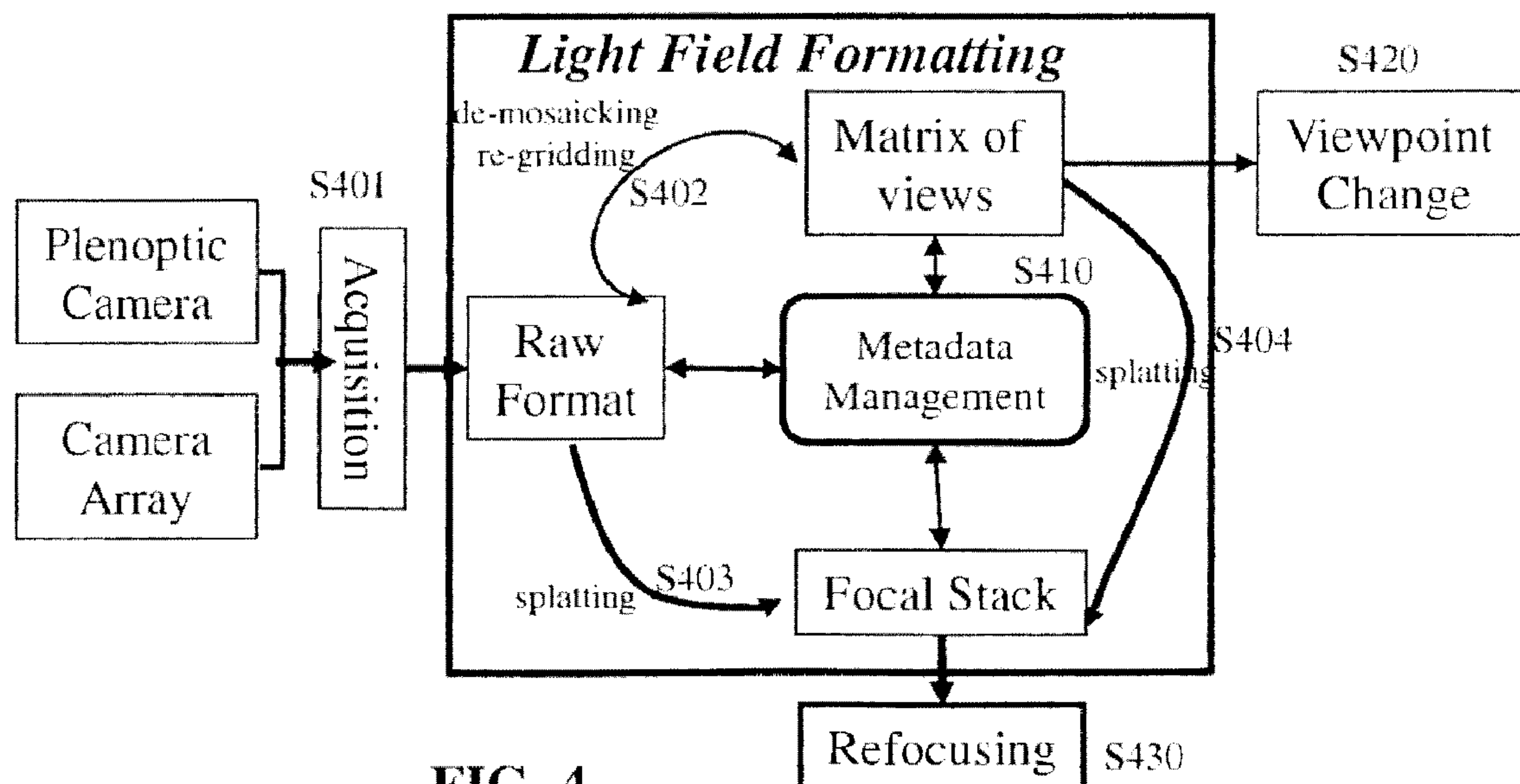
(71) Demandeur/Applicant:  
THOMSON LICENSING, FR

(72) Inventeurs/Inventors:  
BOISSON, GUILLAUME, FR;  
SEIFI, MOZHDEH, FR;  
VANDAME, BENOIT, FR;  
SABATER, NEUS, FR

(74) Agent: FETHERSTONHAUGH & CO.

(54) Titre : METADONNEES DE CHAMP LUMINEUX

(54) Title: LIGHT FIELD METADATA



**FIG. 4**

(57) Abrégé/Abstract:

A method of providing metadata for captured light field data, the method comprising providing a data format aggregating metadata describing characteristics of at least one of: a light field camera used to capture the light field data, a matrix of views derived from the light field data, and a focal stack derived from the light field data.

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

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
9 September 2016 (09.09.2016)

WIPO | PCT

(10) International Publication Number  
**WO 2016/139134 A1**

## (51) International Patent Classification:

*H04N 5/232* (2006.01) *H04N 13/02* (2006.01)  
*H04N 13/00* (2006.01)

## (21) International Application Number:

PCT/EP2016/054041

## (22) International Filing Date:

26 February 2016 (26.02.2016)

## (25) Filing Language:

English

## (26) Publication Language:

English

## (30) Priority Data:

15305340.0 5 March 2015 (05.03.2015) EP

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

(72) Inventors: BOISSON, Guillaume; 975 avenue des Champs Blancs, ZAC des Champs Blancs, CS 17616, 92443 Issy-les-Moulineaux (FR). SEIFI, Mozhdah; 975 avenue des Champs Blancs, ZAC des Champs Blancs, CS 17616, 35576 Cesson-sévigné (FR). VANDAME, Benoit; 975 avenue des Champs Blancs, ZAC des Champs Blancs, CS 17616, 35576 Cesson-sévigné (FR). SABATER, Neus; 975, avenue des Champs Blancs, ZAC des Champs Blancs, CS 176 16, 35576 Cesson-sevigne (FR).

(74) Agent: BARR, Angela; Technicolor, 1-5 rue Jeanne d'Arc, 92443 Issy-les-Moulineaux (FR).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM,

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

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

## Declarations under Rule 4.17:

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

## Published:

— with international search report (Art. 21(3))  
— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))

(54) Title: LIGHT FIELD METADATA

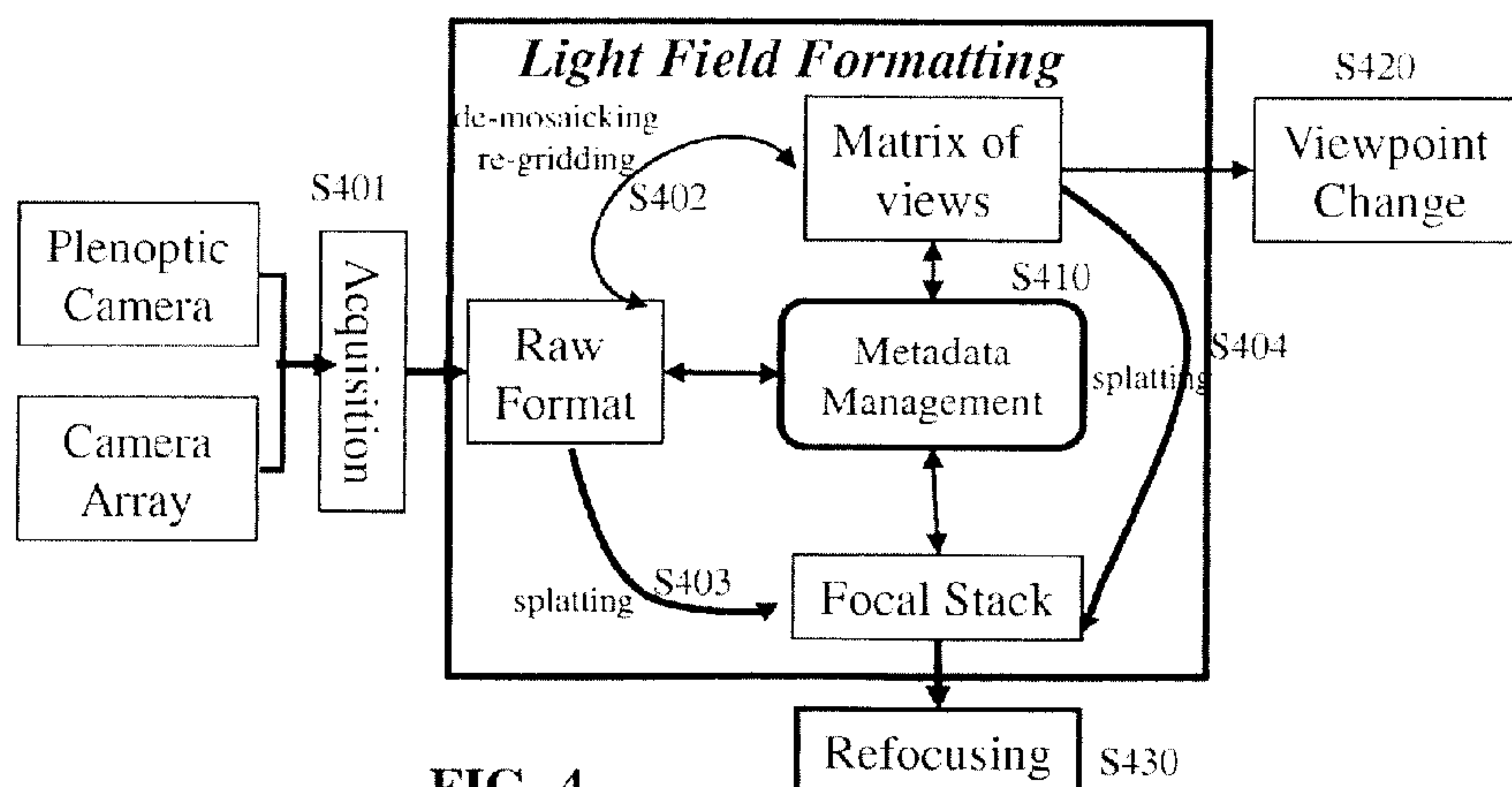


FIG. 4

(57) Abstract: A method of providing metadata for captured light field data, the method comprising providing a data format aggregating metadata describing characteristics of at least one of: a light field camera used to capture the light field data, a matrix of views derived from the light field data, and a focal stack derived from the light field data.



## LIGHT FIELD METADATA

### TECHNICAL FIELD

The present invention relates to metadata for light field data and processing of  
5 light field data.

### BACKGROUND

Conventional cameras capture light from a three-dimensional scene on a two-dimensional sensor device sensitive to visible light. Light sensitive technology used in  
10 such imaging devices is often based on semiconductor technology, capable of converting photons into electrons such as, for example, charge coupled devices (CCD) or complementary metal oxide technology (CMOS). A digital image photosensor, for example, typically includes an array of photosensitive cells, each cell being configured to capture incoming light. A 2D image providing spatial information is obtained from a  
15 measurement of the total amount of light captured by each photosensitive cell of the image sensor device. While the 2D image can provide information on the intensity of the light and the colour of the light at spatial points of the photosensor(s), no information is provided on the direction of the incoming light.

Light field cameras (also known as radiance capturing cameras) can provide  
20 directional information on incoming light rays in addition to 2D spatial information. The directional information may be obtained by the use of an array of micro-lenses, often referred to as a microlens array (MLA) associated with an image sensor. A light field array camera, for example, is provided with a microlens array and a photosensor. A plenoptic camera is provided with a main lens focusing light on a MLA, and a photosensor  
25 associated with the MLA. In other configurations of light field cameras a plurality of cameras each provided with its own lens and sensor may be used to obtain light field data.

A light field is often defined as a 4D function characterizing the light from different directions at different points in a scene. The information on the directional distribution of

the light rays is typically referred to as light field data or 4D data. Since the information provides 4 parameters – two dimensional positional information and two dimensional angular information the directional distribution corresponds to a four-dimensional (4D) function. The light field may be interpreted as a two dimensional collection of 2D images  
5 of a scene.

The light field data obtained can be processed for many applications, for example, to generate refocused images of the scene, to generate images from different viewpoints, to provide depth information on the captured scene or to generate 3D images.

Data obtained from a light field camera should be appropriately managed in order  
10 to be properly exploited in the different applications.

The present invention has been devised with the foregoing in mind.

### SUMMARY OF INVENTION

According to a first aspect of the invention there is provided a method of providing  
15 metadata for captured light field data, the method comprising providing a data format aggregating metadata describing characteristics of at least one of: a light field camera used to capture the light field data, a matrix of views derived from the light field data, and a focal stack derived from the light field data.

In an embodiment the method includes obtaining light field data metadata  
20 describing characteristics of at least one of: a light field camera used to capture the light field data, a matrix of views derived from the light field data, and a focal stack derived from the light field data; and rendering the light field data based on the information provided by the metadata.

In an embodiment, the metadata is aggregated in a light image file.

25 In an embodiment, the metadata is defined in markup language in compliance with XML.

In an embodiment, the metadata describing characteristics of the light field camera used to capture the light field data includes data defining one or more of the following technical features of the light field camera:

30 a photosensor of the light field camera;

a main lens of the light field camera; and  
a micro lens array of the light field camera.

5           In an embodiment, the metadata describing the photosensor of the light field camera includes data representative of one or more of the following parameters:

          spatial resolution of the photosensor;  
          colour sampling of a colour filter of the photosensor;  
          transmittance of the colour filter of the photosensor  
10          dynamic range of the pixels of the photosensor;  
          color transform matrix defining the non-linear relationship between the signal output by a pixel and the incoming light;  
          the pitch between two adjacent pixels of the photosensor; and  
          scaling factor specifying the shape ratio of non-square pixels.

15           In an embodiment, the metadata describing the main lens of the light field camera includes data representative of one or more of the following parameters:

          focal length of the main lens  
          the ratio between the pupil and the focal length; and  
          the vignetting model of the main lens.

20           In an embodiment, the metadata describing the micro lens array of the light field camera includes data representative of one or more of the following parameters:

          pattern of the micro-lens array;  
          pattern of color filters arranged on the micro-lens array;  
          the pitch between two adjacent micro-lenses of the micro-lens array;  
25          scaling factor specifying the shape ratio of micro lens; and  
          the number of calibrated microlens centres  
          the spatial coordinates of one or more microlens centres

          In an embodiment, the metadata describing characteristics of a matrix of views derived from the light field data, includes data representative of one or more of the  
30          following parameters:

          the number of views in the array;



the spatial resolution of a or each view;  
 colors available in a view;  
 the spatial coordinates of a reference point of a view;  
 the horizontal and vertical focal length of a view

5

In an embodiment, the metadata describing characteristics of a focal stack derived from the light field data, includes data representative of one or more of the following parameters:

- 10           the number of focused pictures in the focal stack;  
               the spatial resolution of the focused pictures  
               the distance of focalization of a focused picture

A second aspect of the invention provides a device for providing metadata for captured light field data, the device comprising:

- 15           a light field data acquisition module for acquiring light field data captured by a light field camera and  
               a data formatter for providing aggregating metadata describing characteristics of at least one of:  
               a light field camera used to capture the light field data,  
 20           a matrix of views derived from the light field data, and  
               a focal stack derived from the light field data.

A third aspect of the invention provides a light field imaging device comprising:

- an array of micro lenses arranged in a regular lattice structure;  
               a photosensor configured to capture light projected on the photosensor  
 25           from the array of micro lenses, the photosensor comprising sets of pixels, each set of pixels being optically associated with a respective micro lens of the array of micro lenses; and  
               a device for providing metadata for captured light field data, the device comprising:  
 30           a light field data acquisition module for acquiring light field data captured by a light field camera and

a data formatter for providing aggregating metadata describing characteristics of at least one of:

- a light field camera used to capture the light field data,
- a matrix of views derived from the light field data, and
- 5 a focal stack derived from the light field data.

A further aspect of the invention provides a device for rendering an image from light field data using metadata describing characteristics of at least one of:

- a light field camera used to capture the light field data,
- a matrix of views derived from the light field data, and
- 10 a focal stack derived from the light field data, the device comprising a processor and memory containing a rendering application, the light field data and the metadata wherein the rendering application configures the processor to render the light field data based on information provided in the meta data.

- Another aspect of the invention provides a data package comprising
- 15 light field data captured from a light field camera;
  - metadata describing characteristics of at least one of:
    - a light field camera used to capture the light field data,
    - a matrix of views derived from the light field data, and
    - a focal stack derived from the light field data.

- 20 According to another aspect of the invention there is provided a method of processing light field data, comprising converting light field data from a first light field data format into a second light field data format different to the first light field format, the conversion being performed using parameters obtained from metadata describing the light field data wherein the first light field data format and the second light field data format
- 25 are selected from the group consisting of: raw light field data captured by a light field camera, a matrix of sub-aperture images, and a focal stack.

- Yet another aspect of the invention provides a device for processing light field data, comprising a processor and a memory containing a format conversion application, wherein the format conversion application configures the processor to convert light field
- 30 data from a first light field data format into a second light field data format different to the first light field format, the conversion being performed using parameters obtained from

metadata describing the light field data wherein the first light field data format and the second light field data format are selected from the group consisting of: raw light field data captured by a light field camera, a matrix of sub-aperture images, and a focal stack.

In an embodiment, the memory further comprises a rendering application,  
5 wherein rendering application configures the processor to render the light field data based on information provided in metadata.

A further aspect of the invention provides a light field imaging device comprising:  
an array of micro lenses arranged in a regular lattice structure; a photosensor configured  
to capture light projected on the photosensor from the array of micro lenses, the  
10 photosensor comprising sets of pixels, each set of pixels being optically associated with  
a respective micro lens of the array of micro lenses; and a device for processing light  
field data, comprising a processor and a memory containing a format conversion  
application, wherein the format conversion application configures the processor to  
convert light field data from a first light field data format into a second light field data  
15 format different to the first light field format, the conversion being performed using  
parameters obtained from metadata describing the light field data wherein the first light  
field data format and the second light field data format are selected from the group  
consisting of: raw light field data captured by a light field camera, a matrix of sub-aperture  
images, and a focal stack.

20 In an embodiment, the light-field metadata describes characteristics of at least  
one of: a light field camera used to capture the light field data, a matrix of sub-aperture  
images derived from the light field data, and a focal stack derived from the light field  
data.

In an embodiment, the first light field data format is raw light field data, captured  
25 by a light field camera, the raw light field data comprising an array of micro-images, each  
micro image being projected by a respective micro lens of a micro-lens array and the  
second light field format is a matrix of sub-aperture images, each sub-aperture image  
being generated from collocated samples of the micro-images and wherein at least one  
of the focal lengths and the interaxial baseline distance of the matrix of sub-aperture



images is derived using technical parameters of the light field camera included in the metadata.

In an embodiment, focal lengths are derived from at least one of the focal length of the main lens of the light field camera, the horizontal and scaling factor of the  
5 photosensor of the light field camera, and the pixel pitch of the photosensor.

In an embodiment, the interaxial baseline is derived from at least one of the scaling factors of the photosensor of the light field camera, the pixel pitch of the photosensor, the pitch of the microlens array and the scaling factors of the micro lens array.

10 In an embodiment, wherein the extrinsic matrix of a sub-aperture image  $V_{k,l}$  is composed of the following rotation R and translation T matrices :

$$R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } T = \begin{pmatrix} k.b_x \\ l.b_y \\ 0 \end{pmatrix}$$

15 where  $b_x$  and  $b_y$  are the respective horizontal and vertical inter-axial baseline distances, and  $k$  and  $l$  are integers.

In an embodiment, the first light field data format is raw light field data, captured by a light field camera, the raw light field data comprising an array of micro-images, each micro image being projected by a respective micro lens of a micro-lens array and the second light field format is a focal stack, the focal stack being computed by splatting  
20 shifted micro images of the raw light field data, wherein the focalisation distance of the focused image is derived from at least one of the focal length of the main lens, the pixel pitch of the photosensor and the lens pitch of the microlens array.

In an embodiment, the focalisation distance is determined based on the expression:

25

$$z = \frac{w - p}{p} \times F$$

where  $w$  represents the shift defining a focused picture,  $F$  is the focal distance of the main lens and  $p = \frac{r_{\mu L}}{r_p}$  is the number of pixels of the photosensor corresponding to the diameter of a micro-lens,  $r_{\mu L}$  denoting the microlens pitch of the light field camera and  $r_p$  denoting the photosensor pitch of the light field camera.

5 In an embodiment, the derived focalisation distance is included in the metadata.

In an embodiment, the first light field data format is a matrix of sub-aperture images and the second light field data format is a focal stack, the focal stack being computed by splatting shifted sub-aperture images, wherein the focalisation vergence is derived from at least one of the focal lengths and the inter-axial baseline distances of the matrix of sub-aperture images.

In an embodiment, the focalisation vergence  $v$  is determined by :

$$v = \frac{1}{z} = \frac{d}{b_x \cdot f_x} \text{ or } v = \frac{d}{b_y \cdot f_y}$$

where  $b_x$  and  $b_y$  are the inter-axial baseline distances,  $f_x$  and  $f_y$  are the focal lengths of the matrix of sub-aperture images and  $d$  is the underlying disparity.

15 A further aspect of the invention provides an electronic device comprising memory, one or more processors; and one or more modules stored in the memory and configured for execution by the one or more processors, the one or more modules including instructions for processing or formatting light field data in accordance with embodiments of the invention.

A further aspect of the invention provides a computer readable memory including instructions, that when executed by a processor of a portable device, cause the device to perform operations comprising including instructions for processing or formatting light field data in accordance with embodiments of the invention.

25 Some processes implemented by elements of the invention may be computer implemented. Accordingly, such elements may take the form of an entirely hardware embodiment, an entirely software embodiment (including firmware, resident software, micro-code, etc.) or an embodiment combining software and hardware aspects that may

all generally be referred to herein as a "circuit", "module" or "system". Furthermore, such elements may take the form of a computer program product embodied in any tangible medium of expression having computer usable program code embodied in the medium.

Since elements of the present invention can be implemented in software, the present invention can be embodied as computer readable code for provision to a programmable apparatus on any suitable carrier medium. A tangible carrier medium may comprise a storage medium such as a floppy disk, a CD-ROM, a hard disk drive, a magnetic tape device or a solid state memory device and the like. A transient carrier medium may include a signal such as an electrical signal, an electronic signal, an optical signal, an acoustic signal, a magnetic signal or an electromagnetic signal, e.g. a microwave or RF signal.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Embodiments of the invention will now be described, by way of example only, and with reference to the following drawings in which:

FIG. 1A is a schematic diagram of a light field camera according to a first embodiment of the invention;

FIG. 1B is a schematic diagram of a light field camera according to a second embodiment of the invention;

FIG. 2A is a functional diagram of a light field camera according to an embodiment of the invention;

Figure 2B is a functional diagram of a light field data formator and light field data processor according to an embodiment of the invention;

FIG. 3 is an example of a 2D light field image formed on a photosensor array;

FIG. 4 is an example of steps of formatting and processing of light field data according to an embodiment of the invention;

FIG. 5 and FIG 6 are screen shots illustrating examples of data files in accordance with embodiments of the invention.

#### DETAILED DESCRIPTION

Light-field cameras are typically used to record a 4D light-field on a sensor



composed of an array of pixels. The 4D light-field data provides 2-dimensional spatial information and 2-dimensional angular information on the incoming light. Such light-Field cameras may be for instance: a plenoptic camera 100 comprising a main lens 101, an array of lenses 110 and a photo sensor 120 as illustrated in Figure 1A; or a multi-  
5 camera array comprising an array of lenses 210 and a photosensor 220 without a main lens, as illustrated in Figure 1B. A multi-array camera may be considered as a particular case of a plenoptic camera where the mains lens has an infinite focal.

Embodiments of the invention provide formatting of light field data for further processing applications such as format conversion, refocusing, viewpoint change and  
10 3D image generation.

Figure 1A is a schematic diagram of a light field camera for which one or more embodiments of the invention may be applied.

The light field camera 100 of Figure 1A is a plenoptic type light field camera comprising a micro lens array 110 and an image sensor 120 associated with the micro  
15 lens array 110. Further optics may be provided such as an aperture 102 for enabling a selectable amount of light to enter into the light field camera. The lens arrangement further includes a main (objective) lens 101 for focusing light towards the micro lens array 110.

The micro lens array 110 is composed of an array of microlenses 111, 112, 20 113...11n arranged, in a regular lattice structure. For example, the structure may be a rectangular lattice of rows and columns. The micro lens array may also be referred to as a lenslet array. For purposes of illustratio, the microlens array 110 is shown with a relatively small number of microlenses, but it will be appreciated that the number of microlenses may reach up to several thousand or even several million microlenses.

25 Optionally, spacers might be placed around each micro lens of the micro-lens array 110 between the micro-lens array 110 and the image sensor 120 in order to prevent light from one micro lens from overlapping with light of other microlenses of the microlens array 110 at the image sensor 120.

The image sensor comprises a photosensor array 120 composed of a number m  
30 of photo sensors 121, 122, 123, 124.....12m arranged in a lattice structure. For example the structure may be a rectangular lattice of rows and columns. Each photosensor 121, 122, 123, 124.....12m corresponds to a pixel or a group of pixels of the raw image of the scene captured by the photosensor array 120, each pixel covering a part (also referred

to as a point) of the scene. For purposes of illustration, the photosensor array 120 is illustrated as having a relatively small number of photosensors 121 to 121m. It will be appreciated, however, that, the number of photosensors is not limited to that illustrated in Figure 1A but may be extended to any number of photosensors, for example several  
5 thousand or several million photosensors. As an illustrative example, an image of 12.4 megapixels may be provided by an array of 4088x3040 pixels/photosensors. The image sensor may be for example a charge coupled device (CCD).

The microlenses 111, 112, 11n of the micro lens array 110 are arranged such that each microlens is optically associated with photosensors of the photo sensor array  
10 120. The photosensor array 120 is generally of finer pitch than the microlens array. Accordingly, each microlens is associated with a plurality of photosensors of the photosensor array 120. Optical association between a microlens and a group of photosensors signifies that the light rays passing through a given microlens reach at least one of the group of photosensors optically associated with the given microlens.

15 The interposition of the microlens array 110 between the main lens 101 and the photosensor 120 results in multiple images being formed on the photosensor array 120. Each microlens of the microlens array 110 projects a respective image, onto the associated photosensors of the photosensor array 120. Accordingly the raw image captured by the image sensor 120 is composed of an array of small images, as for  
20 example illustrated in Figure 3, typically referred to as micro-images. Each micro-image corresponds to a partial field of view from a respective different viewpoint. and corresponds to a micro-lens of the microlens array. Each pixel of photosensor 120 may be seen as corresponding to a 4D light field coordinate where two dimensions specify its spatial position on the sensor and two dimensions specify the angular or directional  
25 information of light that is incident upon that pixel according to the 2D position of the pixel within the microimage which it belongs to.

A color filter array (CFA) may in some cases be arranged on the microlens array 110 or on the photosensor array 120. The CFA typically arranges RGB (Red, Green and Blue) color filters on the photosensor or microlens array, the RGB arrangement taking,  
30 for example, the form of a Bayer filter mosaic. One color filter (red, green or blue filter) may be associated with an MLA according to a predetermined pattern, comprising 50 % green, 25 % red and 25 % blue in the example of a Bayer filter, such a pattern also being referred to as a RGBG, GRGB or RGGB pattern. It will be appreciated that, the arrangement of the color filters on the microlens array 110 or photosensor array 120 is



not limited to a RGGB pattern. In other embodiments, the predetermined pattern may be a RGBE pattern with one of the green filters modified to 'Emerald' (for a block of four color filters); a CYYM pattern with one 'Cyan' filter, two 'Yellow' filters and one 'Magenta' filter (for a block of four color filters); a CYGM pattern with one 'Cyan' filter, one 'Yellow' filter, one 'Green' filter and one 'Magenta' filter; a RGBW pattern with one 'Red' filter, one 'Green' filter, one 'Blue' filter and one 'White' filter, several arrangement being possible (for example arranged on a block of four color filters with 'White' for the upper left filter, 'Red' for the upper right filter, 'Blue' for the lower left filter and 'Green' for the lower right filter; or arranged on a block of 4x4 color filters with 'White', 'Blue', 'White', 'Green' for the first line, 'Blue', 'White', 'Green', 'White' for the second line below the first line, 'White', 'Green', 'White', 'Red' for the third line below the second line and 'Green', 'White', 'Red', 'White' for the fourth line below the third line).

The gap between the microlens array and the photosensor array may be composed of air, of an optical material having an index  $n$  (for example a glass layer) or of multiple layers comprising at least one layer air layer and at least one optical material layer. Using a glass layer to form the gap has the advantages of keeping the microlens array 110 at a constant distance from the photosensor array 120 uniformly across the photosensor array 120 and of reducing this distance when needed. If  $d$  is the distance between the output of the microlens array 110 and the photosensor array 120 along a longitudinal axis, having a layer composed of an optical material with an index  $n$  ( $n > 1$ , for example  $n = 1.5$ ) between the micro lens array 110 and the photosensor array 120 enables the distance to be set to  $d/n$  without modifying the distance  $d$ . By adapting/modifying the index of the optical material of the layer forming the gap, it is possible to adapt/modify a parameter representative of the distance between the microlens array 110 and the photosensor array 120 without modifying the distance  $d$ .

Figure 1B is a schematic diagram of a light field camera according to a second embodiment of the invention.

The light field camera 200 according to the second embodiment of the invention is a multi camera array type light field camera comprising a micro lens array 210 and an image sensor 220 associated with the micro lens array 210. In this embodiment a main lens 201 for focusing light towards the micro lens array 210 is not present. Elements such as the photosensor array 220 and micro lens array 210 operate in a similar manner to the corresponding elements of the plenoptic type camera of Figure 1A. The main



difference is that the main lens is not present in the embodiment of Figure 1B. Each micro-image corresponds to a full field of view from a respective different viewpoint.

Figure 2A is a block diagram of a light field camera device in accordance with an embodiment of the invention. The light field camera comprises an aperture/shutter 102, a main (objective) lens 101, a micro lens array 110 and a photosensor array 120 in accordance with the light field camera of Figure 1A. In some embodiments the light field camera includes a shutter release that is activated to capture a light-field image of a subject or scene. It will be appreciated that the functional features may also be applied to the light field camera of Figure 1B.

The photosensor array 120 provides light field image data which is acquired by LF Data acquisition module 140 for formatting by a light field data formatting module 150 and/or processing by light field data processor 155. Light field data may be stored, after acquisition and after processing, in memory 190 in a raw data format, or as sub aperture images or focal stacks.

In the illustrated example, the light field data formatting module 150 and the light field data processor 155 are disposed in or integrated into the light field camera 100. In other embodiments of the invention the light field data formatting module 150 and/or the light field data processor 155 may be provided in a separate component external to the light field capture camera. The separate component may be local or remote with respect to the light field image capture device. It will be appreciated that any suitable wired or wireless protocol may be used for transmitting light field image data to the formatting module 150 or light field data processor 155; for example the light field data processor may transfer captured light field image data and/ or other data via the Internet, a cellular data network, a WiFi network, a Bluetooth communication protocol, and/ or any other suitable means.

The light field data formatting module 150 is configured to provide a data format, in accordance with embodiments of the invention, aggregating metadata describing characteristics of at least one of a light field camera used to capture the light field data, a matrix of views derived from the light field data, and a focal stack derived from the light field data. The light field data formatting module 150 may be implemented in software, hardware or a combination thereof.

The light field data processor 155 configured to operate on raw light field image data received directly from the LF data acquisition module 140 for example to generate focal stacks or a matrix of views in accordance with embodiments of the

invention. Output data, such as, for example, still images, 2D video streams, and the like of the captured scene may be generated. The light field data processor may be implemented in software, hardware or a combination thereof.

In at least one embodiment, the light field camera 100 may also include a  
 5 user interface 160 for enabling a user to provide user input to control operation of camera 100 by controller 170. Control of the camera may include one or more of control of optical parameters of the camera such as shutter speed, or in the case of an adjustable light field camera, control of the relative distance between the microlens array and the photosensor, or the relative distance between the objective lens and the microlens array.  
 10 In some embodiments the relative distances between optical elements of the light field camera may be manually adjusted. Control of the camera may also include control of other light field data acquisition parameters, light field data formatting parameters or light field processing parameters of the camera. The user interface 160 may comprise any suitable user input device(s) such as a touchscreen, buttons, keyboard, pointing device,  
 15 and/ or the like. In this way, input received by the user interface can be used to control and/ or configure the LF data formatting module 150 for controlling the data formatting, the LF data processor 155 for controlling the processing of the acquired light field data and controller 170 for controlling the light field camera 100.

The light field camera includes a power source 180, such as one or more  
 20 replaceable or rechargeable batteries. The light field camera comprises memory 190 for storing captured light-field data and/or rendered final images or other data such as software for implementing methods of embodiments of the invention. The memory can include external and/ or internal memory. In at least one embodiment, the memory can be provided at a separate device and/ or location from camera 100. In one embodiment,  
 25 the memory includes a removable/swappable storage device such as a memory stick.

The light field camera may also include a display unit 165 (e.g., an LCD screen) for viewing scenes in front of the camera prior to capture and/or for viewing previously captured and/or rendered images. The screen 165 may also be used to display one or more menus or other information to the user. The light field camera may further include  
 30 one or more I/O interfaces 195, such as FireWire or Universal Serial Bus (USB) interfaces, or wired or wireless communication interfaces for data communication via the Internet, a cellular data network, a WiFi network, a Bluetooth communication protocol, and/ or any other suitable means. The I/O interface 195 may be used for transferring data, such as meta data formatted by LF data formatting module in accordance with



embodiments of the invention and light field data such as raw light field data or data processed by LF data processor 155, to and from external devices such as computer systems or display units, for rendering applications.

Figure 2B is a block diagram illustrating an embodiment of a potential  
5 implementation for of light field data formatting module and the light field data processor.

The circuit 300 includes memory 390, a memory controller 345 and processing  
circuitry 340 comprising one or more processing units (CPU(s)). The one or more  
processing units 340 are configured to run various software programs and/or sets of  
instructions stored in the memory 390 to perform various functions including light field  
10 data formatting and light field data processing. Software components stored in the  
memory include a data formatting module (or set of instructions) 350 for providing a data  
format aggregating light field metadata in accordance with embodiments of the invention  
and a light field data processing module (or set of instructions) 355 for processing light  
field data in accordance with embodiments of the invention, including a format  
15 conversion module 3555 (or set of instructions) for converting one format of light field  
data to another different format of light field data in accordance with embodiments of the  
invention. Other modules may be included in the memory for applications of the light field  
camera device such as an operating system module 351 for controlling general system  
tasks (e.g. power management, memory management) and for facilitating  
20 communication between the various hardware and software components of the device  
300, and an interface module 352 for controlling and managing communication with other  
devices via I/O interface ports.

Figure 3 illustrates an example of a 2D image formed on the photosensor array  
25 120 of Figure 1A or the photosensor array 220 of Figure 1B. The image is composed of  
an array of micro images  $M_i$ , each micro image being produced by the respective micro  
lens  $(i, j)$  of the microlens array 110, 210. The micro images are arranged in the array in  
a rectangular lattice structure defined by axes  $i$  and  $j$ . A micro lens image is referenced  
by the respective micro lens coordinates  $(i, j)$ . A pixel  $PI$  of the photosensor 120, 220 is  
30 referenced by its spatial coordinates  $(x, y)$ . 4D data associated with a given pixel may  
be referenced as  $(x, y, i, j)$ .



Embodiments of the invention provide a data format that aggregates metadata describing for raw light field data content, for matrices of views, and for focal stacks. In some embodiments the metadata may be used for processing light field data, for example for converting the light field data content from one format to another, e.g. from  
5 raw light field data into matrix of views or focal stacks; or from matrix of views into focal stacks; or for common light field applications such as viewpoint change and refocusing.

Figure 4 outlines steps of a method of processing and formatting light field data in accordance with one or more embodiments of the invention. In step S401 light field data is acquired by data acquisition module 140 from a plenoptic camera such as the  
10 camera shown in Figure 1A or a multi camera array such as the camera shown in Figure 1B. Lightfield data processor 155 receives the light field data in raw format. In step S402 the raw light field data is converted into a matrix of views (sub-aperture images) by a process of de-mosaicking and re-gridding. In step S403 the raw light field data is converted into a focal stack (set of images at different focal planes) by a process of  
15 splatting (accumulating shifted microimages). In step S404 the matrix of views is converted into a focal stack by a process of splatting (accumulating shifted sub-aperture images).

In step S410 metadata for the acquired light field data is provided and formatted. The metadata includes camera metadata defining technical characteristics of the  
20 camera, view matrix metadata defining the matrices of views and focal stack metadata defining the focal stacks; View point changes may be performed in step S420 using the matrix of views and light field meta data defining the matrix of views. Refocusing may be performed in step S430 using the focal stack and metadata defining the focal stack. A 2D image is generated from the micro-images by shifting the micro images by a given  
25 disparity and adding the images to form a single 2D image. In order to refocus the image at a given distance  $z$ , the micro-images are shifted by a disparity associated with the distance  $z$ . In order to perform a change of viewpoint only a part of the micro-images are added (e.g. left/right portions for left/right viewpoints)

In embodiments of the invention formatting syntax for metadata characterising the light field data content is provided. The metadata characterising the light field data content includes:

**Camera characteristic data:** characterising the sensor, the main lens and the  $\mu$ -lens array of the light field camera capturing the light field data and in particular the positions of the centres of the  $\mu$ -lens; The metadata may include data representative of sensor resolution, color sampling, color filter transmittance, dynamic range, color transform matrix, pixel pitch, lens pitch, focal length, vignetting, for example

**View characteristic data** - data characterising matrices of views such as metadata that characterizes the respective intrinsic & extrinsic parameters of each view/sub-aperture image; and color information of each view/sub-aperture image;

**Focal stack data** Data characterising focal stacks such as metadata that characterizes the focalization distance of each focused picture; for example an index map.

The formatting syntax may also include basic information needed to recover file location and file names for the management of video and sequence of frames.

The metadata may be provided as digital data stored in a file or a set of files. In one embodiment of the invention the formatting syntax is presented as a markup language complying for example with XML rules. In what follows the formatting syntax is referred to as *light field markup language* (abbreviated to *lfml*). In the example that follows the formatting syntax for the metadata includes the definition of **sections** (enclosed with **tags**) and of the corresponding **attributes** (the metadata itself).

### **Root section and type attribute**

Table 1 illustrates an example of a root section and attribute of a light field metadata file. The body of an \*.lfml file is enclosed in a **<LightFieldContent>** section. It contains an attribute specifying the type of content, denoted in the example as

**LFtype**="RAW" indicates raw plenoptic picture, indicates "**MATRIX\_OF\_VIEWS**" for arrays of views, and "**FOCAL\_STACK**" for focal stacks.

5

```
<LightFieldContent
lfml:LFtype="RAW"
>
[...]
</LightFieldContent>
```

**Table 1: Root section and attribute of an exemplary \*.lfml file**

10

**Folder and file attributes**

15

File naming attributes can be found at the root level if every file is stored and named consistently. File naming attributes can also be found in lower-level sections if some files do not share the same path or prefix or numbering.

The **path** attribute is a string specifying the folder path.

The **filePrefix** attribute is a string specifying the prefix of filenames.

20

The **storageMode** attribute is a two-part string separated by a colon that describes the file storage mode, and the file format used. It may either of the following: "**FILE\_SEQUENCE:PNG**" ; "**FILE\_SEQUENCE:PGM**" ; "**FILE\_SEQUENCE:RAW**" ; "**VIDEO\_FILE:YUV**"; "**VIDEO\_FILE:AVI**".

25

The **firstFrameIdx** and **lastFrameIdx** attributes are indices of the first and the last frame of the light field sequence respectively. The **digitSuffix** attribute is a string, for example of hash symbols, specifying the number of digits dedicated to frame numbering in the filename suffix. In the case of still-imaging, these three fields may be missing or empty (see Table 2 and Figure 5). In the case of sequence of images, filenames are suffixed with the index of the frames (see Figure 3 and Figure 6).

30

```
<LightFieldContent
lfml:LFtype="RAW"
lfml:path=".\\Lytro\\Gallery"
lfml:filePrefix="Snake"
```



5

```
lfml:storageMode="FILE_SEQUENCE:RAW"
lfml:firstFrameIdx=""
lfml:lastFrameIdx=""
lfml:digitSuffix=""
>
[...]
```

```
</LightFieldContent>
```

Table 2: File naming attributes for a raw light field still image

10

```
<LightFieldContent
  lfml:LfType="MATRIX_OF_VIEWS"
  lfml:path=".\\CGI\\Blender"
  lfml:filePrefix="outdoor"
  lfml:storageMode="FILE_SEQUENCE:PNG"
  lfml:firstFrameIdx="0"
  lfml:lastFrameIdx="249"
  lfml:digitSuffix="####"
>
[...]
```

15

```
</LightFieldContent>
```

20

Table 3: File naming attributes for light field video (here a CGI multi-view video sequence)

25

It may be noted that in the case of matrices of views and focal stacks, files may be fetched by default in folders named for instance respectively \view# and \focus# (see for example Figure 6). Filenames may also be built by default following the prefix and suffix rules predefined. However the attributes path, folder, filename and suffix can be used in subsections in order to specify alternative locations and names.

Coordinate system attributes

30

Dealing with light field imaging requires coordinate system conventions. The coordinate system may be implicit with default values being pre-determined, or explicitly stated in the description file with the attributes CSx and CSy, which can respectively have the following values: “leftward”/“rightward” and “upward”/“downward”.

Coordinate systems attributes can be located at any level of the \*.lfml file, up to the root level provided that all the coordinate systems used are consistent with each other.

### Raw plenoptic pictures

In the case of raw plenoptic pictures (**LFtype**="RAW"), a lightfield metadata file(\*.lfml  
5 file) present a section **<RawPlenopticPictures>**, that may include one or more of the following sections, depending on the completeness of the calibration: **<Sensor>**, **<MainLens>** and **<μLensArray>** (see Table 4).

The **<Sensor>** section may contain one or more of the following attributes describing technical features of the light field camera. Such attributes may be provided by the  
10 manufacturer:

The **raysResolution** attribute specifies the resolution of the sensor ("**WidthxHeight**", e.g. "**4008x2672**" for the Raytrix R11 camera, "**3280x3280**" for the first Lytro '*LightField*' camera, etc.).

The **colorFilter** attribute specifies the color sampling on the sensor (e.g.  
15 "**BayerGRBG**" or "**Foveonx3RGB**" or "**Monochrome**").

The **colorFilterTransmittance** attribute specifies the transmittance of the color filters of the sensor. It has one or three or four values, depending on the color filter. This attribute may be missing or empty.

The **dynamicRange** attribute specifies the range of the pixels. It has two values, the  
20 lower and upper bounds beyond which the signal can be considered as noise.

The **colorTransformMatrix** is a 3×3 matrix that specifies the non-linear relationship between the signal output by a pixel and the incoming light.

The **pixelPitch** attribute specifies the distance between two adjacent pixels, in meters.

The **scaleFactorX** and **scaleFactorY** attributes can be used to specify the shape ratio  
25 of rectangular (not square) pixels.

The **<MainLens>** section contains one or more of the following attributes, that may be provided by the manufacturer:

The **focalLength** attribute specifies the focal length of the main lens, in meters.

5 The **numericalAperture** attribute specifies the ratio between the pupil and the focal length. That attribute may be missing or empty.

The **whiteImage** attribute specifies a folder location from where to retrieve white images that correspond to the actual focal length. White images are used to remove the effect of vignetting. That attribute may be missing or empty.

10 The **vignetting** attribute specifies a model for main lens vignetting. That attribute may be missing or empty.

The **<μLensArray>** section contains the following attributes. Some may be provided by the manufacturer and some may be determined by further calibration.

The **pattern** attribute specifies the pattern of the μ-lens array ("**Square**", or "**HexagonalRowMajor**", or "**HexagonalColumnMajor**").

15 If applicable, a **colorArrangement** attribute specifies the pattern of color filters arranged on the μ-lens array. That attribute may be missing or empty.

The **μLensPitch** attribute specifies the distance between two adjacent μ-lens, in meters.

20 The **μLensArrayRotation** attribute specifies the rotation angle (in degrees) of the μ-lens array with respect to the sensor grid, in the coordinate system defined.

The **scaleFactorX** and **scaleFactorY** attributes can be used to specify the shape ratio of μ-lenses.

The **nbμCentersCalibrated** attribute specifies the numbers of centers of micro-images (μ-centers) (i.e. the number of centers of μ-image) calibrated. It can be equal to zero.



The **<μLensArray>** section contains the corresponding number of **<μCenter>** subsections (see Table 5), that contain the following attributes:

The **μCenterId** attribute is the index of the μ-center.

The **cx** and **cy** attributes respectively specify the x- and y-coordinates (in pixels) of the  
5 μ-center.

10

15

20

25

30

```
<LightFieldContent
  lfml:LfType="RAW"
  lfml:path=".\\Lytro\\Gallery"
  lfml:filePrefix="Snake"
  lfml:storageMode="FILE_SEQUENCE:RAW"
  lfml:firstFrameIdx=""
  lfml:lastFrameIdx=""
  lfml:digitSuffix=""
  lfml:CSx="rightward"
  lfml:CSy="downward"
>

<RawPlencpticPictures>
  <Sensor
    lfml:raysResolution="3280x3280"
    lfml:colorFilter="BayerGRBG"
    lfml:colorFilterTransmittance=""
    lfml:dynamicRange=""
    lfml:colorTransformMatrix=""
    lfml:pixelPitch="0.000001399999976158141876680929"
    lfml:scaleFactorX="1.0"
    lfml:scaleFactorY="1.0"
  ></Sensor>
  <MainLens
    lfml:focalLength="0.05131999969482421875"
    lfml:numericalAperture="1.9"
    lfml:whiteImage=""
    lfml:vignetting=""
```

5	<pre> &gt;&lt;/MainLens&gt;  &lt;μLensArray      lfml:pattern="HexagonalRowMajor"      lfml:μLensPitch="0.00001389861488342285067432158"      lfml:μLensArrayRotation="-0.002579216146841645240783691406"      lfml:scaleFactorX="1.0"      lfml:scaleFactorY="1.00024712085723876953125"      lfml:nbpCentersCalibrated="0"      &gt; </pre>
10	<pre>     &lt;/μLensArray&gt;  &lt;/RawPlenopticPictures&gt;  &lt;/LightFieldContent&gt; </pre>

Table 4: An example of \*.lfml file for a raw light field picture

15	<pre> LightFieldContent      lfml:LfType="RAW"      [...]      &gt; </pre>
20	<pre>         &lt;RawPlenopticPictures&gt;              [...]              &lt;μLensArray                  [...]                  lfml:nbpCentersCalibrated="108900"                  &gt; </pre>
25	<pre>                     [...]                      &lt;μCenter                          lfml:μCenterId="27224"                          lfml:cx="1633.087277"                          lfml:cy="1633.490847"                          &gt;&lt;/μCenter&gt; </pre>
30	<pre>                         [...]                      &lt;/μLensArray&gt;                  &lt;/RawPlenopticPictures&gt; </pre>



</LightFieldContent>

**Table 5: An example of  $\mu$ -center subsection**

### **Matrices of views**

In the case of matrices (or arrays) of views (LFtype="MATRIX\_OF\_VIEWS"), \*.lfrml  
 5 files present a section <MatrixOfViews>, that encloses several <SubApertureImage>  
 subsections. The syntax suits both a set of sub-aperture images reconstructed from  
 raw plenoptic material (see Table 6) and the content acquired from a camera array  
 (see Table 7).

The <MatrixOfViews> section contains the following attributes:

10 The nbSubApertureImages specifies the number of views/sub-aperture images. The  
 <MatrixOfViews> section contains the corresponding number of  
 <SubApertureImage> subsections.

The pixelResolution attribute specifies the resolution of views/sub-aperture images in  
 pixels ("WidthxHeight"), assuming that it is the same for all views. If not, this attribute  
 15 is filled in every subsequent <SubApertureImage> subsection. That attribute may be  
 missing or empty if available in the headers of corresponding picture files.

The <SubApertureImage> subsections contain the following attributes:

The subApertureId attribute is the index of the view / sub-aperture image.

20 The extrinsics attribute specifies the extrinsic matrix of the view/sub-aperture image in  
 the coordinate system defined.

The color attribute specifies the color(s) available in the view/sub-aperture image ("R",  
 "G", "B", "RGB", "L"). This attribute can be filled in the <MatrixOfViews> section if  
 shared by every view/sub-aperture image.

The **cx** and **cy** attributes respectively specify the x- and y-coordinates (in pixels) of the principal point. These attributes may be missing or empty. They can also be filled in the **<MatrixOfViews>** section if shared by every view/sub-aperture image.

The **fx** and **fy** attributes respectively specify the horizontal and vertical focal length (in pixels) of the view/sub-aperture view. These attributes may be missing or empty. They can also be filled in the **<MatrixOfViews>** section if shared by every view/sub-aperture image.

The **numericalAperture** attribute specifies the ratio between the pupil and the focal length, like in the **<MainLens>** section. That attribute may be missing or empty. It can also be filled in the **<MatrixOfViews>** section if shared by every view/sub-aperture image.

The **distortionCoefficients** attribute specifies a distortion model (for example, Bouguet's 5-parameter model of radial and tangential distortions). That attribute may be missing or empty. It can also be filled in the **<MatrixOfViews>** section if shared by every view/sub-aperture image.

```

20 <LightFieldContent
    lfml:LfType="MATRIX_OF_VIEWS"
    lfml:path=".\\PelicanImaging\\Gallery"
    lfml:filePrefix="GoldenGate"
    lfml:storageMode="FILE_SEQUENCE:PGM"
    lfml:firstFrameIdx=""
    lfml:lastFrameIdx=""
    lfml:digitSuffix=""
    lfml:CSx="rightward"
25 lfml:CSy="downward"
>
    <MatrixOfViews
        lfml:nbSubApertureImages="16"
        lfml:pixelResolution="1000x750"
30 >

```



5

10

15

20

25

30

```
<SubApertureImage
    lfml:subApImId="0"
    lfml:extrinsics="[[1, 0, 0, 0], [0, 1, 0, 0], [0, 0, 1, 0], [0, 0, 0, 1]]"
    lfml:color="G"
    lfml:cx="499"
    lfml:cy="374"
    lfml:fx=" 1175.0"
    lfml:fy="1175.0"
    lfml:numericalAperture=""
    lfml:distortionCoefficients=""
></SubApertureImage>
<SubApertureImage
    lfml:subApImId="1"
    lfml:extrinsics="[[1, 0, 0, 0.005], [0, 1, 0, 0], [0, 0, 1, 0], [0, 0, 0,
1]]"
    lfml:color="R"
    lfml:cx="500"
    lfml:cy="374"
    lfml:fx="1175.4"
    lfml:fy="1175.4"
    lfml:numericalAperture=""
    lfml:distortionCoefficients=""
></SubApertureImage>
<SubApertureImage
    lfml:subApImId="2"
    lfml:extrinsics="[[1, 0, 0, 0.01], [0, 1, 0, 0], [0, 0, 1, 0], [0, 0, 0,
1]]"
    lfml:color="G"
    lfml:cx="499"
    lfml:cy="375"
    lfml:fx="1175.9"
    lfml:fy="1175.9"
    lfml:numericalAperture=""
    lfml:distortionCoefficients=""
```

	<pre>&gt;&lt;/SubApertureImage&gt;</pre>
	<pre>&lt;SubApertureImage</pre>
	<pre>    lfml:subApImId="3"</pre>
5	<pre>    lfml:extrinsics="[[1, 0, 0, 0.015], [0, 1, 0, 0], [0, 0, 1, 0], [0, 0, 0,</pre>
	<pre>1]]"</pre>
	<pre>    lfml:color="B"</pre>
	<pre>    lfml:cx="500"</pre>
	<pre>    lfml:cy="375"</pre>
	<pre>    lfml:fx="1175.0"</pre>
10	<pre>    lfml:fy="1175.0"</pre>
	<pre>    lfml:numericalAperture=""</pre>
	<pre>    lfml:distortionCoefficients=""</pre>
	<pre>&gt;&lt;/SubApertureImage&gt;</pre>
	<pre>    [...]</pre>
15	<pre>&lt;/MatrixCfViews&gt;</pre>
	<pre>&lt;/LightFieldContent&gt;</pre>

**Table 6: An example of an excerpt of \*.lfml file describing the matrix of views reconstructed from a Pelican Imaging camera device**

	<pre>&lt;LightFieldContent</pre>
20	<pre>    lfml:LfType="MATRIX_OF_VIEWS"</pre>
	<pre>    lfml:path=".\\CGI\\Blender"</pre>
	<pre>    lfml:filePrefix="outdoor"</pre>
	<pre>    lfml:storageMode="FILE_SEQUENCE:PNG"</pre>
	<pre>    lfml:firstFrameIdx="0"</pre>
25	<pre>    lfml:lastFrameIdx="249"</pre>
	<pre>    lfml:digitSuffix="####"</pre>
	<pre>    lfml:CSx="rightward"</pre>
	<pre>    lfml:CSy="downward"</pre>
	<pre>&gt;</pre>
30	<pre>&lt;MatrixOfViews</pre>
	<pre>    lfml:nbSubApertureImages="121"</pre>
	<pre>    lfml:pixelResolution="1920x1080"</pre>
	<pre>    lfml:color="RGB"</pre>
	<pre>    lfml:cx="960"</pre>



5	>	<pre>lfml:cy="540" lfml:fx="3026.218853" lfml:fy="3026.218853" lfml:numericalAperture="0." &lt;SubApertureImage   lfml:subApImId="0"   lfml:extrinsics="[[1, 0, 0, 0], [0, 1, 0, 0], [0, 0, 1, 0], [0, 0, 0, 1]]"   lfml:distortionCoefficients="" &gt;&lt;/SubApertureImage&gt; &lt;SubApertureImage   lfml:subApImId="1"   lfml:extrinsics="[[1, 0, 0, 0.1], [0, 1, 0, 0], [0, 0, 1, 0], [0, 0, 0, 1]]"   lfml:distortionCoefficients="" &gt;&lt;/SubApertureImage&gt; [...]</pre>
10		
15		
20		<pre>&lt;SubApertureImage   lfml:subApImId="249"   lfml:extrinsics="[[1, 0, 0, 1.0], [0, 1, 0, 1.0], [0, 0, 1, 0], [0, 0, 0, 1]]"   lfml:distortionCoefficients="" &gt;&lt;/SubApertureImage&gt; &lt;/MatrixCfViews&gt; &lt;/LightFieldContent&gt;</pre>

25      **Table 7: An excerpt of \*.lfml file describing a computer-generated matrix of views.**

**Focal stacks**

In the case of focal stacks (LFtype="FOCAL\_STACK"), \*.lfml files present a section <FocalStack>, that encloses several <FocalStack> subsections and optionally a subsection <IndexMap> (see **Erreur ! Source du renvoi introuvable.** 8).

30      The <FocalStack> section contains the following attributes:

The **nbFocusedPictures** specifies the number of focused pictures in the stack. The **<FocalStack>** section contains the corresponding number of **<FocusedPicture>** subsections.

The **focusedPicIdxMap** attribute specifies whether an index map is available or not (**"AVAILABLE"** or **"UNAVAILABLE"**).

As in the case of matrices of views, the **pixelResolution** attribute specifies the resolution of the focused pictures in pixels (**"WidthxHeight"**), assuming that it is the same for everyone. If not, this attribute is filled in every subsequent **<FocusedPicture>** subsection. That attribute may be missing or empty if available in the headers of corresponding picture files.

The **<FocusedPicture>** subsections contain the following attributes:

The **focusPicId** attribute is the index of the focus picture.

The **distance** attribute specifies the distance of focalization of the focused picture, in meters. That attribute may be missing or empty.

The **vergence** attribute specifies the inverse of the distance of focalization of the focused picture (in  $\text{m}^{-1}$ ). That attribute may be missing or empty.

The **<IndexMap>** subsection contain the attributes needed to retrieve the corresponding file (**filename** or **suffix**, optionally **path** or **folder**).

```

20 <LightFieldContent
    lfml:LfType="FOCAL_STACK"
    lfml:path=".\\Lytro\\Gallery"
    lfml:filePrefix="Snake"
    lfml:storageMode="FILE_SEQUENCE:PNG"
25 lfml:firstFrameIdx=""
    lfml:lastFrameIdx=""
    lfml:digitSuffix=""

```

5

10

15

20

25

30

```
lfml:CSx="rightward"
lfml:CSy="downward"
>
<FocalStack
  lfml:nbFocusedPictures="8"
  lfml:focusPicIdxMap="AVAILABLE"
  lfml:pixelResolution="1000x1000"
  >
    <FocusedPicture
      lfml:focusPicId="0"
      lfml:distance=""
      lfml:vergence=""
    ></FccusedPicture>
    <FocusedPicture
      lfml:focusPicId="1"
      lfml:distance=""
      lfml:vergence=""
    ></FccusedPicture>
    <FocusedPicture
      lfml:focusPicId="2"
      lfml:distance=""
      lfml:vergence=""
    ></FccusedPicture>
    <FocusedPicture
      lfml:focusPicId="3"
      lfml:distance=""
      lfml:vergence=""
    ></FccusedPicture>
    <FocusedPicture
      lfml:focusPicId="4"
      lfml:distance=""
      lfml:vergence=""
    ></FccusedPicture>
    <FocusedPicture
```



	<code>lfml:focusPicId="5"</code>
	<code>lfml:distance =""</code>
	<code>lfml:vergence=""</code>
	<code>&gt;&lt;/FccusedPicture&gt;</code>
5	<code>&lt;FocusedPicture</code>
	<code>lfml:focusPicId="6"</code>
	<code>lfml:distance =""</code>
	<code>lfml:vergence=""</code>
	<code>&gt;&lt;/FccusedPicture&gt;</code>
10	<code>&lt;FocusedPicture</code>
	<code>lfml:focusPicId="7"</code>
	<code>lfml:distance =""</code>
	<code>lfml:vergence=""</code>
	<code>&gt;&lt;/FccusedPicture&gt;</code>
15	<code>&lt;IndexMap</code>
	<code>lfml:suffix="indexMap"</code>
	<code>&gt;&lt;/IndexMap&gt;</code>
	<code>&lt;/FocalStack&gt;</code>
	<code>&lt;/LightFieldContent&gt;</code>

20 **Table 8: An example of \*.lfml file for a focal stack**

The metadata may be used for processing light field data, for example converting the light field data content from one format to another, e.g. from raw light field data into matrix of views or focal stacks; or from matrix of views into focal stacks; and for common light field applications such as viewpoint change and refocusing. Examples of processing performed, in accordance with embodiments of the invention, on the LF field data by the light field data processor will now be described.

**Format conversion**

**From raw plenoptic content to matrices of views**

30 Arrays of views (micro images) can be reconstructed by de-mosaicking and re-gridding raw light field data acquired from the light field camera.

$(R_{i,j})$  denotes a set of micro-images in the raw light field picture. A sub-aperture image  $V_{k,l}$  is composed of collocated samples from each  $\mu$ -image:  $V_{k,l}[i,j] = R_{i,j}[k,l]$

The focal lengths (in pixels) of the corresponding array of views, i.e. the horizontal focal length **fx** and the vertical focal **fy** attributes of the metadata describing the matrix of views, are given by:

$$\begin{cases} f_x = \frac{F}{s_x \cdot r_p} \\ f_y = \frac{F}{s_y \cdot r_p} \end{cases}$$

where

- $F$  denotes the focal length of the main lens, in meters (i.e. the value of the attribute **focalLength** of the section **<MainLens>**);
- $s_x$  and  $s_y$  respectively denote horizontal and vertical scaling factors of the sensor (i.e. the values of the attributes **scaleFactorX** and **scaleFactorY** of the section **<Sensor>**);
- and  $r_p$  denotes the pixel pitch, in meters (i.e. the value of the attribute **pixelPitch** of the section **<Sensor>**).

And the underlying inter-axial baseline distances are given by:

$$\begin{cases} b_x = \frac{s_x \cdot s_y \cdot r_p^2}{\sigma_x \cdot r_{\mu L}} \\ b_y = \frac{s_x \cdot s_y \cdot r_p^2}{\sigma_y \cdot r_{\mu L}} \end{cases}$$

where

- $r_{\mu L}$  denotes the  $\mu$ -lens pitch, in meters (i.e. the value of the attribute  **$\mu$ LensPitch**);

- $\sigma_x$  and  $\sigma_y$  respectively denote horizontal and vertical scaling factors of the  $\mu$ -lens array (i.e. the values of the attributes **scaleFactorX** and **scaleFactorY** of the section **< $\mu$ LensArray>**).

The extrinsic matrix of the view  $V_{k,l}$  (that corresponds to the **extrinsics** attribute  
5 of the section **<SubApertureImage>**) is therefore made of the following rotation and translation matrices:

$$R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } T = \begin{pmatrix} k.b_x \\ l.b_y \\ 0 \end{pmatrix}$$

assuming that the optical center of view  $V_{0,0}$  is chosen as the origin of the coordinate system.

## 10 **From raw plenoptic content to focal stacks**

Focal stacks can be computed by accumulating ("*splatting*") shifted  $\mu$ -images.

$w$  denotes the shift defining a focused picture. Considering the contribution of a  $\mu$ -image  $R_{i,j}$ , for each of its pixel  $\begin{pmatrix} x \\ y \end{pmatrix}$ , new coordinates in the focal stack are given by:

$$\begin{cases} X = u.x - u.w.i \\ Y = u.y - u.w.j \end{cases}$$

15 where

- $\begin{pmatrix} x \\ y \end{pmatrix}$  is expressed in the whole raw image coordinate system;
- and  $u$  is a zooming factor.

The underlying focalization distance (in meters), i.e. the **distance** attribute of the section **<FocusedPicture>**, is given by:

20

$$z = \frac{w - p}{p} \times F$$



where  $p = \frac{r_{\mu L}}{r_p}$  is the number of pixels corresponding to the diameter of a  $\mu$ -lens.

The **vergence** attribute is then its inverse.

### From Matrices of views to focal stacks

Focal stacks can also be computed by accumulating ("*splatting*") shifted views,  
5 provided that they are registered.

The same formula can be used, considering this time the collection of sub-aperture images ( $V_{k,l}$ ):

$$\begin{cases} X = u.i - u.d.k \\ Y = u.j - u.d.l \end{cases}$$

where  $\begin{pmatrix} i \\ j \end{pmatrix}$  is expressed in the coordinate system of each view.

10 The underlying focalization vergence (in  $m^{-1}$ ), i.e. the **vergence** attribute of the section **<FocusedPicture>**, is given indifferently by:

$$v = \frac{1}{z} = \frac{d}{b_x \cdot f_x} \text{ or}$$

The corresponding **distance** attribute is the inverse of vergence  $v$ .

The proposed format of metadata describing light field data also suits common  
15 applications of light field data, such as viewpoint change and refocusing.

3D effects can be achieved by changing the viewpoint. This is obtained by view interpolation within the matrix of views.

Refocusing is a common application for focal stacks. The purpose involves stressing the depth of field, in which case the user picks one or several areas of the  
20 picture that he wants to be in focus (the remaining objects of the scene being as blurred as they are far from the focalized depth plane(s)). Another purpose consists in delivering

all-in-focus pictures. In both ways, index maps are locally retrieved in order to fetch samples in one focalized picture of the focal stack.

Although the present invention has been described hereinabove with reference to specific embodiments, the present invention is not limited to the specific  
5 embodiments, and modifications will be apparent to a skilled person in the art which lie within the scope of the present invention.

Many further modifications and variations will suggest themselves to those versed in the art upon making reference to the foregoing illustrative embodiments, which are given by way of example only and which are not intended to limit the scope of the  
10 invention, that being determined solely by the appended claims. In particular the different features from different embodiments may be interchanged, where appropriate.

**CLAIMS:**

1. A method of providing metadata for captured light field data, the method comprising providing a data format aggregating metadata describing characteristics of  
5 at least one of:  
a light field camera used to capture the light field data,  
a matrix of views derived from the light field data, and  
a focal stack derived from the light field data.
- 10 2. A method for rendering an image from light field data, the method comprising obtaining light field data metadata describing characteristics of at least one of:  
a light field camera used to capture the light field data,  
a matrix of views derived from the light field data, and  
a focal stack derived from the light field data; and  
15 rendering the light field data based on the information provided by the metadata
3. A method according to claim 1 or 2, wherein the metadata is aggregated in a light image file
- 20 4. A method according to any preceding claim, wherein the metadata describing characteristics of the light field camera used to capture the light field data includes data defining one or more of the following technical features of the light field camera:  
a photosensor of the light field camera;  
a main lens of the light field camera; and  
25 a micro lens array of the light field camera.
5. A method according to claim 4, wherein the metadata describing the photosensor of the light field camera includes data representative of one or more of the following parameters:  
30 spatial resolution of the photosensor;



colour sampling of a colour filter of the photosensor;  
transmittance of the colour filter of the photosensor  
dynamic range of the pixels of the photosensor;  
color transform matrix defining the non-linear relationship between the signal  
5 output by a pixel and the incoming light;  
the pitch between two adjacent pixels of the photosensor; and  
a scaling factor specifying the shape ratio of non-square pixels.

6. A method according to claim 4 or 5, wherein the metadata describing the main  
10 lens of the light field camera includes data representative of one or more of the following  
parameters:

focal length of the main lens  
the ratio between the pupil and the focal length; and  
the vignetting model of the main lens.

15

7. A method according to any one of claims 4 to 6, wherein the metadata describing  
the micro lens array of the light field camera includes data representative of one or more  
of the following parameters:

a pattern of the micro-lens array;  
20 a pattern of color filters arranged on the micro-lens array;  
the pitch between two adjacent micro-lenses of the micro-lens array;  
a scaling factor specifying the shape ratio of micro lens; and  
the number of calibrated microlens centres; and  
spatial coordinates of one or more microlens centres

25

8. A method according to any preceding claim, wherein the metadata is defined in  
markup language in compliance with XML.

9. A method according to any preceding claim wherein the metadata describing  
30 characteristics of an array of views derived from the light field data, includes data  
representative of one or more of the following parameters:

the number of views in the array;  
 the spatial resolution of a or each view;  
 colors available in a view;  
 the spatial coordinates of a reference point of a view;  
 5 the horizontal and vertical focal length of a view

10. A method according to any preceding claim wherein the metadata describing characteristics of a focal stack derived from the light field data, includes data representative of one or more of the following parameters:

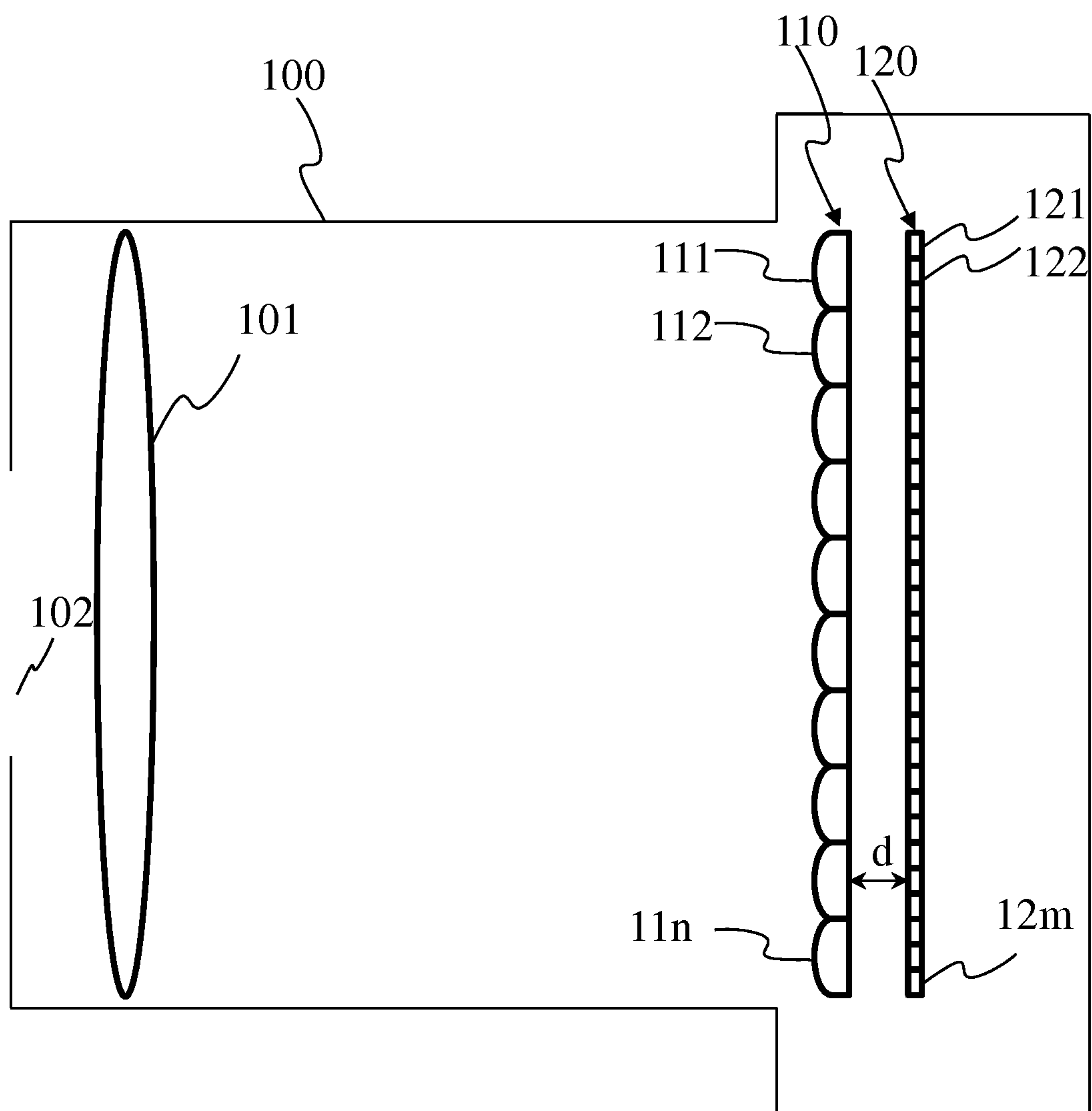
10 the number of focused pictures in the focal stack;  
 the spatial resolution of the focused pictures  
 the distance of focalization of a focused picture

11. A device for providing metadata for captured light field data,  
 15 the device comprising  
     a light field data acquisition module for acquiring light field data captured by a light field camera and  
     a data formatter for providing aggregating metadata describing characteristics of at least one of:  
 20 a light field camera used to capture the light field data,  
 a matrix of views derived from the light field data, and  
 a focal stack derived from the light field data.

12. A light field imaging device comprising:  
 25 an array of micro lenses arranged in a regular lattice structure;  
 a photosensor configured to capture light projected on the photosensor from the array of micro lenses, the photosensor comprising sets of pixels, each set of pixels being optically associated with a respective micro lens of the array of micro lenses; and  
 30 a device for providing metadata in accordance with claim 11.

13. A device for rendering an image from light field data using metadata describing characteristics of at least one of:
- a light field camera used to capture the light field data,
  - a matrix of views derived from the light field data, and
  - 5 a focal stack derived from the light field data, the device comprising a processor and memory containing a rendering application, the light field data and the metadata wherein the rendering application configures the processor to render the light field data based on information provided in the meta data.
- 10 14. A data package comprising light field data captured from a light field camera; and metadata describing characteristics of at least one of:
- a light field camera used to capture the light field data,
  - a matrix of views derived from the light field data, and
  - 15 a focal stack derived from the light field data.
- 15 15. A computer program product for a programmable apparatus, the computer program product comprising a sequence of instructions for implementing a method according to any one of claims 1 to 10 when loaded into and executed by the
- 20 programmable apparatus.



**1/6****FIG. 1A**

2/6

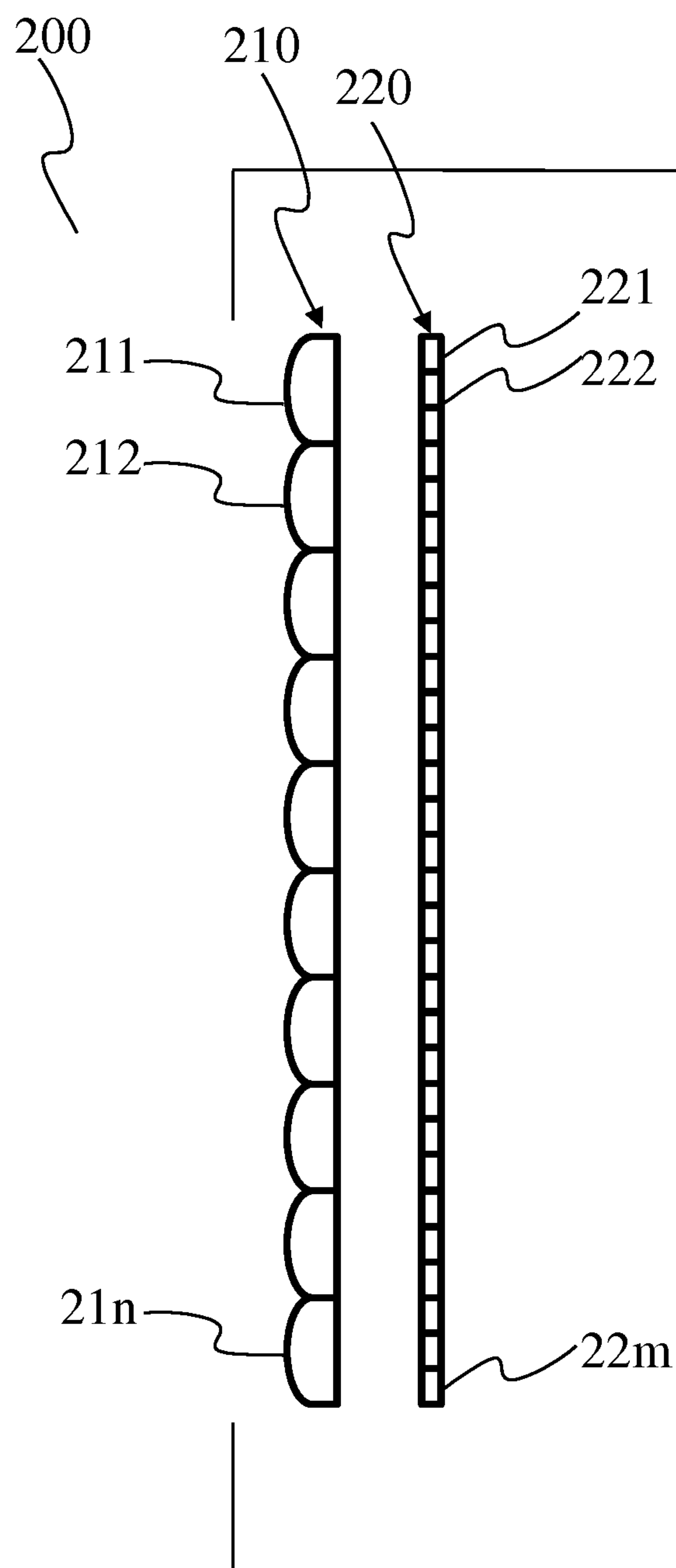


FIG. 1B

3/6

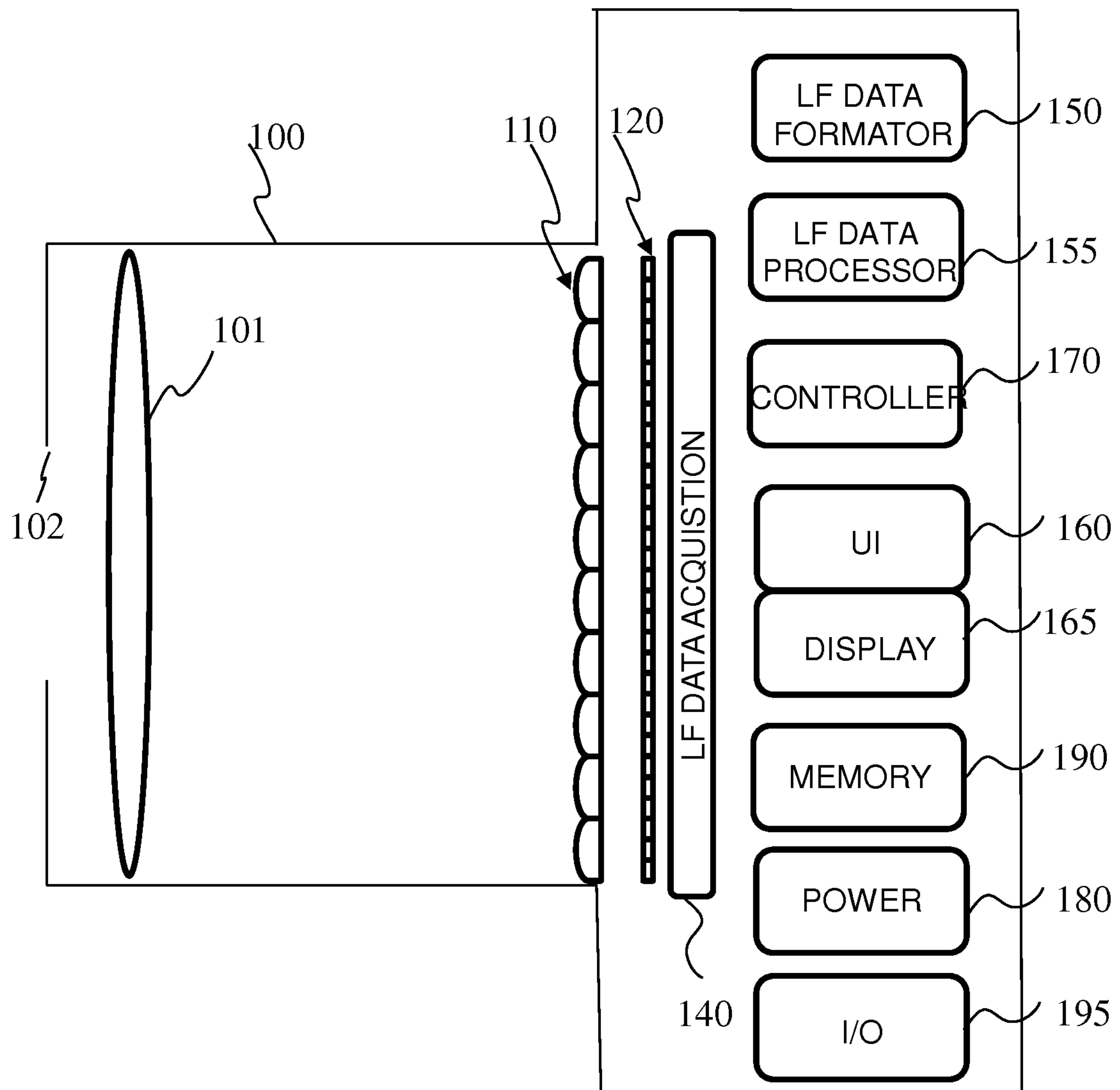
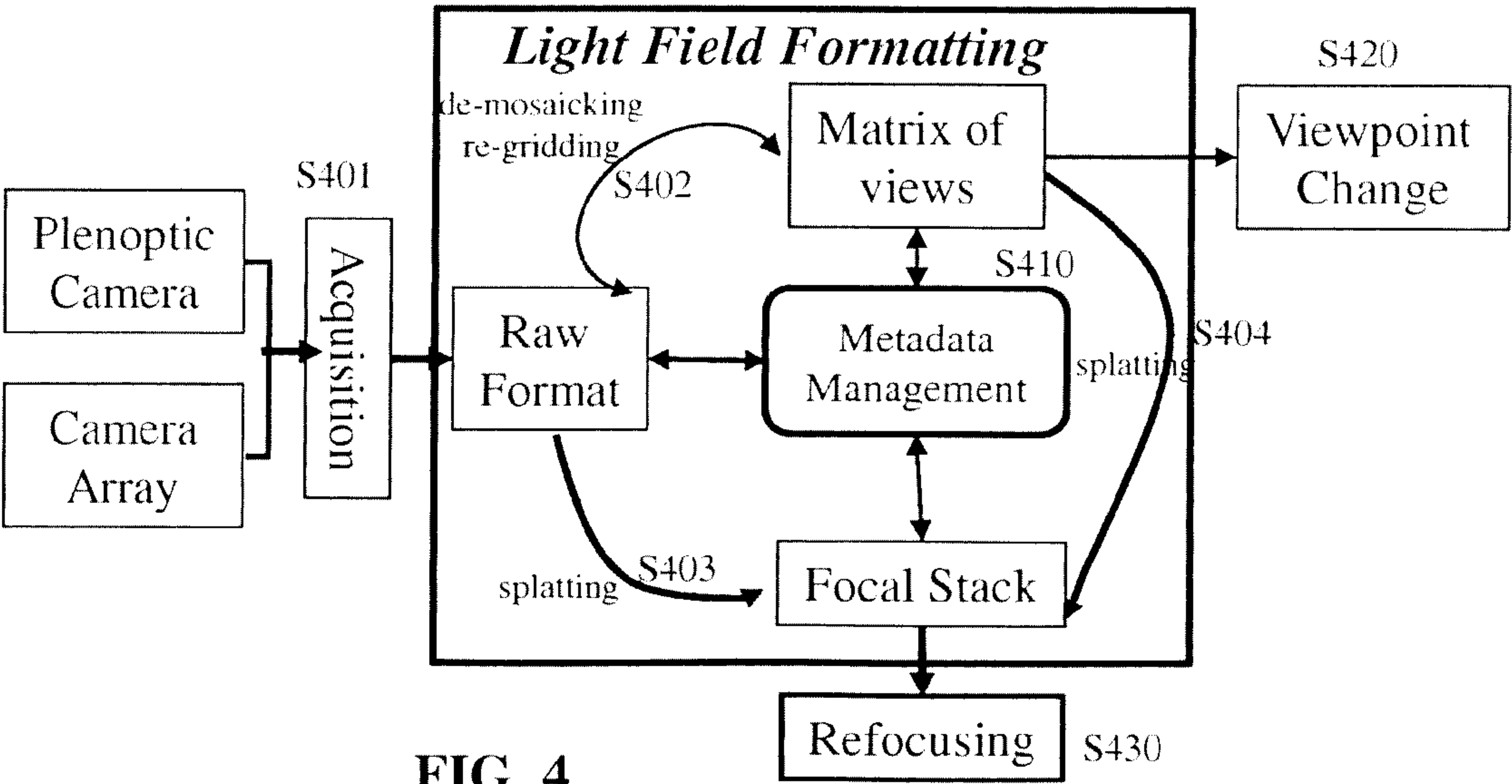
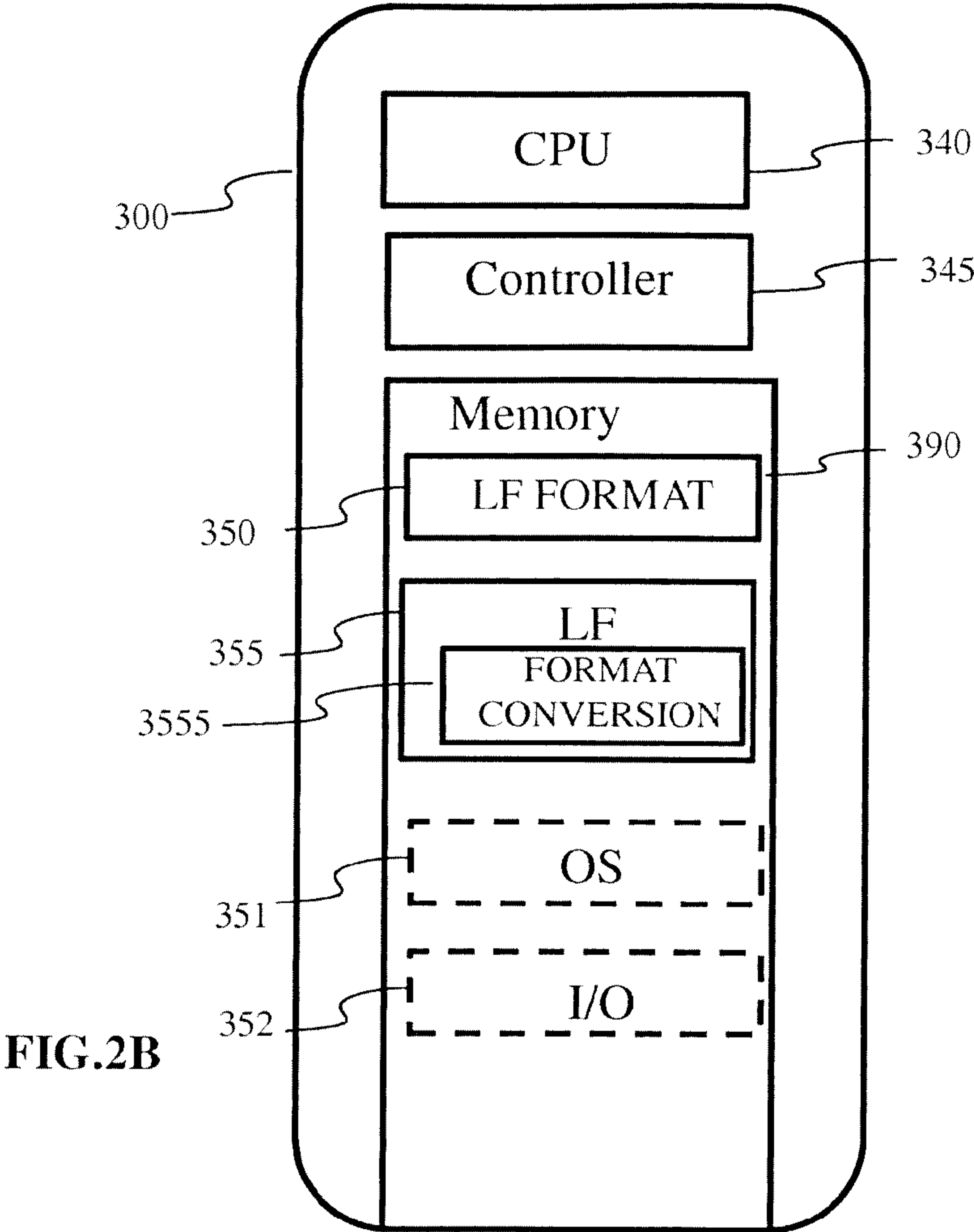
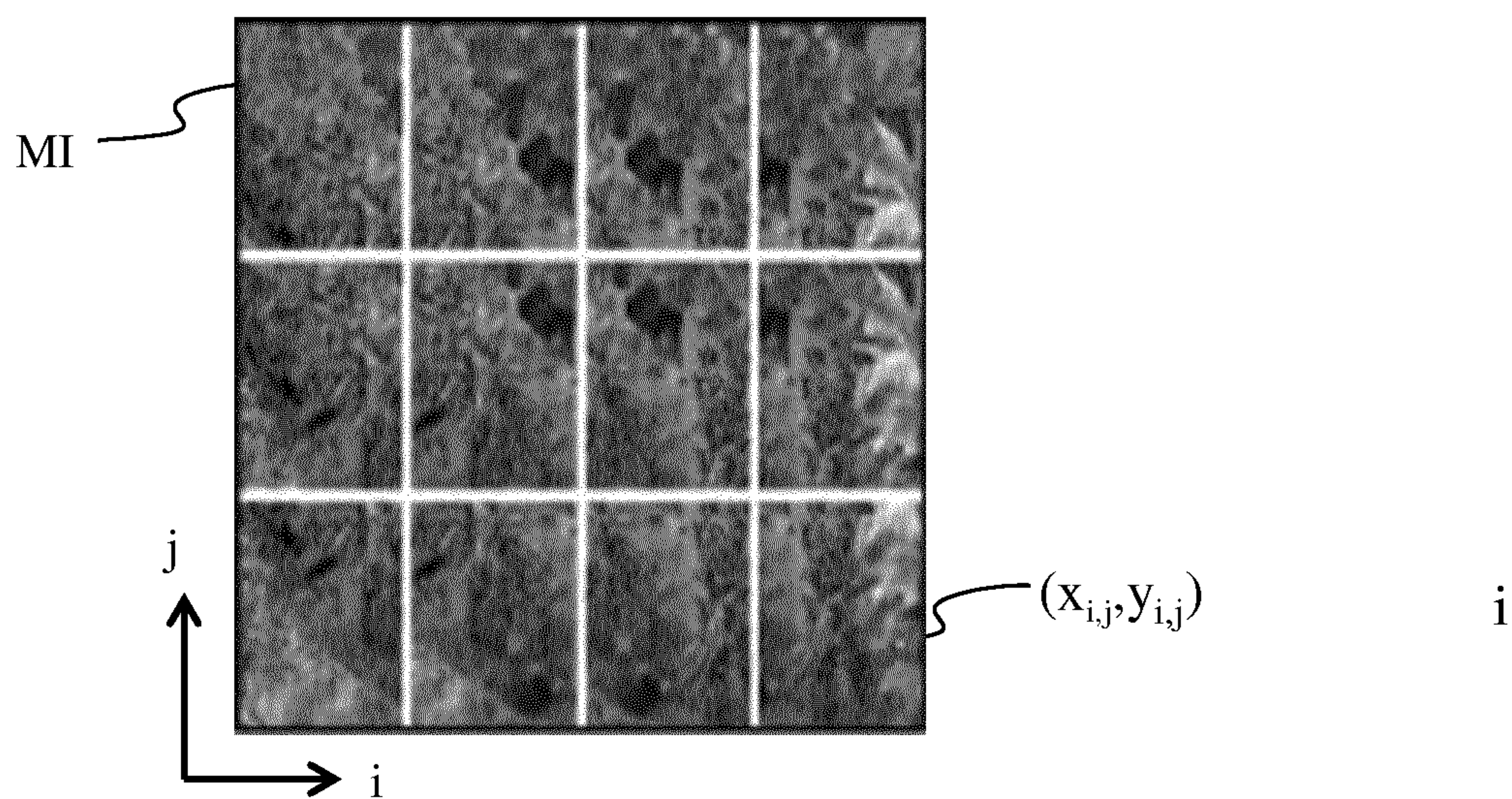


FIG. 2A

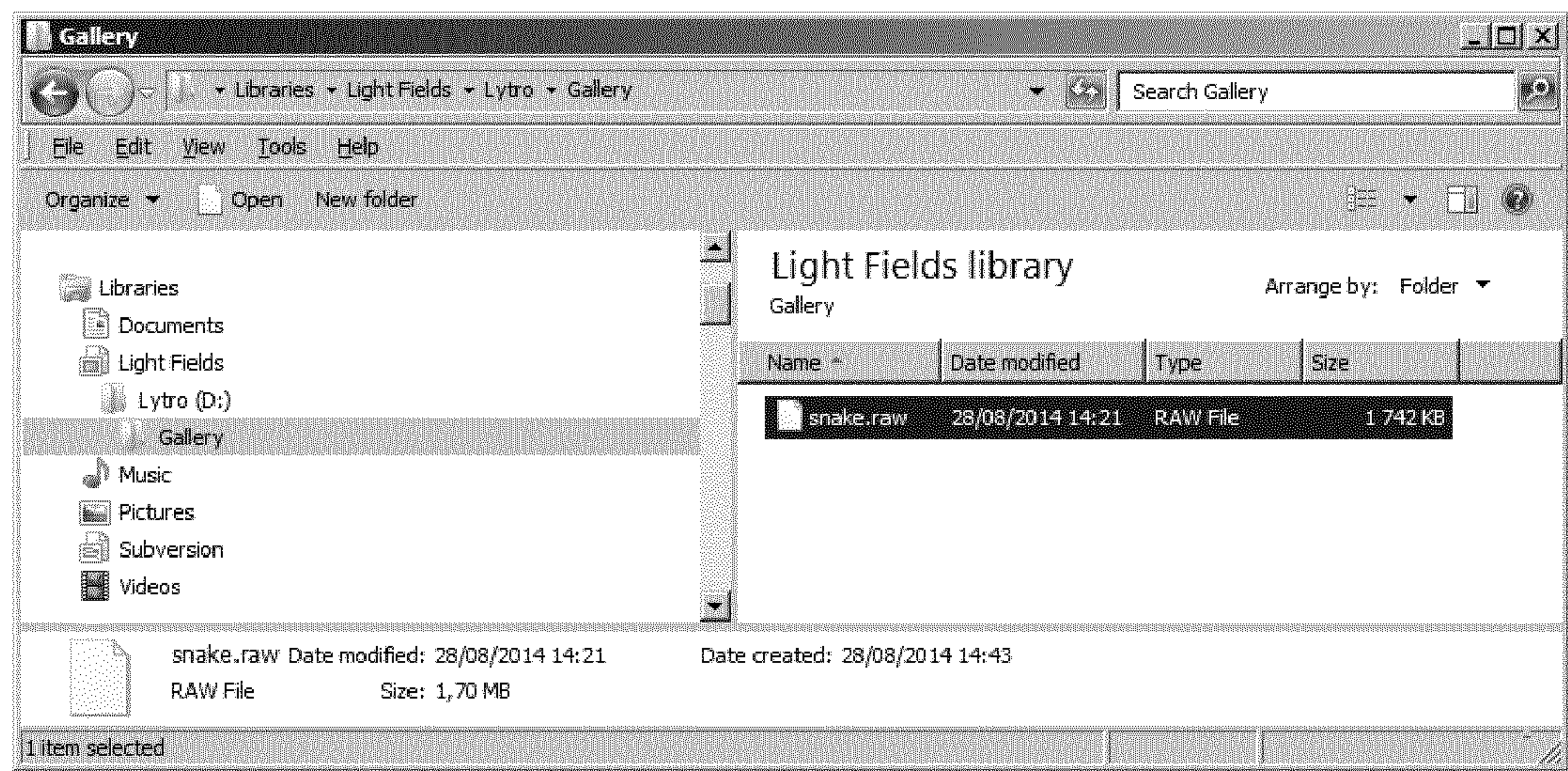


4/6



**5/6****FIG. 3**





i

FIG. 5

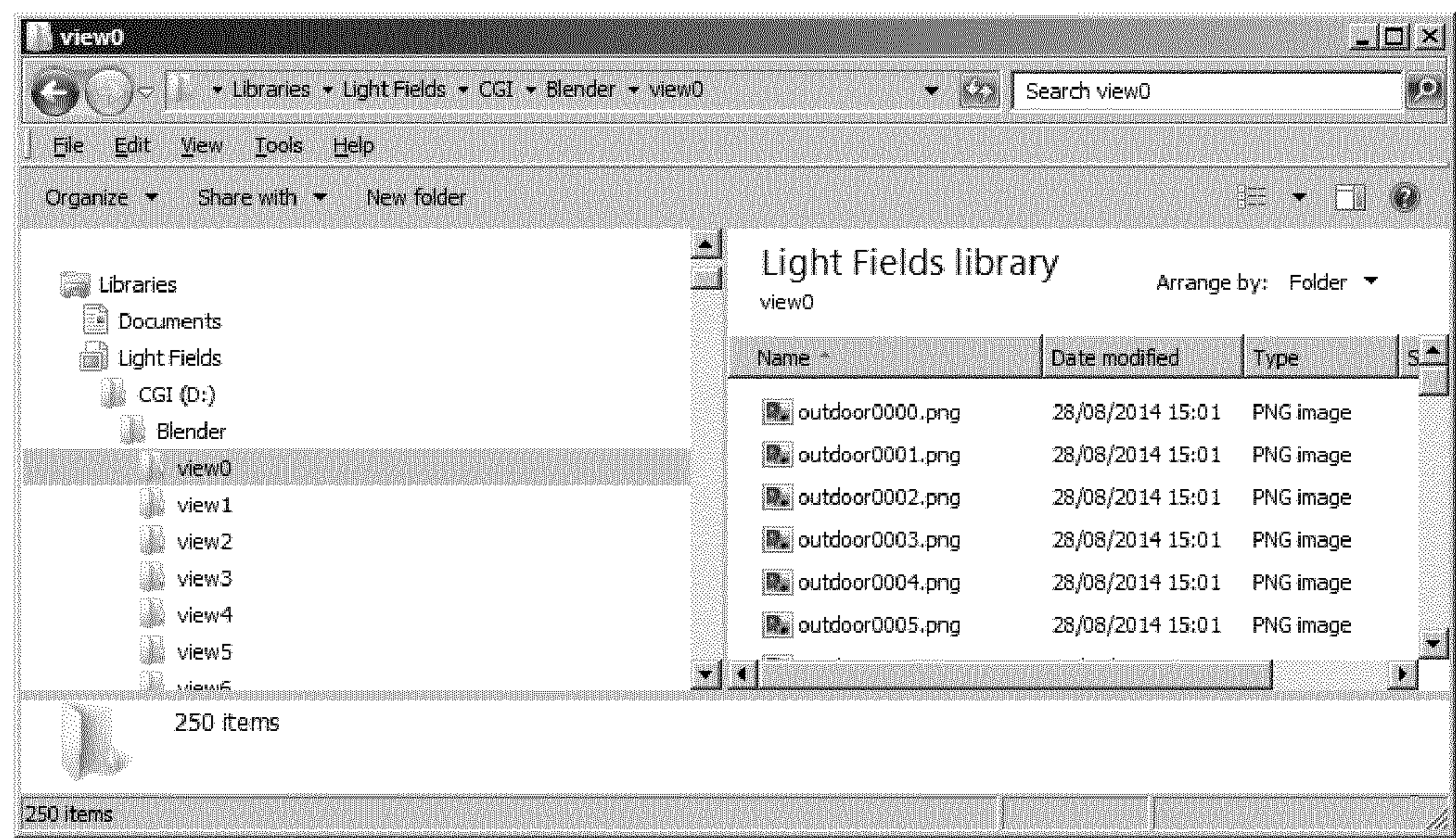
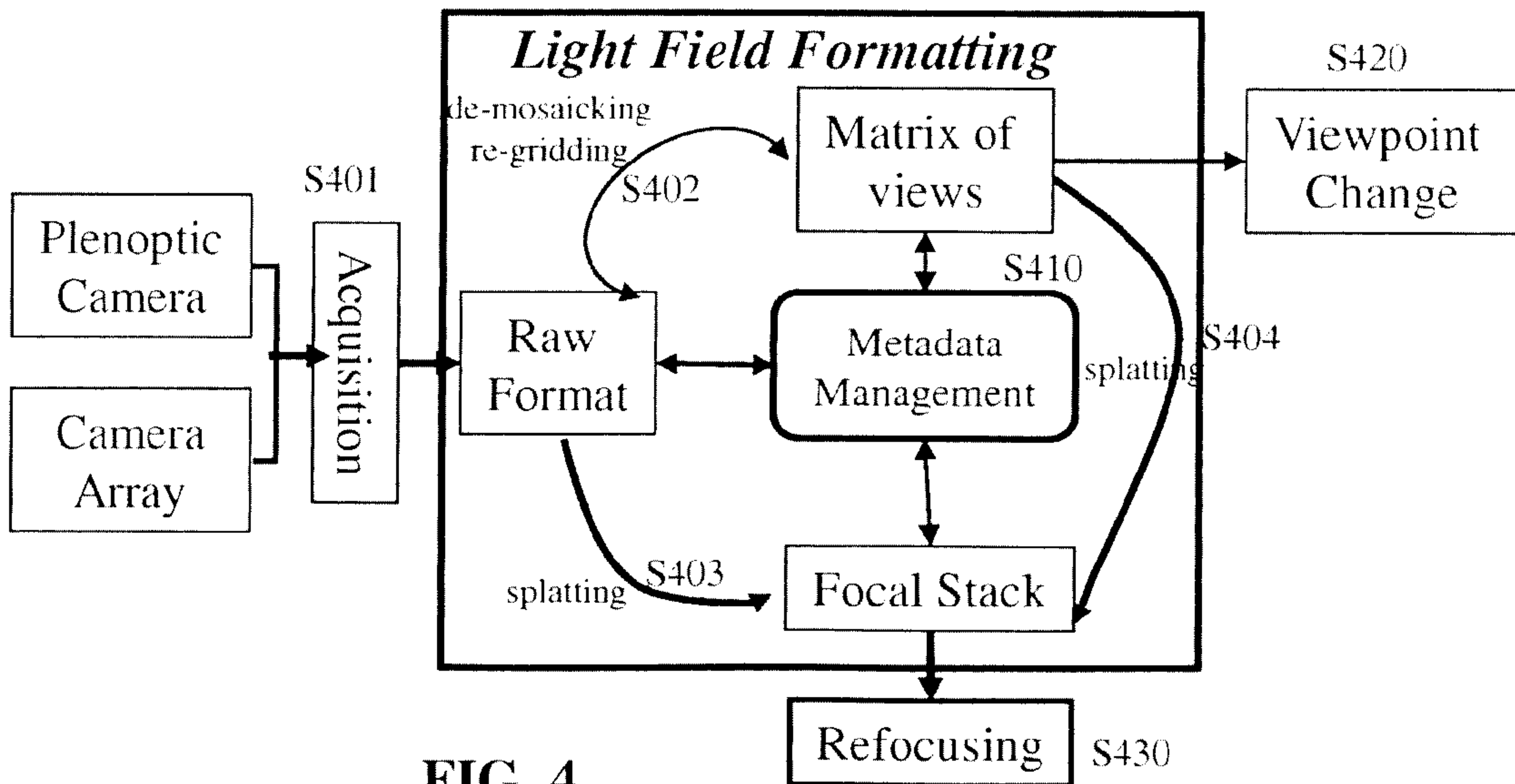


FIG. 6





**FIG. 4**