EXTENDED-DEPTH-OF-FIELD IMAGING

TECHNICAL FIELD

[0001] The present application relates to extended depth-of-field imaging technology, and particularly relates to an image acquisition apparatus for extended depth-of-field imaging, an extended depth-of-field imaging device, an image generation method for an extended depth-of-field image, and an image generation apparatus for an extended depth-of-field image.

BACKGROUND

[0002] An in-focus region included in a single depth-position image is limited by a depth-of-field range of an optical lens. Therefore, to obtain a larger in-focus region, an extended depth-of-field image may be generated by using a depth image sequence captured at a plurality of depth positions in the depth direction of the optical lens. The process of generating the extended depth-of-field image from the depth image sequence may be referred to as extended depth-of-field imaging. The extended depth-of-field image includes a depth map and an all-focus RGB image, such that the extended depth-of-field image can present three-dimensional object information of the target object in the depth image sequence.

SUMMARY

[0003] In view of this, embodiments of the present application provide an image acquisition apparatus for extended depth-of-field imaging, an extended depth-of-field imaging device, an image generation method for an extended depth-of-field image, and an image generation apparatus for an extended depth-of-field image, which help improve accuracy of three-dimensional object information presented by the extended depth-of-field image.

[0004] In one embodiment of the present application, an image acquisition apparatus for extended depth-of-field imaging includes: an adjustable light source module, including a plurality of illumination sources, wherein the plurality of illumination sources are configured to selectively illuminate the target object according to any one of at least two illumination modes, a light-reception state of the target object is configured to determine an imaging exposure condition under which the photosensitive element images the target object, and enabled states of the plurality of illumination sources under different illumination modes are different, so that the light-reception states of the target object under different illumination modes are different.

[0005] In some embodiments, the plurality of illumination sources include: a coaxial illumination source group, including a plurality of first illumination sources, wherein each first illumination source is configured to generate a first single-source light beam parallel to the depth direction, the first single-source light beams respectively generated by the plurality of first illumination sources cover different fan-shaped phase intervals of a circular illuminated region in a field-of-view region of the optical lens, and the circular illuminated region is coaxial with an optical axis of the optical lens; and an annular illumination source group, including a plurality of second illumination sources, wherein each second illumination source is configured to generate a second single-source light beam inclined relative to the depth direction, the second single-source light beams respectively generated by the plurality of second illumination sources cover different annular-sector phase intervals of an annular illuminated region in the field-of-view region, and the annular illuminated region coaxially surrounds the circular illuminated region.

[0006] In some embodiments, the light-reception state of the target object includes a light-reception orientation of the target object, wherein: the light-reception orientation is associated with a deployment position, in the coaxial illumination source group, of at least one enabled first illumination source in the coaxial illumination source group, and/or the light-reception orientation is associated with a deployment position, in the coaxial illumination source group, of at least one enabled second illumination source in the annular illumination source group.

[0007] In some embodiments, the light-reception state of the target object includes light-reception intensity, wherein: the light-reception intensity is associated with an intensity of the first single-source light beam generated by at least one enabled first illumination source in the coaxial illumination source group, and/or the light-reception intensity is associated with an intensity of the second light source beam generated by at least one enabled second illumination source in the annular illumination source group, and/or the light-reception intensity caused by the coaxial illumination source group is greater than that caused by the annular illumination source group.

[0008] In some embodiments, the imaging exposure condition under which the photosensitive element images the target object is further associated with an exposure duration of the photosensitive element.

[0009] In some embodiments, the illumination modes include at least two of a coaxial epi-illumination mode, a coaxial side-illumination mode, an annular epi-illumination mode, and an annular side-illumination mode, wherein: in the coaxial epi-illumination mode, all of the plurality of first illumination sources in the coaxial illumination source group are enabled, and

all of the plurality of second illumination sources in the annular illumination source group are disabled; in the coaxial side-illumination mode, a part of the plurality of first illumination sources in the coaxial illumination source group is enabled, another part of the plurality of first illumination sources in the coaxial illumination source group is disabled, and all of the plurality of second illumination sources in the annular illumination source group are disabled; in the annular epi-illumination mode, all of the plurality of second illumination sources in the annular illumination source group are enabled, and all of the plurality of first illumination sources in the coaxial illumination source group are disabled; in the annular side-illumination mode, a part of the plurality of second illumination sources in the annular illumination source group is enabled, another part of the plurality of second illumination sources in the annular illumination source group is disabled, and all of the plurality of first illumination sources in the coaxial illumination source group are disabled.

[0010] In some embodiments, the coaxial epi-illumination mode includes a normal-exposure coaxial epi-illumination mode and a low-exposure coaxial epi-illumination mode, wherein, in the normal-exposure coaxial epi-illumination mode, the first single-source light beams respectively generated by all the enabled first illumination sources have a first illumination intensity, and/or the exposure duration of the photosensitive element is configured as a first exposure duration; and in the low-exposure coaxial epi-illumination mode, the first single-source light beams respectively generated by all the enabled first illumination sources have a second illumination intensity lower than the first illumination intensity, and/or the exposure duration of the photosensitive element is configured as a second exposure duration shorter than the first exposure duration.

[0011] In some embodiments, the annular epi-illumination mode includes a normal-exposure annular epi-illumination mode and a low-exposure annular epi-illumination mode. In the normal-exposure annular epi-illumination mode, the second single-source light beams respectively generated by all the enabled second illumination sources have a third illumination intensity, and/or the exposure duration of the photosensitive element is configured as a third exposure duration. In the low-exposure annular epi-illumination mode, the second single-source light beams respectively generated by all the enabled second illumination sources have a fourth illumination intensity less than the third illumination intensity, and/or the exposure duration of the photosensitive element is configured as a fourth exposure duration shorter than the third exposure duration.

[0012] In some embodiments, the coaxial side-illumination mode includes at least two directional coaxial side-illumination modes having different side-illumination directions, and

the first illumination sources respectively enabled in any two of the directional coaxial side-illumination modes are not completely identical.

[0013] In some embodiments, the annular side-illumination mode includes at least two directional annular side-illumination modes having different side-illumination directions, and the second illumination sources respectively enabled in any two of the directional annular side-illumination modes are not completely identical.

[0014] In some embodiments, the adjustable light source module is configured to: when a magnification of the optical lens is greater than or equal to a preset magnification threshold, respectively illuminate the target object in the normal-exposure coaxial epi-illumination mode, the low-exposure coaxial epi-illumination mode, and at least two directional coaxial side-illumination modes; and/or, when the magnification of the optical lens is less than the preset magnification threshold, respectively illuminate the target object in the normal-exposure coaxial epi-illumination mode, the normal-exposure annular epi-illumination mode, the low-exposure annular epi-illumination mode, and at least two directional annular side-illumination modes.

[0015] In some embodiments, the movable imaging module includes a module housing, the photosensitive element is located in the module housing, and the optical lens is mounted at an end opening of the module housing facing the target object, wherein: the plurality of first illumination sources are fixedly mounted in the module housing, and the first single-source light beams are constrained by a beam splitter in the module housing to be parallel to the depth direction; and/or, the plurality of second illumination sources are fixedly mounted on an outer peripheral wall of the module housing near the end opening.

[0016] In some embodiments, the first illumination source includes a first light-emitting element group and a first light homogenizing plate covering the first light-emitting element group, wherein a beam cross section of the first single-source light beam generated by the first illumination source is shaped by the first light homogenizing plate, the first light homogenizing plate is fan-shaped, and the first light homogenizing plates of the respective first illumination sources are spliced into a circle; and/or, the second illumination source includes a second light-emitting element group and a second light homogenizing plate covering the second light-emitting element group, wherein a beam cross section of the second single-source light beam generated by the second illumination source is shaped by the second light homogenizing plate, the second light homogenizing plate is in an annular-sector shape, and the second light homogenizing plates of the respective second illumination sources are spliced into a ring.

[0017] In another embodiment of the present application, an extended depth-of-field imaging device includes: a movable imaging module, including an optical lens and a photosensitive element, wherein the photosensitive element is configured to image a target object within a lens field-of-view range of the optical lens, and a movement direction of the movable imaging module is parallel to a depth direction of the optical lens; an adjustable light source module, configured to adjust a light-reception state of the target object; and an image processing module, configured to generate an extended depth-of-field image of the target object based on at least two groups of depth image sequences, wherein the at least two groups of depth image sequences are obtained by imaging the target object by the photosensitive element under different imaging exposure conditions, wherein the imaging exposure condition under which the photosensitive element images the target object is associated with the light-reception state of the target object, and each group of the depth image sequences includes a plurality of images captured by the photosensitive element with the same imaging exposure condition when the movable imaging module is located at a plurality of depth positions in the depth direction, respectively.

[0018] In another embodiment of the present application, an image generation method for an extended depth-of-field image includes: acquiring at least two groups of depth image sequences obtained by imaging a target object by a photosensitive element of a movable imaging module under different imaging exposure conditions, wherein the imaging exposure conditions under which the photosensitive element images the target object are associated with light-reception states of the target object, and each group of the depth image sequences includes a plurality of images captured by the photosensitive element with the same imaging exposure condition when the movable imaging module is located at a plurality of depth positions in the depth direction, respectively; and generating the extended depth-of-field image of the target object based on the at least two groups of depth image sequences.

[0019] In another embodiment of the present application, an image generation apparatus for an extended depth-of-field image includes: an image acquisition module, configured to acquire at least two groups of depth image sequences obtained by imaging a target object by a photosensitive element of a movable imaging module, wherein an imaging exposure condition under which the photosensitive element images the target object is associated with a light-reception state of the target object, each group of the depth image sequences includes a plurality of images captured by the photosensitive element with the same imaging exposure condition when the movable imaging module is located at a plurality of depth positions in the depth direction, respectively, and imaging exposure conditions of any two groups of the depth image sequences are different from each other; and an extended depth-of-field processing module,

configured to generate the extended depth-of-field image of the target object based on at least two groups of the depth image sequences.

[0020] In another embodiment of the present application, a computer program product includes computer-executable instructions, wherein the computer-executable instructions, when executed by a processor, implement the image generation method according to the preceding embodiment.

[0021] In another embodiment of the present application, a non-transitory computer-readable storage medium stores instructions which, when executed by a processor, cause the processor to perform the image generation method according to the preceding embodiment.

[0022] Based on the above embodiments, at least two groups of depth image sequences acquired under different imaging exposure conditions can be provided for extended depth-of-field imaging. When structural features of a target object are prone to localized light-reception abnormality, because positions and degrees of the localized light-reception abnormality differ under different imaging exposure conditions, the probability that the same localized light-reception abnormality is present in all depth image sequences acquired under different imaging exposure conditions is extremely low, that is, local image defects present in at least two groups of depth image sequences acquired under different imaging exposure conditions are not completely identical. Therefore, by generating an extended depth-of-field image based on at least two groups of depth image sequences acquired under different imaging exposure conditions, influence of local image defects caused by localized light-reception abnormality on the extended depth-of-field image can be reduced through mutual supplementation of different object information presented by the target object in the at least two groups of depth image sequences, thereby helping improve the accuracy of the three-dimensional object information presented by the extended depth-of-field image.

BRIEF DESCRIPTION OF DRAWINGS

[0023] The following drawings are provided only for schematic illustration and explanation of the present application, and are not intended to limit the scope of the present application:

[0024] FIG. 1 is an exemplary structural schematic diagram of an image acquisition apparatus for extended depth-of-field imaging according to an embodiment of the present application.

[0025] FIG. 2 is an exemplary structural schematic diagram of an extended depth-of-field imaging device according to an embodiment of the present application.

[0026] FIG. 3 is a schematic diagram of an extended depth-of-field imaging principle of an extended depth-of-field imaging device according to an embodiment of the present application.

[0027] FIG. 4 is an exemplary flowchart of an image generation method for an extended depth-of-field image according to an embodiment of the present application.

[0028] FIG. 5 is a first optimized flowchart of an image generation method for an extended depth-of-field image according to an embodiment of the present application.

[0029] FIG. 6 is a second optimized flowchart of an image generation method for an extended depth-of-field image according to an embodiment of the present application.

[0030] FIG. 7 is an exemplary structural schematic diagram of an image generation apparatus for an extended depth-of-field image according to an embodiment of the present application.

DETAILED DESCRIPTION

[0031] In order to make the objectives, technical solutions, and advantages of the present application clearer, the present application is further described below in detail with reference to the accompanying drawings and embodiments.

[0032] Extended depth-of-field imaging is often used for observing a target object having structural features that abruptly change in a depth direction. However, structural features that abruptly change in the depth direction may cause local light-reception abnormality on the target object. For example, the target object may have a local light-reception blind region due to a recess of a structural feature in the depth direction or due to light shielding between structural features, or the target object may have a local overexposed region due to light reflection from a structural feature. Thus, each image in a depth image sequence may have local image defects due to the local light-reception abnormality of the target object, thereby causing the accuracy of three-dimensional object information presented by the extended depth-of-field image to be low. It can be seen that it is necessary to further improve the accuracy of the three-dimensional object information presented by the extended depth-of-field image.

[0033] FIG. 1 is a schematic diagram of an exemplary structure of an image acquisition apparatus for extended depth-of-field imaging according to an embodiment of the present application. Referring to FIG. 1, in an embodiment of the present application, the image acquisition apparatus for extended depth-of-field imaging may include: a movable imaging module 10 and an adjustable light source module 20.

[0034] In an embodiment of the present application, the movable imaging module 10 may include an optical lens 11 and a photosensitive element 12. A movement direction of the movable imaging module 10 may be parallel to a depth direction of the optical lens 11. For example, both the movement direction of the movable imaging module 10 and the depth direction of the optical lens 11 may be perpendicular to a carrying platform for placing a target object 50. The photosensitive element 12 is configured to image the target object 50 within a

lens field of view of the optical lens 11 to generate a depth image sequence. For example, the photosensitive element 12 may be a photosensitive imaging device such as a CCD (Charge Coupled Device) or a CMOS (Complementary Metal-Oxide Semiconductor).

[0035] Exemplarily, as illustrated in the drawings of the embodiments of the present application, the movable imaging module 10 may further include a module housing 13, wherein the photosensitive element 12 is located in the module housing 13, and the optical lens 11 is mounted at an end opening of the module housing 13 facing the target object 50 (i.e., facing the carrying platform 30).

[0036] In an embodiment of the present application, the depth image sequence generated by the photosensitive element 12 may include a plurality of images generated by imaging respectively when the movable imaging module 10 is located at a plurality of depth positions in the depth direction.

[0037] In an embodiment of the present application, the adjustable light source module 20 may include a plurality of illumination sources. The plurality of illumination sources are configured to selectively illuminate the target object 50 according to any one of at least two illumination modes. A light-reception state of the target object 50 is used for determining an imaging exposure condition under which the photosensitive element 12 images the target object 50, and enabled states of the plurality of illumination sources under different illumination modes are different, so that the light-reception state of the target object 50 differs under different illumination modes.

[0038] Exemplarily, as illustrated in the drawings of the embodiments of the present application, the plurality of illumination sources of the adjustable light source module 20 may include a coaxial illumination source group 21 and an annular illumination source group 22.

[0039] The coaxial illumination source group 21 may include a plurality of first illumination sources 210, each first illumination source 210 is configured to generate a first single-source light beam parallel to the depth direction, the first single-source light beams respectively generated by the plurality of first illumination sources 210 cover different fan-shaped phase intervals of a circular illuminated region in the field-of-view region of the optical lens 11, and the circular illuminated region is coaxial with an optical axis of the optical lens 11.

[0040] In this case, the coaxial illumination source group 21 (i.e., the plurality of first illumination sources 210) may be located on an axial inner side of the optical lens 11, and the first single-source light beams respectively generated by the plurality of first illumination sources 210 pass through the optical lens 11 and are emitted toward corresponding fan-shaped phase intervals of the circular illuminated region in the field-of-view region of the optical lens

11. For example, if the movable imaging module 10 includes the module housing 13 described above, the coaxial illumination source group 21 (i.e., the plurality of first illumination sources 210) may be fixedly mounted in the module housing 13, and the first single-source light beams respectively generated by the plurality of first illumination sources 210 may be constrained by a beam splitter 23 in the module housing 13 to be parallel to the movement direction of the movable imaging module 10, namely the depth direction of the optical lens 11.

[0041] The annular illumination source group 22 may include a plurality of second illumination sources 220, each second illumination source 220 is configured to generate a second single-source light beam inclined relative to the depth direction, the second single-source light beams respectively generated by the plurality of second illumination sources 220 cover different annular-sector phase intervals of an annular illuminated region in the field-of-view region of the optical lens 11, and the annular illuminated region of the annular illumination source group 22 may coaxially surround the circular illuminated region of the coaxial illumination source group 21.

[0042] In this case, the annular illumination source group 22 (i.e., the plurality of second illumination sources 220) may be located on a radial outer side of the optical lens 11, and the second single-source light beams respectively generated by the plurality of second illumination sources 220 may be obliquely emitted from the radial outer side of the optical lens 11 to corresponding annular-sector phase intervals of the annular illuminated region in the field-of-view region of the optical lens 11. For example, if the movable imaging module 10 includes the module housing 13 described above, the annular illumination source group 22 (i.e., the plurality of second illumination sources 220) may be fixedly mounted on an outer peripheral wall of the module housing 13 near the end opening where the optical lens 11 is located.

[0043] In an embodiment of the present application, the circular illuminated region corresponding to the coaxial illumination source group 21 in the field-of-view region of the optical lens 11 and the annular illuminated region corresponding to the annular illumination source group 22 in the field-of-view region of the optical lens 11 may complement each other, that is, the annular illuminated region of the annular illumination source group 22 may be coaxially nested around an outer periphery of the circular illuminated region of the coaxial illumination source group 21.

[0044] In an embodiment of the present application, the light-reception state of the target object 50 may include a light-reception orientation and/or light-reception intensity of the target object 50.

[0045] Exemplarily, in an embodiment of the present application, the light-reception orientation of the target object 50 is associated with a deployment position, in the coaxial illumination source group 21, of at least one enabled first illumination source 210 in the coaxial illumination source group 21, and/or with a deployment position, in the annular illumination source group 22, of at least one enabled second illumination source 220 in the annular illumination source group 22.

[0046] Exemplarily, in an embodiment of the present application, the light-reception intensity of the target object 50 is associated with an intensity of the first single-source light beam generated by at least one enabled first illumination source 210 in the coaxial illumination source group 21, and/or with an intensity of the second light source beam generated by at least one enabled second illumination source 220 in the annular illumination source group 22.

[0047] In an embodiment of the present application, each first illumination source 210 and/or each second illumination source 220 may not be limited to a single light-emitting element. For example, as illustrated in the drawings of the embodiments of the present application, the first illumination source 210 and/or the second illumination source 220 adopts a combined structure of a plurality of light-emitting elements and a light homogenizing plate, namely: the first illumination source 210 may include a first light-emitting element group 211 and a first light homogenizing plate 212 covering the first light-emitting element group 211, wherein the first light-emitting element group 211 may include a plurality of light-emitting elements such as LEDs (Light Emitting Diodes), and a beam cross section of the first single-source light beam generated by the first illumination source 210 may be shaped by the first light homogenizing plate 212; the second illumination source 220 may include a second light-emitting element group 221 and a second light homogenizing plate 222 covering the second light-emitting element group 221, wherein the second light-emitting element group 221 may include a plurality of light-emitting elements such as LEDs, and a beam cross section of the second single-source light beam generated by the second illumination source 220 may be shaped by the second light homogenizing plate 222.

[0048] In this case, the intensity of the first single-source light beam generated by each first illumination source 210 in the coaxial illumination source group 21 and/or the intensity of the second light source beam generated by each second illumination source 220 in the annular illumination source group 22 may be associated with a current magnitude of an energizing current of the light-emitting element when enabled.

[0049] Exemplarily, in an embodiment of the present application, compared with the LED in the second illumination source 220, the LED in the first illumination source 210 may have

greater light intensity and/or better light directivity. In this case, the light-reception intensity of the target object 50 caused by the coaxial illumination source group 21 may be greater than that caused by the annular illumination source group 22.

[0050] It can be understood that the above description of the adjustable light source module 20 is merely for reflecting that a division granularity of the plurality of illumination sources may not be limited to a single light-emitting element, and is also for reflecting that the first illumination source 210 and/or the second illumination source 220 adopts a combined structure of a plurality of light-emitting elements and a light homogenizing plate, so that the first single-source light beam and the second single-source light beam can produce a uniform illumination effect on the target object 50.

[0051] Exemplarily, in the drawing representation of the embodiments of the present application, it is further taken as an example that a beam cross-section shape of the first single-source light beam generated by each first illumination source 210 is fan-shaped and/or a beam cross-section shape of the second single-source light beam generated by each second illumination source 220 is annular-sector-shaped, namely: the first light homogenizing plate 212 of the first illumination source 210 may be fan-shaped, and the first light homogenizing plates 212 of the respective first illumination sources 210 may be combined into a circle, so that the fan-shaped phase intervals corresponding to the first single-source light beams respectively generated by the plurality of first illumination sources 210 can be spliced in a mutually complementary manner into the circular illuminated region corresponding to the coaxial illumination source group 21 in the field-of-view region of the optical lens 11; the second light homogenizing plate 222 of the second illumination source 220 may be annular-sector-shaped, and the second light homogenizing plates 222 of the respective second illumination sources 220 may be combined into a ring, so that the annular-sector phase intervals corresponding to the second single-source light beams respectively generated by the plurality of second illumination sources 220 can be spliced in a mutually complementary manner into the annular illuminated region corresponding to the annular illumination source group 22 in the field-of-view region of the optical lens 11.

[0052] In an embodiment of the present application, selective enabling of the plurality of illumination sources of the adjustable light source module 20 may include: selective enabling of the plurality of first illumination sources 210 and/or selective enabling of the plurality of second illumination sources 220. In this case, selective enabling of the plurality of first illumination sources 210 may be configured to respectively adjust illuminated states of a plurality of fan-shaped phase intervals of the circular illuminated region corresponding to the

coaxial illumination source group 21 in the field-of-view region of the optical lens 11, and/or selective enabling of the plurality of second illumination sources 22 may be configured to respectively adjust illuminated states of a plurality of annular-sector phase intervals of the annular illuminated region corresponding to the annular illumination source group 22 in the field-of-view region of the optical lens 11.

[0053] In an embodiment of the present application, compared with the annular illumination source group 22, the coaxial illumination source group 21 can be more advantageous for improving an observation effect on recessed structural features of the target object 50 in the depth direction, because light illuminating bottoms of the recessed structural features is easily blocked by surroundings, and a recessed structure can be easily observed only by illumination parallel to the depth direction. That is, the circular illuminated region of the coaxial illumination source group 21 has a stronger illumination capability for recesses of structural features of the target object 50 in the depth direction, so as to compensate for local light-reception blind regions on the target object 50 due to recesses of structural features in the depth direction and light shielding between structural features.

[0054] In an embodiment of the present application, selective enabling of the plurality of illumination sources of the adjustable light source module 20 can adjust the light-reception state of the target object 50. For example, the light-reception state of the target object 50 may include light-reception intensity and/or a light-reception orientation of the target object 50, and adjustment of the light-reception state of the target object may be used to cause a change in an imaging exposure condition under which the photosensitive element 12 images the target object 50. Thus, based on the change in the imaging exposure condition under which the target object 50 is imaged, the photosensitive element 12 may be configured to: provide at least two groups of depth image sequences Sq_1 to Sq_m acquired under different imaging exposure conditions for extended depth-of-field imaging, and each group of depth image sequence Sq_i includes: a plurality of images captured by the photosensitive element 12 with the same imaging exposure condition when the movable imaging module 10 is located at a plurality of depth positions in the depth direction, respectively. Here, m is a positive integer greater than or equal to 2, m represents the number of groups of the depth image sequences Sq_1 to Sq_m, and i is a positive integer greater than or equal to 1 and less than or equal to m.

[0055] As can be seen from the above, the image acquisition apparatus for extended depth-of-field imaging in an embodiment of the present application can provide at least two groups of depth image sequences Sq_1 to Sq_m acquired under different imaging exposure conditions for extended depth-of-field imaging. When structural features of the target object 50 are prone

to localized light-reception abnormality, because positions and degrees of the localized light-reception abnormality differ under different imaging exposure conditions, the probability that a same localized light-reception abnormality is present in all depth image sequences Sq_1 to Sq_m acquired under different imaging exposure conditions is extremely low, that is, localized image defects present in at least two groups of depth image sequences Sq_1 to Sq_m acquired under different imaging exposure conditions are not completely identical. Therefore, by generating an extended depth-of-field image based on at least two groups of depth image sequences Sq_1 to Sq_m acquired under different imaging exposure conditions, influence of localized image defects caused by localized light-reception abnormality on the extended depth-of-field image can be reduced through mutual supplementation of different object information presented by the target object in the at least two groups of depth image sequences Sq_1 to Sq_m, thereby helping improve the accuracy of the three-dimensional object information presented by the extended depth-of-field image.

[0056] In an embodiment of the present application, when the adjustable light source module 20 includes the coaxial illumination source group 21 and the annular illumination source group 22, the adjustable light source module 20 may be selectively enabled: the circular illuminated region of the coaxial illumination source group 21 may be in a global illuminating state when all first illumination sources 210 are enabled (also referred to as a coaxial epi-illumination state), or may be in a local side-illumination state when a part of the first illumination sources 210 is enabled (also referred to as a coaxial side-illumination state); the annular illuminated region of the annular illumination source group 22 may be in a global illuminating state when all second illumination sources 220 are enabled (also referred to as an annular epi-illumination state), or may be in a local side-illumination state when a part of the second illumination sources 220 is enabled (also referred to as an annular side-illumination state).

[0057] In an embodiment of the present application, illumination modes of the adjustable light source module 20 may include at least two of a coaxial epi-illumination mode, a coaxial side-illumination mode, an annular epi-illumination mode, and an annular side-illumination mode: in the coaxial epi-illumination mode, the plurality of first illumination sources 210 in the coaxial illumination source group 21 are all enabled, and the plurality of second illumination sources 220 in the annular illumination source group 21 are all disabled; in the coaxial side-illumination mode, one part of the plurality of first illumination sources 210 in the coaxial illumination source group 21 is enabled, the other part of the plurality of first illumination sources 210 in the coaxial illumination source group 21 is disabled, and the plurality of second illumination sources 220 in the annular illumination source group 22 are all

disabled; in the annular epi-illumination mode, the plurality of second illumination sources 220 in the annular illumination source group 22 are all enabled, and the plurality of first illumination sources 210 in the coaxial illumination source group 21 are all disabled; in the annular side-illumination mode, one part of the plurality of second illumination sources 220 in the annular illumination source group 22 is enabled, the other part of the plurality of second illumination sources 220 in the annular illumination source group 22 is disabled, and the plurality of first illumination sources 210 in the coaxial illumination source group 21 are all disabled.

[0058] In an embodiment of the present application, the imaging exposure condition under which the photosensitive element 12 images the target object 50 may further be associated with an exposure duration of the photosensitive element 12, that is, differences in imaging conditions of the at least two groups of depth image sequences Sq_1 to Sq_m may be caused by differences in the light-reception state of the target object 50 and/or in the exposure duration of the photosensitive element 12.

[0059] In an embodiment of the present application, the coaxial epi-illumination mode may include a normal-exposure coaxial epi-illumination mode and a low-exposure coaxial epi-illumination mode. In the normal-exposure coaxial epi-illumination mode, the first single-source light beams respectively generated by all the enabled first illumination sources 210 in the coaxial illumination source group 21 have a first illumination intensity, and/or the exposure duration of the photosensitive element 12 may be configured as a first exposure duration. In the low-exposure coaxial epi-illumination mode, the first single-source light beams respectively generated by all the enabled first illumination sources 210 in the coaxial illumination source group 21 have a second illumination intensity, and/or the exposure duration of the photosensitive element 12 may be configured as a second exposure duration, and the second illumination intensity is less than the first illumination intensity, and the second exposure duration is shorter than the first exposure duration. That is, an exposure-degree difference between the normal-exposure coaxial epi-illumination mode and the low-exposure coaxial epi-illumination mode may be determined by the illumination intensity of the first illumination sources 210 and/or the exposure duration of the photosensitive element 12.

[0060] In an embodiment of the present application, the annular epi-illumination mode may include a normal-exposure annular epi-illumination mode and a low-exposure annular epi-illumination mode. In the normal-exposure annular epi-illumination mode, the second single-source light beams respectively generated by all the enabled plurality of second illumination sources 220 in the annular illumination source group 22 have a third illumination intensity, and/or the exposure duration of the photosensitive element 12 may be configured as a third

exposure duration. In the low-exposure annular epi-illumination mode, the second single-source light beams respectively generated by all the enabled plurality of second illumination sources 220 in the annular illumination source group 22 have a fourth illumination intensity, and/or the exposure duration of the photosensitive element 12 may be configured as a fourth exposure duration. The fourth illumination intensity is less than the third illumination intensity, and the fourth exposure duration is shorter than the third exposure duration. That is, an exposure-degree difference between the normal-exposure annular epi-illumination mode and the low-exposure annular epi-illumination mode may be determined by the illumination intensity of the second illumination sources 220 and/or the exposure duration of the photosensitive element 12.

[0061] In an embodiment of the present application, the illumination intensity of an illumination source may be determined by power of the illumination source. Therefore, if the illumination intensity of the illumination source is used to determine an exposure-degree difference, the exposure degree may be adjusted by controlling the power of the illumination source. Exemplarily: the first illumination source 210 has a first maximum power (for example, 11 W), wherein power of the first illumination source 210 for generating the first illumination intensity in the normal-exposure coaxial epi-illumination mode may be set to a first rated power value between one third and one half of the first maximum power (for example, 4±0.5 W), and power of the first illumination source 210 for generating the second illumination intensity in the low-exposure coaxial epi-illumination mode may be set to between one third and two thirds of the first rated power value (for example, one half of the first rated power value, for example 2 W); the second illumination source 220 has a second maximum power less than the first maximum power value (for example, 9 W), wherein power of the second illumination source 220 for generating the third illumination intensity in the normal-exposure annular epi-illumination mode may be set to a second rated power value between one third and one half of the second maximum power (for example, 3±0.5 W), and power of the second illumination source 220 for generating the fourth illumination intensity in the low-exposure annular epi-illumination mode may be set to between one third and two thirds of the second rated power value (for example, one half of the second rated power value, for example 1.5 W).

[0062] In an embodiment of the present application, if the exposure duration of the photosensitive element 12 is used to determine an exposure-degree difference, exemplarily, the first exposure duration may be any exposure duration (for example, 90 ms) set within an available exposure duration range (for example, 0.5 ms to 100 ms) of the photosensitive

element 12, and the second exposure duration may be any configurable duration less than or equal to two thirds of the first exposure duration (for example, one half of the first exposure duration, for example 45 ms).

[0063] In an embodiment of the present application, the coaxial side-illumination mode may include at least two directional coaxial side-illumination modes having different side-illumination directions (for example, four directions distributed at equal angles of 90 degrees), and the first illumination sources 210 enabled in the coaxial illumination source group 21 under any two directional coaxial side-illumination modes are not completely identical.

[0064] In an embodiment of the present application, the annular side-illumination mode may include at least two directional annular side-illumination modes having different side-illumination directions (for example, four directions distributed at equal angles of 90 degrees), and the second illumination sources 220 enabled in the annular illumination source group 22 under any two directional annular side-illumination modes are not completely identical.

[0065] In an embodiment of the present application, each illumination mode may be set so that overall details of the target object 50 are normally presented in an image without localized light-reception abnormality, and the annular side-illumination mode is more advantageous for compensating texture details of an overexposed region of the target object 50 in any one image, thereby compensating for a local overexposed region on the target object 50 due to light reflection of a structural feature; the low-exposure coaxial epi-illumination mode and the low-exposure annular epi-illumination mode help suppress an overexposed region on the target object 50 in any one image.

[0066] As can be seen from the above, by selecting an illumination mode, the imaging exposure condition under which the photosensitive element 12 images the target object 50 can be adjusted and switched among different target conditions.

[0067] In an embodiment of the present application, the selection of the illumination mode of the adjustable light source module 20 may further take into account magnification of the optical lens 11. For example, when the magnification of the optical lens 11 is at a high magnification level, an illumination effect of the annular illuminated region of the annular illumination source group 22 may be less than that of the circular illuminated region of the coaxial illumination source group 21. Therefore, the adjustable light source module 20 may be configured to: respectively illuminate the target object 50 in a normal-exposure coaxial epi-illumination mode, a low-exposure coaxial epi-illumination mode, and at least two directional coaxial side-illumination modes (for example, at least two directions among four directions distributed at equal angles of 90 degrees) when the magnification of the optical lens 11 is greater than or

equal to a preset magnification threshold; and/or, respectively illuminate the target object 50 in a normal-exposure annular epi-illumination mode, a low-exposure annular epi-illumination mode, and at least two directional annular side-illumination modes (for example, at least two directions among four directions distributed at equal angles of 90 degrees) when the magnification of the optical lens is less than the preset magnification threshold.

[0068] In an embodiment of the present application, as a preferred manner when the magnification of the optical lens 11 is greater than or equal to the preset magnification threshold, four directional coaxial side-illumination modes distributed at equal angles of 90 degrees may be selected to illuminate the target object 50. Thus, the target object 50 may be illuminated by cycling through the four directional coaxial side-illumination modes that are equally distributed at 90 degrees intervals. In this way, the side illumination of the target object 50 generated from different directions may have sufficient intensity and may cover the full 360 degree range. Moreover, cycling through the four directional coaxial side-illumination modes distributed at 90 degrees intervals to illuminate the target object 50 may also be used in combination with other illumination modes. For example, the target object 50 may be illuminated respectively in a normal-exposure coaxial epi-illumination mode, a low-exposure coaxial epi-illumination mode, and by cycling through the four directional coaxial side-illumination modes that are equally distributed at 90° intervals, wherein the normal-exposure coaxial epi-illumination mode may cause the target object 50 to receive the strongest illumination, the low-exposure coaxial epi-illumination mode may cause the target object 50 to receive illumination for suppressing overexposure, such that multiple illuminations of the target object 50 may involve different illumination intensities and different illumination directions covering the full 360° range.

[0069] In an embodiment of the present application, the at least two groups of depth image sequences Sq_1 to Sq_m may be acquired sequentially under different imaging exposure conditions, that is, when the imaging exposure condition matches any one target condition, the respective images of the corresponding group of depth image sequence Sq_i are continuously acquired, and then the imaging exposure condition is switched to match another target condition; or the at least two groups of depth image sequences Sq_1 to Sq_m may be acquired in parallel under different imaging exposure conditions, that is, at each depth position, the imaging exposure condition is polled to match at least two target conditions, so as to sequentially acquire images corresponding to the depth position in the at least two groups of depth image sequences Sq_1 to Sq_m.

[0070] Regardless of the order in which the at least two groups of depth image sequences Sq_1 to Sq_m are acquired under different imaging exposure conditions, the images generated by the photosensitive element 12 may be organized into a sequence form of at least two groups of depth image sequences Sq_1 to Sq_m. For example, the images may be organized by a processing device such as an ISP (Image Signal Processing). In this case, exposure duration of the photosensitive element 12 may be controlled by the processing device, and/or the at least two groups of depth image sequences Sq_1 to Sq_m may be subjected by the processing device to image preprocessing including denoising.

[0071] FIG. 2 is an exemplary structural schematic diagram of an extended depth-of-field imaging device according to an embodiment of the present application. Referring to FIG. 2, in an embodiment of the present application, the extended depth-of-field imaging device may be any electronic device having an extended depth-of-field imaging function, such as an extended depth-of-field microscope. For example, the extended depth-of-field microscope may be applied to observation of three-dimensional models of small objects such as circuit boards, lithium batteries, semiconductors, cultural relics, and specimens, and may also be applied to part roughness measurement, size measurement, defect detection of miniature parts, and the like. In an embodiment of the present application, the extended depth-of-field imaging device may include a movable imaging module 10, an adjustable light source module 20, and an image processing module 60.

[0072] In an embodiment of the present application, the movable imaging module 10 of the extended depth-of-field imaging device may be substantially the same as that described above, and will not be elaborated here.

[0073] In an embodiment of the present application, the adjustable light source module 20 may be substantially the same as that described above, or may not be limited to the foregoing description, that is, the adjustable light source module 20 is only required to have a function of adjusting the light-reception state of the target object 50.

[0074] In an embodiment of the present application, the image processing module 60 may be configured to control an illumination mode of the adjustable light source module 20, and the image processing module 60 may further be configured to control the exposure duration of the photosensitive element 12. For example, the image processing module 60 may include a processing device such as the ISP mentioned above. Moreover, the image processing module 60 may further include other processing devices such as a CPU (Central Processing Unit) and a GPU (Graphic Processing Unit), and/or control devices such as an MCU (Micro Controller

Unit) and an FPGA (Field Programmable Gate Array). The embodiments of the present application do not particularly limit components included in the image processing module 60.

[0075] In an embodiment of the present application, the image processing module 60 may further be configured to: generate an extended depth-of-field image I_sd of the target object 50 based on at least two groups of depth image sequences Sq_1 to Sq_m obtained by imaging the target object 50 by the photosensitive element 12 under different imaging exposure conditions; wherein the imaging exposure condition under which the photosensitive element 12 images the target object 50 is associated with the light-reception state of the target object 50, and each group of depth image sequence Sq_i includes a plurality of images obtained respectively by imaging, through the photosensitive element 12, with the same imaging exposure condition when the movable imaging module 10 is located at a plurality of depth positions in the depth direction.

[0076] As can be seen from the above, the extended depth-of-field imaging device in an embodiment of the present application can implement extended depth-of-field imaging by using at least two groups of depth image sequences Sq_1 to Sq_m acquired under different imaging exposure conditions. When structural features of the target object 50 are prone to localized light-reception abnormality, because positions and degrees of the localized light-reception abnormality differ under different imaging exposure conditions, the probability that the same localized light-reception abnormality is present in all depth image sequences Sq_1 to Sq_m acquired under different imaging exposure conditions is extremely low, that is, local image defects present in at least two groups of depth image sequences Sq_1 to Sq_m acquired under different imaging exposure conditions are not completely identical. Therefore, by generating an extended depth-of-field image based on at least two groups of depth image sequences Sq_1 to Sq_m acquired under different imaging exposure conditions, influence of local image defects caused by localized light-reception abnormality on the extended depth-of-field image can be reduced through mutual supplementation of different object information presented by the target object in the at least two groups of depth image sequences Sq_1 to Sq_m, thereby helping improve the accuracy of the three-dimensional object information presented by the extended depth-of-field image.

[0077] In an embodiment of the present application, generation of the extended depth-of-field image I_sd may use at least two parameters, namely a sharpness evaluation value and a confidence. If one group of depth image sequence Sq_i acquired under each imaging exposure condition is regarded as one data source, each group of depth image sequence Sq_i may have a single-source three-dimensional evaluation tensor $T_{i,merg}$ for correspondingly representing

sharpness evaluation values of the group of depth image sequence Sq_i at respective pixel positions, and a single-source confidence map $C_{i,merg}$ for representing a confidence of the sharpness evaluation values of the group of depth image sequence Sq_i.

[0078] The single-source three-dimensional evaluation tensor $T_{i,merg}$ of any one group of depth image sequence Sq_i may include: single-source evaluation vectors $\{\phi_{x,y}(z)_{i,merg}\}$ at respective pixel positions in the single-source three-dimensional evaluation tensor $T_{i,merg}$ of the group of depth image sequence Sq_i. The single-source evaluation vector $\{\phi_{x,y}(z)_{i,merg}\}$ at each pixel position (x, y) may include a plurality of single-source sharpness evaluation values $\phi_{x,y}(z)_{i,merg}$ at the pixel position (x, y) respectively corresponding to a plurality of depth positions z. The single-source confidence map $C_{i,merg}$ of any one group of depth image sequence Sq_i may include: single-source confidences $C(x, y)_{i,merg}$ of the single-source evaluation vectors $\{\phi_{x,y}(z)_{i,merg}\}$ at respective pixel positions in the single-source three-dimensional evaluation tensor $T_{i,merg}$ of the group of depth image sequence Sq_i.

[0079] Thus, the single-source three-dimensional evaluation tensor $T_{i,merg}$ and the single-source confidence map $C_{i,merg}$ of any one group of depth image sequence Sq_i may be respectively represented by the following expressions:

$$T_{i,merg} = \begin{bmatrix} & \cdots & \\ \vdots & \{\phi_{x,y}(z)_{i,merg}\} & \vdots \\ & \cdots & \end{bmatrix}$$

$$C_{i,merg} = \begin{bmatrix} & \cdots & \\ \vdots & C(x,y)_{i,merg} & \vdots \\ & \cdots & \end{bmatrix}$$

[0080] Moreover, because the single-source three-dimensional evaluation tensor $T_{i,merg}$ has three data dimensions, the single-source three-dimensional evaluation tensor $T_{i,merg}$ may also be referred to as a single-source sharpness evaluation volume.

[0081] The single-source three-dimensional evaluation tensors $T_{1,merg}$ to $T_{m,merg}$ of at least two groups of depth image sequences Sq_1 to Sq_m may be fused into a multi-source three-dimensional evaluation tensor $T_{muti\text{-}merg}$, and the single-source confidence maps $C_{1,merg}$ to $C_{m,merg}$ may be fused into a multi-source confidence map $C_{muti\text{-}merg}$.

[0082] The multi-source three-dimensional evaluation tensor $T_{muti\text{-}merg}$ may include multi-source evaluation vectors $\{\phi_{x,y}(z)_{muti\text{-}merg}\}$ at respective pixel positions. The multi-source evaluation vector $\{\phi_{x,y}(z)_{muti\text{-}merg}\}$ at each pixel position (x, y) may include a plurality of multi-

source sharpness evaluation values $\phi_{x,y}(z)_{muti\text{-}merg}$ at the pixel position (x, y) respectively corresponding to a plurality of depth positions. The multi-source confidence map $C_{muti\text{-}merg}$ may include multi-source confidences $C(x, y)_{muti\text{-}merg}$ of the multi-source evaluation vectors $\{\phi_{x,y}(z)_{muti\text{-}merg}\}$ at respective pixel positions in the multi-source three-dimensional evaluation tensor $T_{muti\text{-}merg}$.

[0083] Thus, the multi-source three-dimensional evaluation tensor $T_{muti\text{-}merg}$ and the multi-source confidence map $C_{muti\text{-}merg}$ may be respectively represented by the following expressions:

$$T_{mult-merg} = \begin{bmatrix} \cdots & & \\ \vdots & \{\phi_{x,y}(z)_{mult-merg}\} & \vdots \\ \cdots & & \end{bmatrix}$$

$$C_{multi-merg} = \begin{bmatrix} \cdots & & \\ \vdots & C(x,y)_{mult-merg} & \vdots \\ \cdots & & \end{bmatrix}$$

[0084] Moreover, because the multi-source three-dimensional evaluation tensor $T_{muti\text{-}merg}$ has three data dimensions, the multi-source three-dimensional evaluation tensor $T_{muti\text{-}merg}$ may also be referred to as a multi-source sharpness evaluation volume.

[0085] In this case, the image processing module 60 may be configured to: determine, based on the at least two groups of depth image sequences Sq_1 to Sq_m obtained through imaging the target object 50 by the photosensitive element 12, a single-source three-dimensional evaluation tensor $T_{i,merg}$ and a single-source confidence map $C_{i,merg}$ of each group of depth image sequence Sq_i; determine a multi-source three-dimensional evaluation tensor $T_{muti\text{-}merg}$ based on the single-source three-dimensional evaluation tensors $T_{1,merg}$ to $T_{m,merg}$ of the at least two groups of depth image sequences Sq_1 to Sq_m, and determine the multi-source confidence map $C_{muti\text{-}merg}$ based on the single-source confidence maps $C_{1,merg}$ to $C_{m,merg}$ of the at least two groups of depth image sequences Sq_1 to Sq_m; and generate an extended depth-of-field image I_sd of the target object 50 based on the multi-source three-dimensional evaluation tensor $T_{muti\text{-}merg}$ and the multi-source confidence map $C_{muti\text{-}merg}$.

[0086] Thus, the image processing module 60 may be configured to: based on the at least two groups of depth image sequences Sq_1 to Sq_m, for each group of depth image sequence Sq_i and for each pixel position (x, y), determine, under the imaging exposure condition of the group, single-source sharpness evaluation values $\phi_{x,y}(z)_{i,merg}$ respectively corresponding to a plurality of depth positions, and the single-source confidence $C(x, y)_{i,merg}$ of the single-source sharpness evaluation values $\varphi_{x,y}(z)_{i,merg}$; fuse the single-source sharpness evaluation values

$\varphi_{x,y}(z)_{i,merg}$ at each pixel position (x, y) under the imaging exposure conditions of the at least two groups of depth image sequences Sq_1 to Sq_m into multi-source sharpness evaluation values $\varphi_{x,y}(z)_{muti\text{-}merg}$ at the pixel position (x, y) respectively corresponding to a plurality of depth positions, and fuse the single-source confidence $C(x, y)_{i,merg}$ at each pixel position (x, y) under the imaging exposure conditions of the at least two groups of depth image sequences Sq_1 to Sq_m into a multi-source confidence $C(x, y)_{muti\text{-}merg}$ of the multi-source sharpness evaluation values $\varphi_{x,y}(z)_{muti\text{-}merg}$ at the pixel position (x, y) respectively corresponding to a plurality of depth positions, wherein a fusion weight of the single-source sharpness evaluation value $\varphi_{x,y}(z)_{i,merg}$ at each pixel position (x, y) under the imaging exposure condition of each group of depth image sequence Sq_i is associated with the single-source confidence $C(x, y)_{i,merg}$ at the pixel position (x, y) under the imaging exposure condition of the group of depth image sequence Sq_i; and generate an extended depth-of-field image I_sd of the target object 50 based on the multi-source sharpness evaluation values $\varphi_{x,y}(z)_{muti\text{-}merg}$ respectively corresponding to the plurality of depth positions and the multi-source confidence $C(x, y)_{muti\text{-}merg}$.

[0087] In an embodiment of the present application, the multi-source confidence $C(x, y)_{muti\text{-}merg}$ at each pixel position (x, y) may be used for determining smooth filtering processing on a plurality of multi-source sharpness evaluation values $\phi_{x,y}(z)_{i,merg}$ at the pixel position (x, y) respectively corresponding to a plurality of depth-of-field positions (for example, determining whether smoothing processing is required and/or whether a smoothing threshold of the smoothing processing is reached). Pixel information at each pixel position (x, y) in the extended depth-of-field image I_sd may be determined based on pixel information in images corresponding to a selected depth position at the pixel position (x, y) in the at least two groups of depth image sequences Sq_1 to Sq_m, and the selected depth position corresponding to each pixel position (x, y) may be determined according to a plurality of multi-source sharpness evaluation values $\phi_{x,y}(z)_{muti\text{-}merg}$ respectively corresponding to a plurality of depth-of-field positions at the pixel position (x, y).

[0088] In an embodiment of the present application, generating the extended depth-of-field image I_sd of the target object 50 based on the multi-source three-dimensional evaluation tensor $T_{muti\text{-}merg}$ and the multi-source confidence map $C_{muti\text{-}merg}$ may be a process of synthesizing image information of the at least two groups of depth image sequences Sq_1 to Sq_m based on the multi-source three-dimensional evaluation tensor $T_{muti\text{-}merg}$ and the multi-source confidence map $C_{muti\text{-}merg}$, wherein the process of synthesizing the image information includes at least depth information in the at least two groups of depth image sequences Sq_1 to Sq_m, and a

result of synthesizing the image information may include the extended depth-of-field image I_sd. For example, the extended depth-of-field image I_sd may include: a depth map (structural information of the target object 50) reconstructed based on the depth information in the at least two groups of depth image sequences Sq_1 to Sq_m, and/or an all-focus image (color information of the target object 50) obtained based on image fusion of the at least two groups of depth image sequences Sq_1 to Sq_m.

[0089] In an embodiment of the present application, as one embodiment, generation of the extended depth-of-field image I_sd may further use a highlight mask for representing an overexposed region. In this case: each group of depth image sequence Sq_i may have a corresponding single-source highlight mask $M_{i,merg}$, wherein the single-source highlight mask $M_{i,merg}$ is used for representing an overexposed region of the target object 50 under the imaging exposure condition of the group of depth image sequence Sq_i.

[0090] Exemplarily, in an embodiment of the present application, If, for any one pixel position (x, y), in any set of depth image sequences Sq_i, the pixel values in a plurality of images, the number of which is greater than or equal to a preset image-count threshold, reach a preset highlight threshold (for example, 255), for example, if in the set of depth image sequences Sq_i, at least one image has a pixel value of 255 at any one pixel position in that image, then it may be determined that the pixel position is located within the single-source highlight mask $M_{i,merg}$ corresponding to the set of depth image sequences Sq_i. Moreover, the single-source highlight mask $M_{i,merg}$ corresponding to each group of depth image sequence Sq_i may further be optimized through further image processing such as denoising. For example, the denoising processing may be performed through connected-component detection, by setting local highlight overexposed regions whose connected-component areas are smaller than a preset area threshold to non-highlight, thereby eliminating suspected noise regions in the single-source highlight mask $M_{i,merg}$ that have excessively small connected areas.

[0091] The single-source highlight masks $M_{1,merg}$ to $M_{m,\ merg}$ respectively corresponding to the at least two groups of depth image sequences Sq_1 to Sq_m may be fused into a multi-source highlight mask $M_{muti\text{-}merg}$, wherein the multi-source highlight mask $M_{muti\text{-}merg}$ is used for representing overexposed regions of the target object 50 under different imaging exposure conditions, and the multi-source highlight mask $M_{muti\text{-}merg}$ may be used to optimize the extended depth-of-field image I_sd generated based on the multi-source three-dimensional evaluation tensor $T_{muti\text{-}merg}$ and the multi-source confidence map $C_{muti\text{-}merg}$. For example, the multi-source highlight mask $M_{muti\text{-}merg}$ may be used for identifying an overexposed region in the process of

generating the extended depth-of-field image I_sd based on the multi-source three-dimensional evaluation tensor $T_{muti\text{-}merg}$ and the multi-source confidence map $C_{muti\text{-}merg}$.

[0092] Moreover, as one embodiment, the multi-source confidence map $C_{muti\text{-}merg}$ may further be used to calibrate the multi-source highlight mask $M_{muti\text{-}merg}$.

[0093] In this case, the image processing module 60 may further be configured to: generate, based on the at least two groups of depth image sequences Sq_1 to Sq_m obtained by imaging the target object 50 by the photosensitive element 12, a single-source highlight mask $M_{i,merg}$ corresponding to each group of depth image sequence Sq_i; and generate the multi-source highlight mask $M_{muti\text{-}merg}$ based on the single-source highlight masks $M_{1,merg}$ to $M_{muti\text{-}merg}$ respectively corresponding to the at least two groups of depth image sequences Sq_1 to Sq_m.

[0094] Thus, the image processing module 60 may further be configured to: determine, based on the at least two groups of depth image sequences Sq_1 to Sq_m, an overexposed region of the target object under the imaging exposure condition of each group of depth image sequence Sq_i, wherein fusion of the single-source sharpness evaluation value $\Phi_{x,y}(z)_{i,merg}$ and/or the single-source confidence $C(x, y)_{i,merg}$ at each pixel position (x, y) under the imaging exposure condition of each group of depth image sequence Sq_i is associated with a positional relationship indicating whether the pixel position (x, y) is located within the overexposed region under the imaging exposure condition of the group of depth image sequence Sq_i.

[0095] In an embodiment of the present application, at least two sharpness evaluation metrics may be adopted for the sharpness evaluation values of each group of depth image sequence Sq_i. For example, a combination of at least two spatial-domain-based sharpness evaluation metrics, a combination of at least two transform-domain-based sharpness evaluation metrics, or a combination of a spatial-domain-based sharpness evaluation metric and a transform-domain-based sharpness evaluation metric may be adopted.

[0096] Exemplarily, the embodiments of the present application may adopt at least two spatial-domain-based sharpness evaluation metrics, including at least two of a first sharpness evaluation metric based on gradient, a second sharpness evaluation metric based on Laplacian-transform, and a third sharpness evaluation metric based on statistics. The first sharpness evaluation metric uses GRA operators (Gradient-based operators), the second sharpness evaluation metric uses LAP operators (Laplacian-based operators), and the third sharpness evaluation metric uses STA operators (Statistics-based operators).

[0097] An evaluation principle in which the first sharpness evaluation metric uses GRA operators may be represented as:

$$\phi_{x,y}(z)_{i,1} = \sum_{(p,q)\in\Omega(X1,Y1)} \{[Dx_i(z)_{p,q}]^2 + [Dy_i(z)_{p,q}]^2\}$$

[0098] $Dx_i(z)_{p,q}$ represents a first-order derivative in an x direction (i.e., a row direction) at any neighboring pixel position (p, q) within a first neighborhood window X1×Y1 at any one pixel position (x, y) in any one image corresponding to a depth position z in the depth image sequence Sq_i, and $Dy_i(z)_{p,q}$ represents a first-order derivative in a y direction (i.e., a column direction) at any neighboring pixel position (p, q) within the first neighborhood window X1×Y1 at any one pixel position (x, y) in any one image corresponding to the depth position z in the depth image sequence Sq_i.

[0099] An evaluation principle in which the second sharpness evaluation metric uses LAP operators may be represented as:

$$\phi_{x,y}(z)_{i,2} = \sum_{(p,q)\in\Omega(X2,Y2)} E_i(z)_{p,q}{}^2$$

[0100] Ei(z)p,q represents Laplacian energy at any neighboring pixel position (p, q) within a second neighborhood window X2×Y2 at any one pixel position (x, y) in any one image corresponding to a depth position z in the depth image sequence Sq_i.

[0101] An evaluation principle in which the third sharpness evaluation metric uses STA operators may be represented as:

$$\phi_{x,y}(z)_{i,3} = \sum_{(p,q)\in\Omega(X3,Y3)} [L_i(z)_{p,q} - \overline{L_i(z)}]^2$$

[0102] Li(z)p,q represents variance at any neighboring pixel position (p, q) within a third neighborhood window X3×Y3 at any one pixel position (x, y) in any one image corresponding to a depth position z in the depth image sequence Sq_i; and $\overline{L_\iota(z)}$ represents a mean of variances of an image corresponding to the depth position z in the i-th group of depth image sequence within the neighborhood window X3×Y3 at the pixel position (x, y).

[0103] In the above expressions, $\phi_{x,y}(z)_{i,j}$ represents a result obtained by evaluating the group of depth image sequence Sq_i by using a j-th sharpness evaluation metric, that is, a single-metric sharpness evaluation value corresponding to any one depth position z at any one pixel position (x, y), wherein j is a positive integer greater than or equal to 1 and less than or equal to n, n is a positive integer greater than or equal to 2, and n indicates a total number of types of the at least two sharpness evaluation metrics. For example, if the at least two sharpness evaluation metrics include the first sharpness evaluation metric using the GRA operators, the

second sharpness evaluation metric using the LAP operators, and the third sharpness evaluation metric using the STA operators, n is 3.

[0104] FIG. 3 is a schematic diagram of an extended depth-of-field imaging principle of an extended depth-of-field imaging device according to an embodiment of the present application. Referring to FIG. 3, in an embodiment of the present application, in a process of generating an extended depth-of-field image I_sd of a target object 50 based on the at least two groups of depth image sequences Sq_1 to Sq_m, a process by which the image processing module 60 determines a single-source three-dimensional evaluation tensor $T_{i,merg}$ and a single-source confidence map $C_{i,merg}$ of each group of depth image sequence Sq_i may include: determining, based on at least two groups of depth image sequences Sq_1 to Sq_m, single-metric three-dimensional evaluation tensors $T_{i,1}$ to $T_{i,n}$ and single-metric confidence maps $C_{i,1}$ to $C_{i,n}$ of each group of depth image sequence Sq_i respectively corresponding to at least two (i.e., n types of) sharpness evaluation metrics; and determining a single-source three-dimensional evaluation tensor $T_{i,merg}$ and a single-source confidence map $C_{i,merg}$ of each group of depth image sequence Sq_i based on the single-metric three-dimensional evaluation tensors $T_{i,1}$ to $T_{i,n}$ and the single-metric confidence maps $C_{i,1}$ to $C_{i,n}$ of each group of depth image sequence Sq_i respectively corresponding to the at least two (i.e., n types of) sharpness evaluation metrics.

[0105] The single-metric three-dimensional evaluation tensor $T_{i,j}$ of any one group of depth image sequence Sq_i corresponding to any one sharpness evaluation metric may include: single-metric evaluation vectors $\{\Phi_{x,y}(z)_{i,j}\}$ at respective pixel positions. The single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ at each pixel position (x, y) may include a plurality of single-metric sharpness evaluation values $\Phi_{x,y}(z)_{i,j}$ at the pixel position (x, y) respectively corresponding to a plurality of depth positions z. The single-metric confidence map $C_{i,j}$ of any one group of depth image sequence Sq_i may include: single-metric confidences $C(x, y)_{i,j}$ of the single-metric evaluation vectors $\{\Phi_{x,y}(z)_{i,j}\}$ at respective pixel positions in the single-metric three-dimensional evaluation tensor $T_{i,j}$ of the group of depth image sequence Sq_i.

[0106] In this case, in the single-source three-dimensional evaluation tensor $T_{i,merg}$ of any one group of depth image sequence Sq_i: the single-source evaluation vectors $\{\Phi_{x,y}(z)_{i,merg}\}$ at respective pixel positions may be obtained by fusing the single-metric evaluation vectors $\{\Phi_{x,y}(z)_{i,j}\}$ in the single-metric three-dimensional evaluation tensors $T_{i,1}$ to $T_{i,n}$ of the group of depth image sequence Sq_i respectively corresponding to at least two (i.e., n types of) sharpness evaluation metrics.

[0107] That is, in an embodiment of the present application, the image processing module 60 may be configured to: determine, based on the at least two groups of depth image sequences

Sq_1 to Sq_m and by using each sharpness evaluation metric among at least two sharpness evaluation metrics, single-metric sharpness evaluation values $\{\Phi_{x,y}(z)_{i,j}\}$ at each pixel position (x, y) respectively corresponding to a plurality of depth positions under the imaging exposure condition of each group of depth image sequence Sq_i, and the single-metric confidence $C(x, y)_{i,j}$ of the single-metric sharpness evaluation values $\{\Phi_{x,y}(z)_{i,j}\}$; fuse, the single-metric sharpness evaluation values $\{\Phi_{x,y}(z)_{i,j}\}$ and the single-metric confidence $C(x, y)_{i,j}$ at each pixel position (x, y) under the imaging exposure condition of each group of depth image sequence Sq_i and determined by using at least two sharpness evaluation metrics, into the single-source sharpness evaluation values $\Phi_{x,y}(z)_{i,merg}$ at the pixel position (x, y) respectively corresponding to a plurality of depth positions under the imaging exposure condition of the group of depth image sequence Sq_i, and the single-source confidence $C(x, y)_{i,merg}$ of the single-source sharpness evaluation values $\Phi_{x,y}(z)_{i,merg}$; wherein a fusion weight of the single-metric sharpness evaluation values $\Phi_{x,y}(z)_{i,j}$ of each sharpness evaluation metric for each pixel position (x, y) under the imaging exposure condition of each group of depth image sequence Sq_i may be associated with the single-metric confidence $C(x, y)_{i,j}$ of the same sharpness evaluation metric at the pixel position (x, y) under the imaging exposure condition of the group of depth image sequence Sq_i.

[0108] Moreover, if the embodiment of the present application also uses a highlight mask, the single-metric evaluation vector $\Phi_{x,y}(z)_{i,j}$ may further be used to calibrate the single-source highlight mask $M_{i,merg}$, and the single-metric confidence maps $C_{i,1}$ to $C_{i,n}$ of any one group of depth image sequence Sq_i respectively corresponding to at least two sharpness evaluation metrics may be used to calibrate the single-source highlight mask $M_{i,merg}$. For example, calibrating the single-source highlight mask $M_{i,merg}$ of any one group of depth image sequence Sq_i may include: if any one pixel position (x, y) is present in the overexposed region, and the single-metric confidences $C(x, y)_{i,j}$ at the pixel position (x, y) in the single-metric confidence maps $C_{i,1}$ to $C_{i,n}$ of the at least two sharpness evaluation metrics are all greater than a preset single-metric high-confidence threshold (for example, all being 1), removing the pixel position (x, y) from the single-source highlight mask $M_{i,merg}$ of the group of depth image sequence Sq_i. Then, the single-source highlight mask $M_{i,merg}$ may be subjected to morphological operations (for example, first dilation and then erosion, namely first expanding an overexposed region in the image and then shrinking the overexposed region in the image), and/or pixel position regions having small areas may be removed from the single-source highlight masks $M_{i,merg}$ of the at least two groups of depth image sequences Sq_i.

[0109] In an embodiment of the present application, the process by which the image processing module 60 determines the single-metric three-dimensional evaluation tensor $T_{i,j}$ and the single-metric confidence map $C_{i,j}$ of any one group of depth image sequence Sq_i corresponding to any one sharpness evaluation metric may include: determining, based on the at least two groups of depth image sequences Sq_1 to Sq_m, single-metric three-dimensional evaluation tensors $T_{i,1}$ to $T_{i,n}$ of each group of depth image sequence Sq_i corresponding to at least two sharpness evaluation metrics, wherein, in the single-metric three-dimensional evaluation tensor $T_{i,j}$ of any one group of depth image sequence Sq_i corresponding to any one sharpness evaluation metric, the single-metric sharpness evaluation values $\Phi_{x,y}(z)_{i,j}$of the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ at each pixel position (x, y) are obtained by evaluating the group of depth image sequence Sq_i by using the sharpness evaluation metric, namely, the single-metric sharpness evaluation value $\Phi_{x,y}(z)_{i,j}$ corresponding to any one depth position z in the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ at each pixel position (x, y) is obtained by evaluating an image corresponding to the depth position z in the group of depth image sequence Sq_i by using the sharpness evaluation metric; and determining, based on fusion of the single-metric three-dimensional evaluation tensors $T_{i,1}$ to $T_{i,n}$ respectively of each group of depth image sequence Sq_i corresponding to at least two sharpness evaluation metrics, single-metric confidence maps $C_{i,1}$ to $C_{i,n}$ of each group of depth image sequence Sq_i respectively corresponding to at least two sharpness evaluation metrics, wherein, in the single-metric confidence map $C_{i,j}$ of any one group of depth image sequence Sq_i corresponding to any one sharpness evaluation metric, the single-metric confidence $C(x, y)_{i,j}$ at each pixel position (x, y) is determined based on a vector feature of the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ at the pixel position (x, y).

[0110] In an embodiment of the present application, in order to determine the single-metric confidence maps $C_{i,1}$ to $C_{i,n}$ of each group of depth image sequence Sq_i respectively corresponding to at least two sharpness evaluation metrics, the image processing module 60 may be configured to: detect, in the single-metric three-dimensional evaluation tensor $T_{i,j}$ of each group of depth image sequence Sq_i corresponding to each sharpness evaluation metric, an evaluation value distribution characteristic of the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ at each pixel position, wherein the vector feature used for determining the single-metric confidence $C(x, y)_{i,j}$ may include the evaluation value distribution characteristic of the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$; determine, based on a preset reference distribution characteristic and the evaluation value distribution characteristic detected in the single-metric three-dimensional evaluation tensor $T_{i,j}$ of each group of depth image sequence Sq_i

corresponding to each sharpness evaluation metric, the single-metric confidence maps $C_{i,1}$ to $C_{i,n}$ of each group of depth image sequence Sq_i respectively corresponding to at least two sharpness evaluation metrics; wherein, in the single-metric three-dimensional evaluation tensor $T_{i,j}$ of any one group of depth image sequence Sq_i corresponding to any one sharpness evaluation metric: if the evaluation value distribution characteristic of the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ at any one pixel position (x, y) matches the reference distribution characteristic, the single-metric confidence $C(x, y)_{i,j}$ at the pixel position (x, y) in the single-metric confidence map $C_{i,j}$ of the group of depth image sequence Sq_i corresponding to the sharpness evaluation metric is set to a value representing high confidence (for example, a first value equal to 1); if the evaluation value distribution characteristic of the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ at any one pixel position (x, y) fails to match the reference distribution characteristic, the single-metric confidence $C(x, y)_{i,j}$ at the pixel position (x, y) in the single-metric confidence map $C_{i,j}$ of the group of depth image sequence Sq_i corresponding to the sharpness evaluation metric is set to a value representing low confidence (for example, a second value equal to 0).

[0111] In an embodiment of the present application, the single-metric confidence $C(x,y)_{i,j}$ may not be limited to a binary confidence having a first value and a second value. For example, different sizes may be set for the neighborhood windows used when evaluating each of the sharpness evaluation metrics mentioned above with an operator. Thus, multi-scale evaluation may be performed on any one group of depth image sequences Sq_i based on each sharpness evaluation metric, each pixel position (x, y) may have single-metric evaluation vectors $\{\Phi_{x,y}(z)_{i,j}\}$ at at least two scales. Further, based on distribution characteristics of the evaluation values of the single-metric evaluation vectors $\{\Phi_{x,y}(z)_{i,j}\}$ at the at least two scales, values corresponding respectively to the at least two scales may be determined for the single-metric confidence $C(x,y)_{i,j}$; the sum of the values corresponding respectively to the at least two scales, or the average of the values, may be determined as a non-binary final value of the single-metric confidence $C(x,y)_{i,j}$.

[0112] In an embodiment of the present application, the reference distribution characteristic used for determining the single-metric confidence $C(x, y)_{i,j}$ may include at least one of the following:

[0113] (a) the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ has a unimodal distribution characteristic;

[0114] (b) the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ has a fluctuation amplitude greater than a preset peak-valley fluctuation threshold, that is, a value distribution of the single-metric

sharpness evaluation values $\Phi_{x,y}(z)_{i,j}$ in the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ has sufficiently large separability;

[0115] (c) each value peak of the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ has a value-peak distribution width in the depth direction located within a preset value-peak width interval, that is, the width of the value peak needs to be moderate.

[0116] In an embodiment of the present application, a value-peak determination threshold may be used for determining a value peak, that is, the value peak may be considered as: an evaluation value set of single-metric sharpness evaluation values $\Phi_{x,y}(z)_{i,j}$ in the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ that are greater than or equal to the value-peak determination threshold, and the value-peak determination threshold corresponding to each single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ may be determined according to the single-metric sharpness evaluation values $\Phi_{x,y}(z)_{i,j}$ in the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$. For example, for each single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$, a corresponding value-peak determination threshold may be determined according to a maximum value, a minimum value, and an average value of the single-metric sharpness evaluation values $\Phi_{x,y}(z)_{i,j}$ in the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$, and may be specifically represented by the following expression:

$$\max\left\{\frac{\left(\max\left(\phi_{x,y}(z)_{i,j}\right)+\min\left(\Phi_{x,y}(z)_{i,j}\right)\right)}{2}, mean\left(\phi_{x,y}(z)_{i,j}\right)\right\}$$

[0117] Here, $\max(\Phi_{x,y}(z)_{i,j})$ represents, in the single-metric three-dimensional evaluation tensor $T_{i,j}$ of any one group of depth image sequence Sq_i corresponding to any one sharpness evaluation metric, a maximum value, namely a maximum sharpness evaluation value, of the single-metric sharpness evaluation values $\Phi_{x,y}(z)_{i,j}$ in the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ at any one pixel position (x, y); $\min(\Phi_{x,y}(z)_{i,j})$ represents, in the single-metric three-dimensional evaluation tensor $T_{i,j}$ of any one group of depth image sequence Sq_i corresponding to any one sharpness evaluation metric, a minimum value of the single-metric sharpness evaluation values $\Phi_{x,y}(z)_{i,j}$ in the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ at any one pixel position (x, y); $\text{mean}(\Phi_{x,y}(z)_{i,j})$ represents, in the single-metric three-dimensional evaluation tensor $T_{i,j}$ of any one group of depth image sequence Sq_i corresponding to any one sharpness evaluation metric, an average value of the single-metric sharpness evaluation values $\Phi_{x,y}(z)_{i,j}$ in the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ at any one pixel position (x, y).

[0118] In an embodiment of the present application, with respect to the reference distribution characteristic (a) described above, the unimodal distribution characteristic does not limit that a single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ only has one value peak. That is, the unimodal

distribution characteristic means that all value peaks determined in one single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ are distributed within a same continuous depth position range, a single-metric sharpness evaluation value $\Phi_{x,y}(z)_{i,j}$ corresponding to a depth position in the continuous depth position range is the maximum value $\max(\Phi_{x,y}(z)_{i,j})$ described above, and a range size of the continuous depth position range may be preset and positioned by the depth position corresponding to the maximum value $\max(\Phi_{x,y}(z)_{i,j})$. For example, the image processing module 60 may count the number W1 of value peaks present in each single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$, and then count the number W2 of value peaks in the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ present within the continuous depth position range. Thus, if W2-W1<1 is determined, it is determined that the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ has a unimodal distribution characteristic, that is, the unimodal distribution characteristic (a) is matched successfully; otherwise, it is determined that matching of the unimodal distribution characteristic (a) fails.

[0119] In an embodiment of the present application, with respect to the reference distribution characteristic (b) described above, the fluctuation amplitude of the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ may be determined according to a maximum value and a minimum value of the single-metric sharpness evaluation values $\Phi_{x,y}(z)_{i,j}$ in the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$. For example, the fluctuation amplitude of the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ may be represented by the following expression:

$$\frac{\max\left(\Phi_{x,y}(z)_{i,j}\right)-\min\left(\Phi_{x,y}(z)_{i,j}\right)}{\max\left(\Phi_{x,y}(z)_{i,j}\right)}$$

[0120] In this case, the reference distribution characteristic (b) described above may be represented by the following expression:

$$\frac{\max\left(\Phi_{x,y}(z)_{i,j}\right)-\min\left(\Phi_{x,y}(z)_{i,j}\right)}{\max\left(\Phi_{x,y}(z)_{i,j}\right)}>\mathrm{Thr_}\Phi$$

[0121] Here, $\mathrm{Thr_}\Phi$ represents the preset peak-valley fluctuation threshold.

[0122] In an embodiment of the present application, with respect to the reference distribution characteristic (c) described above, the value-peak distribution width Wth of each value peak of the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ in the depth direction may be constrained by the following expressions: Wth<a×B, and Wth>b×B, wherein B represents the number of images in a group of depth image sequence Sq_i, a and b are respectively configurable coefficients, and b is greater than a.

[0123] In an embodiment of the present application, when light shielding between structural features exists in the target object 50, the image processing module 60 may further be configured to: in a case that any one single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ has two adjacent peaks with an inter-peak interval located within a preset interval range, perform peak clipping suppression on one value peak of the adjacent two peaks, that is located within a position region subject to light shielding, so that a top of the value peak is suppressed to the corresponding value-peak determination threshold.

[0124] The reason why the two adjacent peaks exist may be that: when the optical lens 11 is at a high magnification level, a shielded structural feature is captured with relatively high sharpness at one depth position and another shielding structural feature is unclear, and then the other shielding structural feature is captured with relatively high sharpness at another depth position and the shielded structural feature is unclear. Therefore, in an embodiment of the present application, the preset interval range may be used to represent that the two adjacent peaks are caused by light shielding between structural features of the target object 50 in the depth direction. In the two depth positions respectively corresponding to the two adjacent peaks, the depth position corresponding to the one value peak subjected to peak clipping suppression is located downstream in light reception relative to the depth position corresponding to the other value peak, that is, the value peak subjected to peak clipping suppression represents the shielded structural feature captured with relatively high sharpness, while the unprocessed peak represents the other shielding structural feature captured with relatively high sharpness. Moreover, in an embodiment of the present application, one value peak in each single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ subjected to peak clipping suppression may be suppressed to the value-peak determination threshold corresponding to the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$.

[0125] In an embodiment of the present application, fusion of the single-metric three-dimensional evaluation tensors $T_{i,1}$ to $T_{i,n}$ and the single-metric confidence maps $C_{i,1}$ to $C_{i,n}$ of each group of depth image sequence Sq_i respectively corresponding to at least two (i.e., n types of) sharpness evaluation metrics may adopt an associated fusion manner of the single-metric three-dimensional evaluation tensors $T_{i,1}$ to $T_{i,n}$ and the single-metric confidence maps $C_{i,1}$ to $C_{i,n}$. In this case, the image processing module 60 may be configured to: perform vector intra-source fusion on the single-metric evaluation vectors $\{\Phi_{x,y}(z)_{i,j}\}$ in the single-metric three-dimensional evaluation tensors $T_{i,1}$ to $T_{i,n}$ at each pixel position (x, y) respectively corresponding to at least two sharpness evaluation metrics, wherein, in the single-source three-dimensional evaluation tensor $T_{i,j}$ of any one group of depth image sequence Sq_i, the single-

source sharpness evaluation values $\Phi_{x,y}(z)_{i,merg}$ of the single-source evaluation vector $\{\Phi_{x,y}(z)_{i,merg}\}$ at each pixel position (x, y) are all determined by the vector intra-source fusion; and perform confidence intra-source fusion on the single-metric confidences $C(x, y)_{i,j}$ in the single-metric confidence maps $C_{i,1}$ to $C_{i,n}$ at each pixel position (x, y) respectively corresponding to at least two sharpness evaluation metrics, wherein, in the single-source confidence map $C_{i,j}$ of any one group of depth image sequence Sq_i, the single-source confidence $C(x, y)_{i,merg}$ at each pixel position (x, y) is determined by the confidence intra-source fusion and is bound to and associated with a depth position Z_win_ix,y that contributes to the confidence intra-source fusion; wherein the vector intra-source fusion and the confidence intra-source fusion described above are associated with each other.

[0126] In an embodiment of the present application, the mutually associated vector intra-source fusion and confidence intra-source fusion may employ a voting mechanism, and the depth position Z_win_ix,y contributing to the confidence intra-source fusion may be the depth-of-field position that wins the voting. In this case, the image processing module 60 may be configured to: perform confidence voting on a plurality of depth positions in the depth direction at a granularity of pixel positions based on the single-metric three-dimensional evaluation tensors $T_{i,1}$ to $T_{i,n}$ and the single-metric confidence maps $C_{i,1}$ to $C_{i,n}$ of each group of depth image sequence Sq_i respectively corresponding to at least two sharpness evaluation metrics; and determine, based on a voting result of the confidence voting, the single-source three-dimensional evaluation tensor $T_{i,merg}$ and the single-source confidence map $C_{i,merg}$ of each group of depth image sequence Sq_i. Here: in the single-metric three-dimensional evaluation tensors $T_{i,1}$ to $T_{i,n}$ of any one group of depth image sequence Sq_i respectively corresponding to at least two sharpness evaluation metrics, a fusion weight of the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ at any one pixel position (x, y) in the vector intra-source fusion is associated with the voting result; the single-source confidence $C(x, y)_{i,merg}$ at any one pixel position (x, y) in the single-source confidence map $C_{i,merg}$ of any one group of depth image sequence Sq_i is associated, in the confidence intra-source fusion, with a voting value of the depth position $Z_win_i_{x,y}$ that wins the voting; and the single-source confidence $C(x, y)_{i,merg}$ at any one pixel position (x, y) in the single-source confidence map $C_{i,merg}$ of any one group of depth image sequence Sq_i is associated, in the confidence intra-source fusion, with the depth position $Z_win_i_{x,y}$ that wins the voting.

[0127] Exemplarily, for the confidence voting, the image processing module 60 may be configured to: assign, at each pixel position (x, y), one voting right for the confidence voting to each sharpness evaluation metric; determine, by using the single-metric three-dimensional

evaluation tensor $T_{i,j}$, a voting range set by the confidence voting for each of at least two sharpness evaluation metrics at each pixel position (x, y), wherein, for the single-metric three-dimensional evaluation tensor $T_{i,j}$ of any one group of depth image sequence Sq_i corresponding to any one sharpness evaluation metric, the voting range set by the confidence voting for the sharpness evaluation metric at each pixel position (x, y) includes: a depth position corresponding to a maximum sharpness evaluation value $\max(\Phi_{x,y}(z)_{i,j})$ in the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ at the pixel position (x, y), and other depth positions in a position neighborhood of the depth position; and determine, by using the single-metric confidence map $C_{i,j}$, a confidence voting value contributed by each of at least two sharpness evaluation metrics at each pixel position (x, y), wherein, for the single-metric confidence map $C_{i,j}$ of any one group of depth image sequence Sq_i corresponding to any one sharpness evaluation metric: if the single-metric confidence $C(x, y)_{i,j}$ at any one pixel position (x, y) is greater than or equal to a preset voting value determination threshold, that is, the single-metric confidence $C(x, y)_{i,j}$ at the pixel position (x, y) is at a high confidence level, the voting value contributed by the confidence voting for the sharpness evaluation metric at the pixel position (x, y) is a first voting value (for example, 1); if the single-metric confidence $C(x, y)_{i,j}$ at any one pixel position (x, y) is less than the preset voting value determination threshold, that is, the single-metric confidence $C(x, y)_{i,j}$ at the pixel position (x, y) is at a low confidence level, the voting value contributed by the confidence voting for the sharpness evaluation metric at the pixel position (x, y) is a second voting value smaller than the first voting value (for example, 0.1).

[0128] In an embodiment of the present application, for any one group of depth image sequence Sq_i, maximum sharpness evaluation values $\max(\Phi_{x,y}(z)_{i,j})$ in the single-metric evaluation vectors $\{\Phi_{x,y}(z)_{i,j}\}$ at the same pixel position (x, y) in the single-metric three-dimensional evaluation tensors $T_{i,j}$ respectively corresponding to different sharpness evaluation metrics are not necessarily the same. Therefore, for the single-metric three-dimensional evaluation tensor $T_{i,j}$ of any one group of depth image sequence Sq_i corresponding to any one sharpness evaluation metric, voting ranges set by the confidence voting for at least two sharpness evaluation metrics at each pixel position (x, y) may be different.

[0129] Assuming that voting ranges set by the confidence voting for at least two sharpness evaluation metrics at a certain pixel position (x, y) are different: if a certain depth position only falls within a voting range set by the confidence voting for one sharpness evaluation metric at the certain pixel position (x, y), and because the single-metric confidence $C(x, y)_{i,j}$ at the pixel position (x, y) is at a low confidence level, the depth position obtains only a second voting value (for example, 0.1) within a voting range where it is located, then the voting value obtained

by the depth position is only one second voting value (for example, 0.1); if another depth position simultaneously falls within voting ranges set by the confidence voting for all (n) sharpness evaluation metrics at the certain pixel position (x, y), and because the single-metric confidence $C(x, y)_{i,j}$ at the pixel position (x, y) is at a high confidence level, the depth position obtains a first voting value (for example, 1) in all voting ranges, then the voting value obtained by the depth position is a sum of n first voting values (for example, 1).

[0130] The above two extreme cases may represent: when any one depth position falls within a voting range set by the confidence voting for at least one sharpness evaluation metric at any one pixel position (x, y), a lower limit value (namely one second voting value) and an upper limit value (namely n first voting values) of voting values that the depth position may obtain. In the confidence voting, the highest voting value obtained by the depth position $Z_win_i_{x,y}$ that wins the voting may be located within a voting value interval bounded by the lower limit value and the upper limit value described above.

[0131] Exemplarily, for the confidence voting, the image processing module 60 may further be configured to: in the single-source confidence map $C_{i,merg}$ of each group of depth image sequence Sq_i, set the single-source confidence $C(x, y)_{i,merg}$ at each pixel position (x, y) as the highest voting value obtained by the depth position $Z_win_i_{x,y}$ that wins in the confidence voting; in the single-source three-dimensional evaluation tensor $T_{i,merg}$ of each group of depth image sequence Sq_i, perform the vector intra-source fusion by using, as weights, voting values contributed by respective pixel positions to the depth position winning the voting, wherein, in the single-source evaluation vector $\{\Phi_{x,y}(z)_{i,merg}\}$ at any one pixel position (x, y), the single-source sharpness evaluation value $\Phi_{x,y}(z)_{i,merg}$ corresponding to any one depth position z is a weighted value, using the voting values contributed by confidence voting at the pixel position (x, y) to the winning depth position as weights, of the normalized results of the single-metric sharpness evaluation value $\Phi_{x,y}(z)_{i,j}$ corresponding to the depth position $Z_win_i_{x,y}$ in the single-metric evaluation vectors $\{\Phi_{x,y}(z)_{i,j}\}$ of the single-metric three-dimensional evaluation tensors $T_{i,j}$ respectively corresponding to at least two sharpness evaluation metrics.

[0132] For example, in the single-source evaluation vector $\{\Phi_{x,y}(z)_{i,merg}\}$ at any one pixel position (x, y), the single-source sharpness evaluation value $\Phi_{x,y}(z)_{i,merg}$ corresponding to any one depth position z may be represented by the following expression:

$$\phi_{x,y}(z)_{i,merg} = \sum \widehat{\Phi}_{x,y}(z)_{i,j} \times V_{x,y}(Z_win_i_{x,y})_{i,j}$$

$$\widehat{\Phi}_{x,y}(z)_{i,j} = \frac{\phi_{x,y}(z)_{i,j}}{sum\{\phi_{x,y}(z)_{i,j}\}}$$

[0133] Here: $\hat{\varphi}_{x,y}(z)_{i,j}$ represents a normalized result of the single-metric sharpness evaluation value $\Phi_{x,y}(z)_{i,j}$ in the single-metric evaluation vector $\{\Phi_{x,y}(z)_{i,j}\}$ at any one pixel position (x, y) in the single-metric three-dimensional evaluation tensor $T_{i,j}$ of any one group of depth image sequence Sq_i corresponding to any one sharpness evaluation metric; $Vx, y(Z_win_i_{x,y})_{i,j}$ represents, for any one group of depth image sequence Sq_i, a voting value contributed by any one pixel position (x, y) to a depth-of-field position $Z_win_i_{x,y}$ winning the voting (i.e., having the highest voting value) for any one sharpness evaluation metric; for example, if the depth-of-field position $Z_win_i_{x,y}$ winning the voting (i.e., having the highest voting value) falls within a voting range set by the confidence voting for the sharpness evaluation metric at the pixel position (x, y), the voting value contributed by the pixel position (x, y) to the depth-of-field position $Z_win_i_{x,y}$ winning the voting (i.e., having the highest voting value) for the sharpness evaluation metric may be one first voting value or one second voting value described above; alternatively, if the depth-of-field position $Z_win_i_{x,y}$ winning the voting (i.e., having the highest voting value) does not fall within the voting range set by the confidence voting for the sharpness evaluation metric at the pixel position (x, y), the voting value contributed by the pixel position (x, y) to the depth-of-field position $Z_win_i_{x,y}$ winning the voting (i.e., having the highest voting value) for the sharpness evaluation metric may be 0.

[0134] In an embodiment of the present application, fusion of the single-source three-dimensional evaluation tensors $T_{1,merg}$ to $T_{m,merg}$ and the single-source confidence maps $C_{1,merg}$ to $C_{m,merg}$ of at least two groups of depth image sequences Sq_1 to Sq_m may adopt an associated fusion manner. In this case, the image processing module 60 may be configured to: perform vector multi-source fusion on the single-source evaluation vectors $\{\Phi_{x,y}(z)_{i,merg}\}$ at each pixel position (x, y) in the single-source three-dimensional evaluation tensors $T_{1,merg}$ to $T_{m,merg}$ of the at least two groups of depth image sequences Sq_1 to Sq_m, wherein the multi-source sharpness evaluation values $\Phi_{x,y}(z)_{muti\text{-}merg}$ of the multi-source evaluation vector $\{\Phi_{x,y}(z)_{muti\text{-}merg}\}$ at each pixel position (x, y) are determined by vector multi-source fusion; perform confidence multi-source fusion on the single-source confidences $C(x, y)_{i,merg}$ at each pixel position (x, y) in the single-source confidence maps $C_{1,merg}$ to $C_{m,merg}$ of the at least two groups of depth image sequences Sq_1 to Sq_m, wherein the multi-source confidence $C(x, y)_{muti\text{-}merg}$ at each pixel position (x, y) in the multi-source confidence map $C_{muti\text{-}merg}$ is determined by confidence multi-source fusion; wherein the vector multi-source fusion and the confidence multi-source fusion described above are associated with each other.

[0135] In an embodiment of the present application, the mutually associated vector multi-source fusion and confidence multi-source fusion may be implemented by evaluating a confidence level. In this case, the image processing module 60 may be configured to: determine, based on the single-source confidence maps $C_{1,merg}$ to $C_{m,merg}$ of the at least two groups of depth image sequences Sq_1 to Sq_m, a maximum single-source confidence among at least two single-source confidences $C(x, y)_{i,merg}$ at each pixel position (x, y) in the single-source confidence maps $C_{1,merg}$ to $C_{m,merg}$; and perform the above vector multi-source fusion and confidence multi-source fusion based on a matching confidence level of the maximum single-source confidence at each pixel position (x, y) among preset confidence levels.

[0136] The preset confidence levels may include a high confidence level, a medium confidence level, and a low confidence level.

[0137] If, for any one pixel position (x, y), the matching confidence level of the maximum single-source confidence is the high confidence level or the medium confidence level, the vector multi-source fusion at the pixel position (x, y) may be configured, such that a multi-source sharpness evaluation value $\Phi_{x,y}(z)_{muti\text{-}merg}$ corresponding to any one depth position z in the multi-source evaluation vector $\{\Phi_{x,y}(z)_{muti\text{-}merg}\}$, is determined by an average of single-source sharpness evaluation values $\Phi_{x,y}(z)_{i,merg}$ corresponding to the depth position z in k single-source evaluation vectors $\{\Phi_{x,y}(z)_{i,merg}\}$ whose single-source confidences $C(x, y)_{i,merg}$ reach the matching confidence level, for example the vector multi-source fusion at the pixel position (x, y) is set as the average, wherein k is a positive integer less than or equal to m.

[0138] If the matching confidence level of the maximum single-source confidence at any one pixel position (x, y) is the low confidence level, the vector multi-source fusion at the pixel position (x, y) is configured, such that a multi-source sharpness evaluation value $\phi_{x,y}(z)_{muti\text{-}merg}$ corresponding to any one depth position z in the multi-source evaluation vector $\{\phi_{x,y}(z)_{muti\text{-}merg}\}$ is a single-source sharpness evaluation value $\Phi_{x,y}(z)_{i,merg}$ corresponding to the depth position z in the single-source evaluation vector $\{\Phi_{x,y}(z)_{i,merg}\}$ having the maximum single-source confidence.

[0139] For any one pixel position (x, y), regardless of whether the matching confidence level of the maximum single-source confidence thereof is the high confidence level, the medium confidence level, or the low confidence level, the confidence multi-source fusion at the pixel position (x, y) is configured to: determine the multi-source confidence $C(x, y)_{muti\text{-}merg}$ at the pixel position (x, y) in the multi-source confidence map $C_{muti\text{-}merg}$ based on the matching confidence level at the pixel position (x, y).

[0140] In an embodiment of the present application, the image processing module 60 may be configured to determine the multi-source confidence $C(x, y)_{muti\text{-}merg}$ at any one pixel position (x, y) in the multi-source confidence map $C_{muti\text{-}merg}$ in the following manner: if the matching confidence level of the maximum single-source confidence at any one pixel position (x, y) is the high confidence level or the medium confidence level, verifying the matching confidence level at the pixel position (x, y) based on a high-confidence-level matching number k of single-source confidences $C(x, y)_{i,merg}$ at the pixel position (x, y) that match the high confidence level and associated depth positions $Z_win_i_{x, y}$ of the single-source confidences $C(x, y)_{i,merg}$ at the pixel position (x, y) that match the high confidence level, wherein when verification of the matching confidence level at any one pixel position (x, y) succeeds, the multi-source confidence $C(x, y)_{muti\text{-}merg}$ at the pixel position (x, y) in the multi-source confidence map $C_{muti\text{-}merg}$ is set to meet the matching confidence level at the pixel position; and when verification of the matching confidence level at any one pixel position (x, y) fails, the multi-source confidence $C(x, y)_{muti\text{-}merg}$ at the pixel position (x, y) in the multi-source confidence map $C_{muti\text{-}merg}$ is set to meet a next lower confidence level relative to the matching confidence level at the pixel position (x, y), that is, downgraded from the high confidence level to the medium confidence level or from the medium confidence level to the low confidence level; if the matching confidence level of the maximum single-source confidence at any one pixel position (x, y) is the low confidence level, the multi-source confidence $C(x, y)_{muti\text{-}merg}$ at the pixel position (x, y) in the multi-source confidence map $C_{muti\text{-}merg}$ is set according to an average of at least two single-source confidences $C(x, y)_{i,merg}$ at the pixel position (x, y) in the single-source confidence maps $C_{1,merg}$ to $C_{m,merg}$.

[0141] Exemplarily, if the matching confidence level of the maximum single-source confidence at any one pixel position (x, y) is the high confidence level or the medium confidence level, the image processing module 60 may be configured to implement the verification described above in the following manner: if the high-confidence-level matching number k at any one pixel position (x, y) is greater than or equal to a confidence determination threshold confthre, and a position distribution span of the associated depth positions $Z_win_i_{x, y}$ of the single-source confidences $C(x, y)_{i,merg}$ at the pixel position (x, y) that match the high confidence level is less than or equal to a preset interval threshold, determining that verification of the matching confidence level at the pixel position (x, y) succeeds; if the high-confidence-level matching number k at any one pixel position (x, y) is less than the confidence determination threshold confthre, and/or the position distribution span of the associated depth positions $Z_win_i_{x, y}$ of the single-source confidences $C(x, y)_{i,merg}$ at the pixel position (x, y)

that match the high confidence level is greater than the preset interval threshold, determining that verification of the matching confidence level at the pixel position (x, y) fails.

[0142] Exemplarily, in a case that the matching confidence level of the maximum single-source confidence at any one pixel position (x, y) is the medium confidence level, the above verification may further include: verifying whether the pixel position (x, y) is located within the single-source highlight mask $M_{i,merg}$ of the group of depth image sequence Sq_i corresponding to the maximum single-source confidence thereof, wherein, if the high-confidence-level matching number k at the pixel position (x, y) is greater than or equal to the confidence determination threshold confthre, the position distribution span of the associated depth positions $Z_win_i_{x,y}$ of the single-source confidences $C(x, y)_{i,merg}$ at the pixel position (x, y) that match the high confidence level is less than or equal to the preset interval threshold, and the pixel position (x, y) is located outside the single-source highlight mask $M_{i,merg}$ of the group of depth image sequence Sq_i corresponding to the maximum single-source confidence thereof, determining that verification of the matching confidence level at the pixel position (x, y) succeeds; if the high-confidence-level matching number k at the pixel position (x, y) is less than the confidence determination threshold confthre, and/or the position distribution span of the associated depth positions $Z_win_i_{x,y}$ of the single-source confidences $C(x, y)_{i,merg}$ at the pixel position (x, y) that match the high confidence level is greater than the preset interval threshold, and/or the pixel position (x, y) is located within the single-source highlight mask $M_{i,merg}$ of the group of depth image sequence Sq_i corresponding to the maximum single-source confidence thereof, determining that verification of the matching confidence level at the pixel position (x, y) fails.

[0143] Exemplarily, the confidence determination threshold *confthre* used in the above verification process may be an adaptive value associated with the single-source confidence $C(x, y)_{i,merg}$. In this case, the image processing module 60 may be configured to: determine, based on the single-source confidence maps $C_{1,merg}$ to $C_{m,merg}$ of the at least two groups of depth image sequences Sq_1 to Sq_m, a single-source confidence sum $\sum C(x,y)_{i,merg}$ of at least two single-source confidences $C(x, y)_{i,merg}$ at each pixel position (x, y); and determine the confidence determination threshold confthre meet the pixel position (x, y) based on the single-source confidence sum $\sum C(x,y)_{i,merg}$ at each pixel position (x, y).

[0144] For example, determination of the confidence determination threshold *confthre* may refer to a matching result between the single-source confidence sum $\sum C(x,y)_{i,merg}$ and preset confidence grading intervals, wherein an example of determining the confidence

determination threshold confthre with reference to the matching result between the single-source confidence sum $\sum C(x,y)_{i,merg}$ and the preset confidence grading intervals is shown in the following expression:

$$confthre = \begin{cases} \max(1, 0.2 \times m), & \sum C(x,y)_{i,merg} < 0.5 \\ \max(1, 0.3 \times m), & 0.5 \leq \sum C(x,y)_{i,merg} < 0.8 \\ \max(1, 0.4 \times m), & 0.8 \leq \sum C(x,y)_{i,merg} \end{cases}$$

[0145] In the example shown in the above expression, the preset confidence grading intervals are merely taken as including (0, 0.5), [0.5, 0.8), and [0.8, +infinity), and the preset confidence grading intervals may be arbitrarily set as required. Similarly, coefficients multiplied by m in piecewise values of the confidence determination threshold confthre may also be arbitrarily set as required, wherein m is the number of groups of the depth image sequences, and the value "1" as a lower limit value for constraining the confidence determination threshold confthre may also be set to another value as required.

[0146] In an embodiment of the present application, if generation of the extended depth-of-field image adopts a highlight mask, the image processing module 60 may be further configured to implement mask fusion of the single-source highlight masks $M_{1,merg}$ to $M_{muti\text{-}merg}$ in the following manner: determining whether each pixel position (x, y) is located within the multi-source highlight mask $M_{muti\text{-}merg}$ based on the number of times the pixel position (x, y) is present in the single-source highlight masks $M_{1,merg}$ to $M_{muti\text{-}merg}$ respectively corresponding to the at least two groups of depth image sequences Sq_1 to Sq_m, that is, preliminary fusion of the single-source highlight masks $M_{1,merg}$ to $M_{muti\text{-}merg}$, wherein if the number of times present in the single-source highlight masks is greater than or equal to a preset number threshold, the pixel position (x, y) is determined to be located within the multi-source highlight mask $M_{muti\text{-}merg}$, otherwise the pixel position (x, y) is determined to be outside the multi-source highlight mask $M_{muti\text{-}merg}$; and/or, in the single-source highlight mask $M_{i,merg}$ of the same group of depth image sequence Sq_i corresponding to the maximum single-source confidence at each pixel position (x, y) (for example, generally represented by $C(x, y)_{i,merg}$), if the pixel position (x, y) is located outside the single-source highlight mask $M_{i,merg}$ of the group of depth image sequence Sq_i, and the multi-source confidence $C(x, y)_{muti\text{-}merg}$ at the pixel position (x, y) in the multi-source confidence map $C_{muti\text{-}merg}$ meet the above high confidence level, prohibiting the pixel position (x, y) from being placed in the multi-source highlight mask $M_{muti\text{-}merg}$, that is, calibrating the multi-source highlight mask $M_{muti\text{-}merg}$ by using the multi-source confidence map $C_{muti\text{-}merg}$.

[0147] FIG. 4 is an exemplary flowchart of an image generation method for an extended depth-of-field image according to an embodiment of the present application. Referring to FIG. 4, an embodiment of the present application further provides an image generation method for an extended depth-of-field image, and the image generation method may include steps S410 to S430.

[0148] S410: acquiring at least two groups of depth image sequences obtained by imaging a target object by a photosensitive element of a movable imaging module under different imaging exposure conditions, wherein the imaging exposure condition under which the photosensitive element images the target object is associated with a light-reception state of the target object, and each group of the depth image sequences includes a plurality of images captured by the photosensitive element with the same imaging exposure condition when the movable imaging module is located at a plurality of depth positions in the depth direction, respectively.

[0149] S430: generating the extended depth-of-field image of the target object based on the at least two groups of depth image sequences.

[0150] For the specific description of the extended depth-of-field image, reference may be made to the foregoing description, and repeated description is omitted here.

[0151] As can be seen from the above, the image generation method in the embodiment of the present application generates an extended depth-of-field image based on at least two groups of depth image sequences acquired under different imaging exposure conditions. When structural features of the target object are prone to localized light-reception abnormality, because positions and degrees of the localized light-reception abnormality differ under different imaging exposure conditions, the probability that the same localized light-reception abnormality is present in all depth image sequences acquired under different imaging exposure conditions is extremely low, that is, local image defects present in at least two groups of depth image sequences acquired under different imaging exposure conditions are not completely identical. Therefore, by generating an extended depth-of-field image based on at least two groups of depth image sequences acquired under different imaging exposure conditions, influence of local image defects caused by localized light-reception abnormality on the extended depth-of-field image can be reduced through mutual supplementation of different object information presented by the target object in the at least two groups of depth image sequences, thereby helping improve the accuracy of three-dimensional object information presented by the extended depth-of-field image.

[0152] FIG. 5 is a first optimized flowchart of an image generation method for an extended depth-of-field image according to an embodiment of the present application. Referring to FIG.

5, in an embodiment of the present application, if sharpness evaluation values and a confidence are introduced, S430 in the process flow illustrated in FIG. 4 may include S510 to S550.

[0153] S510: determining, based on at least two groups of depth image sequences, for each pixel and for each group of depth image sequences, single-source sharpness evaluation values at the pixel position respectively corresponding to a plurality of depth positions under the imaging exposure condition of the group of depth image sequences, and a single-source confidence of the single-source sharpness evaluation values.

[0154] For example, S510 may determine, based on the at least two groups of depth image sequences, a single-source three-dimensional evaluation tensor and a single-source confidence map of each group of depth image sequences; wherein the single-source three-dimensional evaluation tensor of any one group of depth image sequences includes: single-source evaluation vectors at respective pixel positions, each single-source evaluation vector including a plurality of single-source sharpness evaluation values at the pixel position respectively corresponding to a plurality of depth positions; and the single-source confidence map of any one group of depth image sequences includes: single-source confidences of the single-source evaluation vectors of respective pixel positions in the single-source three-dimensional evaluation tensor of the group of depth image sequences.

[0155] S530: for each pixel position, fusing the single-source sharpness evaluation values and the single-source confidence at the pixel position under the imaging exposure conditions of the at least two groups of depth image sequences respectively into multi-source sharpness evaluation values at the pixel position respectively corresponding to a plurality of depth positions and a multi-source confidence of the multi-source sharpness evaluation values, wherein a fusion weight of the single-source sharpness evaluation values at each pixel position under the imaging exposure condition of each group of depth image sequences is associated with the single-source confidence at the pixel position under the imaging exposure condition of the group of depth image sequences.

[0156] For example, in S530, a multi-source three-dimensional evaluation tensor and a multi-source confidence map can be determined based on the single-source three-dimensional evaluation tensors and the single-source confidence maps of the at least two groups of depth image sequences; wherein the multi-source three-dimensional evaluation tensor includes multi-source evaluation vectors at respective pixel positions, the multi-source evaluation vector at each pixel position including a plurality of multi-source sharpness evaluation values at the pixel position respectively corresponding to a plurality of depth positions; and the multi-source

confidence map includes multi-source confidences of the multi-source evaluation vectors at respective pixel positions in the multi-source three-dimensional evaluation tensor.

[0157] S550: generating an extended depth-of-field image of the target object based on the multi-source confidence and the multi-source sharpness evaluation values corresponding to the plurality of depth positions.

[0158] For example, in S550, the extended depth-of-field image of the target object can be generated based on the multi-source three-dimensional evaluation tensor and the multi-source confidence map corresponding to the plurality of depth positions.

[0159] In an embodiment of the present application, the multi-source confidence at each pixel position may be used to determine smooth filtering processing on a plurality of multi-source sharpness evaluation values at the pixel position respectively corresponding to a plurality of depth-of-field positions (for example, determining whether smoothing processing is required and/or a smoothing threshold of the smoothing processing). Pixel information at each pixel position in the extended depth-of-field image I_sd may be determined based on pixel information in images at a selected depth position corresponding to the pixel position in the at least two groups of depth image sequences, and the selected depth position corresponding to each pixel position may be determined according to a plurality of multi-source sharpness evaluation values at the pixel position respectively corresponding to a plurality of depth-of-field positions.

[0160] In an embodiment of the present application, if an overexposed region is further introduced to optimize the extended depth-of-field image, S430 in the process flow illustrated in FIG. 4 may further include: determining, based on at least two groups of depth image sequences, an overexposed region of the target object under the imaging exposure condition of each group of depth image sequences, wherein fusion of the single-source sharpness evaluation values and/or the single-source confidence at each pixel position under the imaging exposure condition of each group of depth image sequences is associated with a positional relationship indicating whether the pixel position is located within the overexposed region under the imaging exposure condition of the group of depth image sequences. For example, a single-source highlight mask corresponding to each group of depth image sequences is generated based on the at least two groups of depth image sequences, wherein the single-source highlight mask corresponding to any one group of depth image sequences is used to represent an overexposed region of the target object under the imaging exposure condition of the group of depth image sequences; for example, if pixel values at any one pixel position in several images greater than or equal to a preset image number threshold in any one group of depth image

sequences reach a preset highlight threshold, the pixel position may be determined to be located within the single-source highlight mask corresponding to the group of depth image sequences; a multi-source highlight mask is generated based on mask fusion of the single-source highlight masks respectively corresponding to the at least two groups of depth image sequences, wherein the multi-source highlight mask is used to represent overexposed regions of the target object under different imaging exposure conditions, and the multi-source highlight mask is used to optimize the extended depth-of-field image generated based on multi-source fusion, that is, S550 of the process flow illustrated in FIG. 5 may further use the multi-source highlight mask when generating the extended depth-of-field image. In addition, the multi-source confidence map may be used to calibrate the multi-source highlight mask.

[0161] FIG. 6 illustrates a second optimized flowchart of an image generation method for an extended depth-of-field image according to an embodiment of the present application. Referring to FIG. 6, in an embodiment of the present application, if at least two sharpness evaluation metrics are introduced to determine the single-source three-dimensional evaluation tensor and the single-source confidence map, S510 of the process flow illustrated in FIG. 5 may include steps S610 to S630.

[0162] S610: determining, based on at least two groups of depth image sequences and by respectively using at least two sharpness evaluation metrics, single-metric sharpness evaluation values at each pixel position respectively corresponding to a plurality of depth positions under the imaging exposure condition of each group of depth image sequences, and a single-metric confidence of the single-metric sharpness evaluation values.

[0163] For example, in S610, based on the at least two groups of depth image sequences, single-metric three-dimensional evaluation tensors and single-metric confidence maps of each group of depth image sequences respectively corresponding to at least two sharpness evaluation metrics can be determined; wherein the single-metric three-dimensional evaluation tensor of any one group of depth image sequences includes: single-metric evaluation vectors at respective pixel positions, each single-metric evaluation vector including a plurality of single-metric sharpness evaluation values at the pixel position respectively corresponding to a plurality of depth positions; and the single-metric confidence map of any one group of depth image sequences includes: single-metric confidences of the single-metric evaluation vectors at respective pixel positions in the single-metric three-dimensional evaluation tensor of the group of depth image sequences.

[0164] For example, in S610, at least two of a gradient-based first sharpness evaluation metric, a Laplacian-based second sharpness evaluation metric, and a statistics-based third sharpness

evaluation metric may be used to determine a single-metric three-dimensional evaluation tensor of each group of depth image sequences respectively corresponding to each of them.

[0165] S630: fusing, for each pixel position under the imaging exposure condition of each group of depth image sequences, the single-metric sharpness evaluation values and the single-metric confidence respectively determined by using at least two sharpness evaluation metrics into the single-source sharpness evaluation values at the pixel position respectively corresponding to a plurality of depth positions under the imaging exposure condition of the group of depth image sequences and the single-source confidence of the single-source sharpness evaluation values, wherein a fusion weight of the single-metric sharpness evaluation values of each sharpness evaluation metric at each pixel position under the imaging exposure condition of each group of depth image sequences may be associated with the single-metric confidence of the same sharpness evaluation metric at the pixel position under the imaging exposure condition of the group of depth image sequences.

[0166] For example, in S630, the single-source three-dimensional evaluation tensor and the single-source confidence map of each group of depth image sequences can be determined based on fusion of the single-metric three-dimensional evaluation tensors and the single-metric confidence maps of each group of depth image sequences respectively corresponding to at least two sharpness evaluation metrics.

[0167] In an embodiment of the present application, if an overexposed region is further introduced to optimize the extended depth-of-field image, the single-metric confidence maps may be used to calibrate the single-source highlight mask. Exemplarily, S610 described above may include: determining, based on the at least two groups of depth image sequences, single-metric three-dimensional evaluation tensors of each group of depth image sequences respectively corresponding to at least two sharpness evaluation metrics, wherein, in the single-metric three-dimensional evaluation tensor of any one group of depth image sequences corresponding to any one sharpness evaluation metric, the single-metric sharpness evaluation values of the single-metric evaluation vector at each pixel position are obtained by evaluating the group of depth image sequences by using the sharpness evaluation metric; and determining, based on the single-metric three-dimensional evaluation tensors of each group of depth image sequences respectively corresponding to at least two sharpness evaluation metrics, single-metric confidence maps of each group of depth image sequences respectively corresponding to at least two sharpness evaluation metrics, wherein, in the single-metric confidence map of any one group of depth image sequences corresponding to any one sharpness evaluation metric, the

single-metric confidence at each pixel position is determined based on a vector feature of the single-metric evaluation vector at the pixel position.

[0168] In an embodiment of the present application, if an overexposed region is further introduced to optimize the extended depth-of-field image, and the single-metric confidence maps may be used to calibrate the single-source highlight mask, then, the single-source highlight mask can be calibrated by using the single-metric confidence maps of each group of depth image sequences respectively corresponding to at least two sharpness evaluation metrics. For example, calibration of the single-source highlight mask of any one group of depth image sequences includes: if any one pixel position is present in the single-source highlight mask, and single-metric confidences at the pixel position in the single-metric confidence maps of at least two sharpness evaluation metrics are all greater than a preset single-metric high-confidence threshold, removing the pixel position from the single-source highlight mask.

[0169] In an embodiment of the present application, the process in which determining the single-metric confidence maps of each group of depth image sequences respectively corresponding to at least two sharpness evaluation metrics in S610 may include: detecting, in the single-metric three-dimensional evaluation tensor of each group of depth image sequences corresponding to each sharpness evaluation metric, an evaluation value distribution characteristic of the single-metric evaluation vector at each pixel position, wherein the vector feature includes the evaluation value distribution characteristic; determining, based on a preset reference distribution characteristic and the evaluation value distribution characteristic detected in the single-metric three-dimensional evaluation tensor of each group of depth image sequences corresponding to each sharpness evaluation metric, the single-metric confidence maps of each group of depth image sequences respectively corresponding to at least two sharpness evaluation metrics; wherein, in the single-metric three-dimensional evaluation tensor of any one group of depth image sequences corresponding to any one sharpness evaluation metric: if evaluation value distribution characteristic at any one pixel position matches reference distribution characteristic, the single-metric confidence at the pixel position in the single-metric confidence map of the group of depth image sequences corresponding to the sharpness evaluation metric is set to a value representing high confidence; if the evaluation value distribution characteristic at any one pixel position fails to match the reference distribution characteristic, the single-metric confidence at the pixel position in the single-metric confidence map of the group of depth image sequences corresponding to the sharpness evaluation metric is set to a value representing low confidence.

[0170] For example, the reference distribution characteristic used in S610 may include at least one of the following: the single-metric evaluation vector has a unimodal distribution characteristic, the single-metric evaluation vector has a fluctuation amplitude greater than a preset peak-valley fluctuation threshold, and each value peak of the single-metric evaluation vector has a value-peak distribution width in the depth direction located within a preset value-peak width interval. For the definition of the value peak and the explanation of the reference distribution characteristic described above, reference may be made to the foregoing description.

[0171] In an embodiment of the present application, S610 may further include: if any one single-metric evaluation vector has two adjacent peaks with an inter-peak interval located within a preset interval range, performing peak clipping suppression on one value peak, among the two adjacent peaks, that is located within a position region subject to light shielding, so that a top of the value peak is suppressed to the corresponding value-peak determination threshold.

[0172] In an embodiment of the present application, S630 may include: performing vector intra-source fusion on the single-metric evaluation vectors at each pixel position in the single-metric three-dimensional evaluation tensors respectively corresponding to at least two sharpness evaluation metrics, wherein, in the single-source three-dimensional evaluation tensor of any one group of depth image sequences, the single-source sharpness evaluation values of the single-source evaluation vector at each pixel position are all determined by vector intra-source fusion; performing confidence intra-source fusion on the single-metric confidence at each pixel position in the single-metric confidence maps respectively corresponding to at least two sharpness evaluation metrics, wherein, in the single-source confidence map of any one group of depth image sequences, the single-source confidence at each pixel position is determined by confidence intra-source fusion and is bound to a depth position contributing to the confidence intra-source fusion; wherein the vector intra-source fusion and the confidence intra-source fusion are associated with each other.

[0173] In an embodiment of the present application, the vector intra-source fusion and the confidence intra-source fusion that are mutually associated may include: performing confidence voting respectively on a plurality of depth positions in the depth direction at a granularity of pixel positions, based on the single-metric three-dimensional evaluation tensors and the single-metric confidence maps of each group of depth image sequences respectively corresponding to at least two sharpness evaluation metrics; and determining, based on a voting result of the confidence voting, the single-source three-dimensional evaluation tensor and the single-source confidence map of each group of depth image sequences; wherein, in the single-metric three-dimensional evaluation tensor of any one group of depth image sequences

respectively corresponding to at least two sharpness evaluation metrics, a fusion weight of the single-metric evaluation vector at any one pixel position in the vector intra-source fusion is associated with the voting result; the single-source confidence at any one pixel position in the single-source confidence map of any one group of depth image sequences is associated, in the confidence intra-source fusion, with a voting value of a depth position that wins the voting; and the single-source confidence at any one pixel position in the single-source confidence map of any one group of depth image sequences is associated, in the confidence intra-source fusion, with the depth position that wins the voting.

[0174] Exemplarily, a specific implementation manner of the confidence voting described above may include: assigning, at each pixel position, one voting right for the confidence voting to each sharpness evaluation metric; determining, by using the single-metric three-dimensional evaluation tensor, a voting range set by the confidence voting for each of at least two sharpness evaluation metrics at each pixel position, wherein, for the single-metric three-dimensional evaluation tensor of any one group of depth image sequences corresponding to any one sharpness evaluation metric, the voting range set by the confidence voting for the sharpness evaluation metric at each pixel position includes: a depth position corresponding to a maximum sharpness evaluation value in the single-metric evaluation vector at the pixel position, and other depth positions in a position neighborhood of the depth position; determining, by using the single-metric confidence map, a voting value contributed by the confidence voting for each of at least two sharpness evaluation metrics at each pixel position, wherein, for the single-metric confidence map of any one group of depth image sequences corresponding to any one sharpness evaluation metric: if the single-metric confidence at any one pixel position is greater than or equal to a preset voting value determination threshold, the voting value contributed by the confidence voting for the sharpness evaluation metric at the pixel position is a first voting value; if the single-metric confidence at any one pixel position is less than the preset voting value determination threshold, the voting value contributed by the confidence voting for the sharpness evaluation metric at the pixel position is a second voting value smaller than the first voting value.

[0175] For example, in the single-source confidence map of each group of depth image sequences, the single-source confidence at each pixel position is set as the highest voting value obtained by the depth position that wins in the confidence voting; also for example, in the single-source three-dimensional evaluation tensor of each group of depth image sequences, vector intra-source fusion is performed by using, as weights, voting values contributed by respective pixel positions to the depth position winning the voting, wherein, in the single-

source evaluation vector at any one pixel position, the single-source sharpness evaluation value corresponding to any one depth position is a weighted value, using the voting values (i.e., a first voting value, a second voting value, or 0) contributed by confidence voting at the pixel position (x, y) to the winning depth position as weights, of the normalized results of the single-metric sharpness evaluation values corresponding to the same depth position in the single-metric evaluation vectors of the single-metric three-dimensional evaluation tensors respectively corresponding to at least two sharpness evaluation metrics.

[0176] Based on the specific processing manner described above, S530 in the process flow illustrated in FIG. 5 may include: performing vector multi-source fusion on the single-source evaluation vectors at each pixel position in the single-source fusion three-dimensional evaluation tensors of the at least two groups of depth image sequences, wherein the multi-source sharpness evaluation values of the multi-source evaluation vector at each pixel position are determined by vector multi-source fusion; performing confidence multi-source fusion on the single-source confidence at each pixel position in the single-source confidence maps of the at least two groups of depth image sequences, wherein the multi-source confidence at each pixel position in the multi-source confidence map is determined by confidence multi-source fusion; wherein the vector multi-source fusion and the confidence multi-source fusion are associated with each other.

[0177] Exemplarily, mutual association between the vector multi-source fusion and the confidence multi-source fusion may be represented as: determining, based on the single-source confidence maps of the at least two groups of depth image sequences, a maximum single-source confidence among at least two single-source confidences at each pixel position; performing the vector multi-source fusion and the confidence multi-source fusion based on a matching confidence level of the maximum single-source confidence at each pixel position among preset confidence levels; wherein the preset confidence levels include a high confidence level, a medium confidence level, and a low confidence level; if the matching confidence level of the maximum single-source confidence is the high confidence level or the medium confidence level, the vector multi-source fusion is configured such that, for each pixel position, the multi-source sharpness evaluation value in the multi-source evaluation vector corresponding to any one depth position, is determined by an average of single-source sharpness evaluation values corresponding to the depth position in the single-source evaluation vectors whose single-source confidences reach the matching confidence level; if the matching confidence level of the maximum single-source confidence is the low confidence level, the vector multi-source fusion is configured such that, for each pixel position, the multi-source sharpness evaluation value

corresponding to any one depth position in the multi-source evaluation vector is the single-source sharpness evaluation value corresponding to the depth position in the single-source evaluation vector having the maximum single-source confidence; and the confidence multi-source fusion is configured to determine the multi-source confidence at the pixel position in the multi-source confidence map based on the matching confidence level at each pixel position.

[0178] Exemplarily, determining the multi-source confidence at the pixel position in the multi-source confidence map based on the matching confidence level at each pixel position may include: if the matching confidence level of the maximum single-source confidence at any one pixel position is the high confidence level or the medium confidence level, verifying the matching confidence level at the pixel position based on a high-confidence-level matching number of single-source confidences at the pixel position that match the high confidence level and associated depth positions of the single-source confidences at the pixel position that match the high confidence level, wherein when verification of the matching confidence level at any one pixel position succeeds, the multi-source confidence of the pixel position is set to meet the matching confidence level at the pixel position; when verification of the matching confidence level at any one pixel position fails, the multi-source confidence at the pixel position is set to meet a next lower confidence level relative to the matching confidence level at the pixel position; if the matching confidence level of the maximum single-source confidence at any one pixel position is the low confidence level, setting the multi-source confidence at the pixel position according to an average of at least two single-source confidences at the pixel position.

[0179] Exemplarily, a specific implementation manner of the verification described above may include: if the high-confidence-level matching number at any one pixel position is greater than or equal to a confidence determination threshold, and a position distribution span of the associated depth positions is less than or equal to a preset interval threshold, determining that verification of the matching confidence level at the pixel position succeeds; if the high-confidence-level matching number at any one pixel position is less than the confidence determination threshold, and/or the position distribution span of the associated depth positions is greater than the preset interval threshold, determining that verification of the matching confidence level at the pixel position fails.

[0180] Exemplarily, the confidence determination threshold described above may be an adaptively configurable threshold, namely S530 may further include: determining, based on the single-source confidence maps of the at least two groups of depth image sequences, a single-source confidence sum of at least two single-source confidences at each pixel position;

and determining, based on the single-source confidence sum of the single-source confidences at each pixel position, a confidence determination threshold adapted to the pixel position.

[0181] In an embodiment of the present application, if an overexposed region is further introduced to optimize the extended depth-of-field image, mask fusion of the single-source highlight masks includes: determining whether each pixel position is located within the multi-source highlight mask based on the number of times the pixel position is present in the single-source highlight masks respectively corresponding to the at least two groups of depth image sequences, that is, preliminary fusion of the multi-source highlight mask; and/or, in a single-source highlight mask of the same group of depth image sequence corresponding to the maximum single-source confidence at each pixel position, if the pixel position is located outside the single-source highlight mask of the group of depth image sequence, and the multi-source confidence at the pixel position in the multi-source confidence map meets the high confidence level, prohibiting the pixel position from being placed in the multi-source highlight mask, that is, calibrating the multi-source highlight mask by using the multi-source confidence map.

[0182] FIG. 7 is an exemplary structural schematic diagram of an image generation apparatus for an extended depth-of-field image according to an embodiment of the present application. Referring to FIG. 7, an embodiment of the present application further provides an image generation apparatus for an extended depth-of-field image, and the image generation apparatus may include: an image acquisition module 710, configured to acquire at least two groups of depth image sequences obtained by imaging a target object by a photosensitive element of a movable imaging module under different imaging exposure conditions, wherein the imaging exposure condition under which the photosensitive element images the target object is associated with a light-reception state of the target object, and each group of the depth image sequences includes a plurality of images captured by the photosensitive element with the same imaging exposure condition when the movable imaging module is located at a plurality of depth positions in the depth direction, respectively; and an extended depth-of-field processing module 730, configured to generate the extended depth-of-field image of the target object based on the at least two groups of depth image sequences. In the embodiments of the present application, for the specific operation principle of the extended depth-of-field processing module 730, reference may be made to the description of S430 above, and repeated description is omitted here.

[0183] As can be seen from the above, the image generation apparatus in an embodiment of the present application may generate an extended depth-of-field image based on at least two

groups of depth image sequences acquired under different imaging exposure conditions. When structural features of the target object are prone to localized light-reception abnormality, because positions and degrees of the localized light-reception abnormality differ under different imaging exposure conditions, the probability that the same localized light-reception abnormality is present in all depth image sequences acquired under different imaging exposure conditions is extremely low, that is, local image defects present in at least two groups of depth image sequences acquired under different imaging exposure conditions are not completely identical. Therefore, by generating an extended depth-of-field image based on at least two groups of depth image sequences acquired under different imaging exposure conditions, influence of local image defects caused by localized light-reception abnormality on the extended depth-of-field image can be reduced through mutual supplementation of different object information presented by the target object in the at least two groups of depth image sequences, thereby helping improve the accuracy of three-dimensional object information presented by the extended depth-of-field image.

[0184] An embodiment of the present application further provides a computer program product including computer-executable instructions, wherein the computer-executable instructions, when executed by a processor, implement the image generation method according to the foregoing embodiment.

[0185] An embodiment of the present application further provides a non-transitory computer-readable storage medium storing instructions, wherein the instructions, when executed by a processor, cause the processor to perform the image generation method according to the foregoing embodiment.

[0186] The above descriptions are merely some embodiments of the present application and are not intended to limit the present application. Any modification, equivalent replacement, improvement, and the like made within the spirit and principle of the present application shall fall within the protection scope of the present application.

CLAIMS

1. An image acquisition apparatus for extended depth-of-field imaging, comprising:

a movable imaging module, comprising an optical lens and a photosensitive element, wherein a movement direction of the movable imaging module is parallel to a depth direction of the optical lens, and the photosensitive element is configured to image a target object within a field-of-view range of the optical lens to generate a depth image sequence; and

an adjustable light source module, comprising a plurality of illumination sources, wherein the plurality of illumination sources are configured to selectively illuminate the target object according to any one of at least two illumination modes, a light-reception state of the target object is used for determining an imaging exposure condition under which the photosensitive element images the target object, and enabled states of the plurality of illumination sources under different illumination modes are different, so that the light-reception state of the target object differs under the different illumination modes.

2. The image acquisition apparatus according to claim 1, wherein the plurality of illumination sources comprise:

a coaxial illumination source group, comprising a plurality of first illumination sources, wherein each of the plurality of first illumination sources is configured to generate a first single-source light beam parallel to the depth direction, the first single-source light beams respectively generated by the plurality of first illumination sources cover different fan-shaped phase intervals of a circular illuminated region in a field-of-view region of the optical lens, and the circular illuminated region is coaxial with an optical axis of the optical lens; and

an annular illumination source group, comprising a plurality of second illumination sources, wherein each of the plurality of second illumination sources is configured to generate a second single-source light beam inclined relative to the depth direction, the second single-source light beams respectively generated by the plurality of second illumination sources cover different annular-sector phase intervals of an annular illuminated region in the field-of-view region, and the annular illuminated region coaxially surrounds the circular illuminated region.

3. The image acquisition apparatus according to claim 2, wherein:

the light-reception state of the target object comprises a light-reception orientation of the

target object, wherein:

the light-reception orientation is associated with a deployment position, in the coaxial illumination source group, of at least one enabled first illumination source in the coaxial illumination source group, and/or

the light-reception orientation is associated with a deployment position, in the coaxial illumination source group, of at least one enabled second illumination source in the annular illumination source group;

and/or the light-reception state of the target object comprises light-reception intensity, wherein:

the light-reception intensity is associated with an intensity of the first single-source light beam generated by at least one enabled first illumination source in the coaxial illumination source group, and/or

the light-reception intensity is associated with an intensity of the second single-source light beam generated by at least one enabled second illumination source in the annular illumination source group, and/or

the light-reception intensity caused by the coaxial illumination source group is greater than the light-reception intensity caused by the annular illumination source group;

and/or the imaging exposure condition under which the photosensitive element images the target object is further associated with an exposure duration of the photosensitive element.

4. The image acquisition apparatus according to claim 2, wherein the illumination modes comprise at least two of a coaxial epi-illumination mode, a coaxial side-illumination mode, an annular epi-illumination mode, and an annular side-illumination mode:

in the coaxial epi-illumination mode, all of the plurality of first illumination sources in the coaxial illumination source group are enabled, and all of the plurality of second illumination sources in the annular illumination source group are disabled;

in the coaxial side-illumination mode, a part of the plurality of first illumination sources in the coaxial illumination source group is enabled, another part of the plurality of first illumination sources in the coaxial illumination source group is disabled, and all of the plurality of second illumination sources in the annular illumination source group are disabled;

in the annular epi-illumination mode, all of the plurality of second illumination sources in the annular illumination source group are enabled, and all of the plurality of first illumination sources in the coaxial illumination source group are disabled; and

in the annular side-illumination mode, a part of the plurality of second illumination sources in the annular illumination source group is enabled, another part of the plurality of second illumination sources in the annular illumination source group is disabled, and all of the plurality of first illumination sources in the coaxial illumination source group are disabled.

5. The image acquisition apparatus according to claim 4, wherein:

the coaxial epi-illumination mode comprises a normal-exposure coaxial epi-illumination mode and a low-exposure coaxial epi-illumination mode, wherein,

in the normal-exposure coaxial epi-illumination mode, the first single-source light beams respectively generated by the plurality of first illumination sources that are all enabled have a first illumination intensity, and/or exposure duration of the photosensitive element is configured as a first exposure duration,

in the low-exposure coaxial epi-illumination mode, the first single-source light beams respectively generated by the plurality of first illumination sources that are all enabled have a second illumination intensity less than the first illumination intensity, and/or the exposure duration of the photosensitive element is configured as a second exposure duration shorter than the first exposure duration;

and/or the annular epi-illumination mode comprises a normal-exposure annular epi-illumination mode and a low-exposure annular epi-illumination mode, wherein,

in the normal-exposure annular epi-illumination mode, the second single-source light beams respectively generated by the plurality of second illumination sources that are all enabled have a third illumination intensity, and/or the exposure duration of the photosensitive element is configured as a third exposure duration,

in the low-exposure annular epi-illumination mode, the second single-source light beams respectively generated by the plurality of second illumination sources that are all enabled have a fourth illumination intensity less than the third illumination intensity, and/or the exposure duration of the photosensitive element is configured as a fourth exposure duration shorter than the third exposure duration;

and/or the coaxial side-illumination mode comprises at least two directional coaxial side-illumination modes having different side-illumination directions, and the first illumination sources respectively enabled in any two of the directional coaxial side-illumination modes are not completely identical;

and/or the annular side-illumination mode comprises at least two directional annular side-illumination modes having different side-illumination directions, and the second illumination sources respectively enabled in any two of the directional annular side-illumination modes are not completely identical.

6. The image acquisition apparatus according to claim 5, wherein the adjustable light source module is configured to:

when a magnification of the optical lens is greater than or equal to a preset magnification threshold, respectively illuminate the target object in the normal-exposure coaxial epi-illumination mode, the low-exposure coaxial epi-illumination mode, and at least two directional coaxial side-illumination modes;

and/or,

when the magnification of the optical lens is less than the preset magnification threshold, respectively illuminate the target object in the normal-exposure coaxial epi-illumination mode, the normal-exposure annular epi-illumination mode, the low-exposure annular epi-illumination mode, and at least two directional annular side-illumination modes.

7. The image acquisition apparatus according to claim 2, wherein:

the plurality of first illumination sources are located on an axial inner side of the optical lens, and the first single-source light beams pass through the optical lens and are emitted toward the circular illuminated region; and/or,

the plurality of second illumination sources are located on a radial outer side of the optical lens, and the second single-source light beams are obliquely emitted from the radial outer side of the optical lens toward the annular illuminated region.

8. The image acquisition apparatus according to claim 7, wherein the movable imaging

module comprises a module housing, the photosensitive element is located in the module housing, and the optical lens is mounted at an end opening of the module housing facing the target object, wherein:

the plurality of first illumination sources are fixedly mounted in the module housing, and the first single-source light beams are constrained by a beam splitter in the module housing to be parallel to the depth direction; and/or,

the plurality of second illumination sources are fixedly mounted on an outer peripheral wall of the module housing near the end opening.

9. The image acquisition apparatus according to claim 2, wherein:

each of the plurality of first illumination sources comprises a first light-emitting element group and a first light homogenizing plate covering the first light-emitting element group, wherein a beam cross section of the first single-source light beam generated by the first illumination source is shaped by the first light homogenizing plate, the first light homogenizing plate is fan-shaped, and the first light homogenizing plates of the respective first illumination sources are spliced into a circle; and/or,

each of the plurality of second illumination sources comprises a second light-emitting element group and a second light homogenizing plate covering the second light-emitting element group, wherein a beam cross section of the second single-source light beam generated by the second illumination source is shaped by the second light homogenizing plate, the second light homogenizing plate is in an annular-sector shape, and the second light homogenizing plates of the respective second illumination sources are spliced into a ring.

10. An extended depth-of-field imaging device, comprising:

a movable imaging module, comprising an optical lens and a photosensitive element, wherein the photosensitive element is configured to image a target object within a field-of-view range of the optical lens, and a movement direction of the movable imaging module is parallel to a depth direction of the optical lens;

an adjustable light source module, configured to adjust a light-reception state of the target object; and

an image processing module, configured to generate an extended depth-of-field image of the target object based on at least two groups of depth image sequences, wherein the at least two groups of depth image sequences are obtained by imaging the target object by the photosensitive element under different imaging exposure conditions, wherein the imaging exposure condition under which the photosensitive element images the target object is associated with the light-reception state of the target object, and each group of the depth image sequences comprises a plurality of images captured by the photosensitive element with a same imaging exposure condition when the movable imaging module is located at a plurality of depth positions in the depth direction, respectively.

11. The extended depth-of-field imaging device according to claim 10, wherein the adjustable light source module comprises a plurality of illumination sources configured to selectively illuminate the target object according to any one of at least two illumination modes, the light-reception state of the target object is used for determining the imaging exposure condition under which the photosensitive element images the target object, enabled states of the plurality of illumination sources under different illumination modes are different, so that the light-reception state of the target object differs under the different illumination modes, and the plurality of illumination sources comprise:

a coaxial illumination source group, comprising a plurality of first illumination sources, wherein each of the plurality of first illumination sources is configured to generate a first single-source light beam parallel to the depth direction, the first single-source light beams respectively generated by the plurality of first illumination sources cover different fan-shaped phase intervals of a circular illuminated region in a field-of-view region of the optical lens, and the circular illuminated region is coaxial with an optical axis of the optical lens; and

an annular illumination source group, comprising a plurality of second illumination sources, wherein each of the plurality of second illumination sources is configured to generate a second single-source light beam inclined relative to the depth direction, the second single-source light beams respectively generated by the plurality of second illumination sources cover different annular-sector phase intervals of an annular illuminated region in the field-of-view region, and the annular illuminated region coaxially surrounds the circular illuminated region.

12. The extended depth-of-field imaging device according to claim 10, wherein the image

processing module is configured to:

based on the at least two groups of depth image sequences, for each group of depth image sequences and for each pixel position, determine, under the imaging exposure condition of the group, single-source sharpness evaluation values respectively corresponding to the plurality of depth positions, and a single-source confidence of the single-source sharpness evaluation values;

for each pixel position, fuse, under the imaging exposure conditions of the at least two groups of depth image sequences, the single-source sharpness evaluation values and the single-source confidence into multi-source sharpness evaluation values at the pixel position respectively corresponding to the plurality of depth positions and a multi-source confidence of the multi-source sharpness evaluation values, wherein a fusion weight of the single-source sharpness evaluation values at each pixel position under the imaging exposure condition of each group of depth image sequences is associated with the single-source confidence at the pixel position under the imaging exposure condition of the group of depth image sequences; and

generate the extended depth-of-field image based on the multi-source sharpness evaluation values respectively corresponding to the plurality of depth positions and the multi-source confidence.

13. The extended depth-of-field imaging device according to claim 12, wherein the image processing module is further configured to:

based on the at least two groups of depth image sequences, for each group of depth image sequences and for each pixel position, determine, by respectively using at least two sharpness evaluation metrics and under the imaging exposure condition of the group, single-metric sharpness evaluation values corresponding to the plurality of depth positions, and a single-metric confidence of the single-metric sharpness evaluation values; and

for each pixel position under the imaging exposure condition of each group of depth image sequences, fuse the single-metric sharpness evaluation values and the single-metric confidence, respectively determined by using the at least two sharpness evaluation metrics, into the single-source sharpness evaluation values at the pixel position corresponding to the plurality of depth positions under the imaging exposure condition of the group of depth image sequences and the single-source confidence of the single-source sharpness evaluation values, wherein a fusion

weight of the single-metric sharpness evaluation values at each pixel position under the imaging exposure condition of each group of depth image sequences is associated with the single-metric confidence at the pixel position under the imaging exposure condition of the group of depth image sequences.

14. The extended depth-of-field imaging device according to claim 12, wherein the image processing module is further configured to:

determine, based on the at least two groups of depth image sequences, an overexposed region of the target object under the imaging exposure condition of each group of depth image sequences; and

wherein fusion of the single-source sharpness evaluation values and/or the single-source confidence at each pixel position under the imaging exposure condition of each group of depth image sequences is associated with a positional relationship indicating whether the pixel position is located within the overexposed region under the imaging exposure condition of the group of depth image sequences.

15. The extended depth-of-field imaging device according to claim 12, wherein the image processing module is configured to:

determine, based on the at least two groups of depth image sequences, a single-source three-dimensional evaluation tensor and a single-source confidence map of each group of depth image sequences; wherein the single-source three-dimensional evaluation tensor of any one group of depth image sequences comprises single-source evaluation vectors at respective pixel positions, the single-source evaluation vector at each pixel position comprising a plurality of single-source sharpness evaluation values at the pixel position respectively corresponding to the plurality of depth positions; and the single-source confidence map of any one group of depth image sequences comprises single-source confidences of the single-source evaluation vectors at respective pixel positions in the single-source three-dimensional evaluation tensor of the group of depth image sequences;

determine a multi-source three-dimensional evaluation tensor and a multi-source confidence map based on the single-source three-dimensional evaluation tensors and the single-source confidence maps of the at least two groups of depth image sequences; wherein the multi-

source three-dimensional evaluation tensor comprises multi-source evaluation vectors at respective pixel positions, the multi-source evaluation vector at each pixel position comprising a plurality of multi-source sharpness evaluation values at the pixel position respectively corresponding to the plurality of depth positions; and the multi-source confidence map comprises multi-source confidences of the multi-source evaluation vectors at respective pixel positions in the multi-source three-dimensional evaluation tensor; and

generate the extended depth-of-field image based on the multi-source three-dimensional evaluation tensor and the multi-source confidence map.

16. The extended depth-of-field imaging device according to claim 15, wherein the image processing module is configured to:

based on the at least two groups of depth image sequences, determine, for each group of depth image sequences, a single-metric three-dimensional evaluation tensor and a single-metric confidence map corresponding to each of at least two sharpness evaluation metrics; wherein the single-metric three-dimensional evaluation tensor of any one group of depth image sequences comprises single-metric evaluation vectors at respective pixel positions, the single-metric evaluation vector at each pixel position comprising a plurality of single-metric sharpness evaluation values at the pixel position respectively corresponding to the plurality of depth positions; and the single-metric confidence map of any one group of depth image sequences comprises the single-metric confidences of the single-metric evaluation vectors at respective pixel positions in the single-metric three-dimensional evaluation tensor of the group of depth image sequences; and

determine the single-source three-dimensional evaluation tensor and the single-source confidence map of each group of depth image sequences based on fusion of the single-metric three-dimensional evaluation tensors and the single-metric confidence maps corresponding to the at least two sharpness evaluation metrics for the group of depth image sequences.

17. The extended depth-of-field imaging device according to claim 16, wherein the image processing module is configured to:

based on the at least two groups of depth image sequences, determine, for each group of depth image sequences, the single-metric three-dimensional evaluation tensors corresponding

to at least two sharpness evaluation metrics, wherein, in the single-metric three-dimensional evaluation tensor of any one group of depth image sequences corresponding to any one sharpness evaluation metric, the single-metric sharpness evaluation values of the single-metric evaluation vector at each pixel position are obtained by evaluating the group of depth image sequences by using the sharpness evaluation metric; and

determine, based on the single-metric three-dimensional evaluation tensors of each group of depth image sequences respectively corresponding to the at least two sharpness evaluation metrics, the single-metric confidence maps of each group of depth image sequences respectively corresponding to the at least two sharpness evaluation metrics, wherein, in the single-metric confidence map of any one group of depth image sequences corresponding to any one sharpness evaluation metric, the single-metric confidence at each pixel position is determined based on a vector feature of the single-metric evaluation vector at the pixel position.

18. The extended depth-of-field imaging device according to claim 17, wherein the image processing module is further configured to:

generate, based on the at least two groups of depth image sequences, a single-source highlight mask corresponding to each group of depth image sequences, wherein the single-source highlight mask corresponding to any one group of depth image sequences represents an overexposed region of the target object under the imaging exposure condition of the group of depth image sequences;

calibrate the single-source highlight mask by using the single-metric confidence maps of each group of depth image sequences respectively corresponding to the at least two sharpness evaluation metrics;

generate a multi-source highlight mask based on mask fusion of the single-source highlight masks respectively corresponding to the at least two groups of depth image sequences, wherein the multi-source highlight mask is used to represent overexposed regions of the target object under different imaging exposure conditions, and the multi-source highlight mask is used to optimize the extended depth-of-field image generated based on multi-source fusion; and

wherein calibrating the single-source highlight mask of any one group of depth image sequences comprises: if any one pixel position is present in the single-source highlight mask, and the single-metric confidences at the pixel position in the single-metric confidence maps of the at least two sharpness evaluation metrics are all greater than a preset single-metric high-

confidence threshold, removing the pixel position from the single-source highlight mask.

19. The extended depth-of-field imaging device according to claim 17, wherein the image processing module is configured to:

detect, in the single-metric three-dimensional evaluation tensor of each group of depth image sequences corresponding to each sharpness evaluation metric, evaluation value distribution characteristic of the single-metric evaluation vector at each pixel position, wherein the vector feature comprises the evaluation value distribution characteristic;

determine, based on a preset reference distribution characteristic and the evaluation value distribution characteristic detected in the single-metric three-dimensional evaluation tensor of each group of depth image sequences corresponding to each sharpness evaluation metric, the single-metric confidence map of each group of depth image sequences respectively corresponding to the at least two sharpness evaluation metrics;

wherein, in the single-metric three-dimensional evaluation tensor of any one group of depth image sequences corresponding to any one sharpness evaluation metric:

if the evaluation value distribution characteristic at any one pixel position matches the reference distribution characteristic, the single-metric confidence at the pixel position in the single-metric confidence map of the group of depth image sequences corresponding to the sharpness evaluation metric is set to a value representing high confidence; and

if the evaluation value distribution characteristic at any one pixel position fails to match the reference distribution characteristic, the single-metric confidence at the pixel position in the single-metric confidence map of the group of depth image sequences corresponding to the sharpness evaluation metric is set to a value representing low confidence.

20. The extended depth-of-field imaging device according to claim 19, wherein the image processing module is configured to:

determine a value peak in the single-metric evaluation vector at each pixel position in the single-metric three-dimensional evaluation tensor of each group of depth image sequences corresponding to each sharpness evaluation metric, wherein the value peak is a set of the single-metric sharpness evaluation values in the single-metric evaluation vector that are greater than or equal to a corresponding peak determination threshold, and the peak determination threshold

corresponding to the single-metric evaluation vector is determined according to a maximum value, a minimum value, and an average value of the single-metric sharpness evaluation values in the single-metric evaluation vector;

wherein:

the reference distribution characteristic comprises at least one of: the single-metric evaluation vector having a unimodal distribution characteristic, the single-metric evaluation vector having a fluctuation amplitude greater than a preset peak-valley fluctuation threshold, or each value peak of the single-metric evaluation vector having a peak distribution width in the depth direction within a preset peak width interval;

and/or,

the image processing module is further configured to: if any one single-metric evaluation vector has two adjacent peaks with an inter-peak interval located within a preset interval range, perform peak clipping suppression on one value peak, among the two adjacent peaks, that is located within a position region subject to light shielding, so that a top of the value peak is suppressed to the corresponding peak determination threshold.

21. The extended depth-of-field imaging device according to claim 17, wherein the image processing module is configured to:

perform confidence voting respectively on the plurality of depth positions in the depth direction at a granularity of pixel positions, based on the single-metric three-dimensional evaluation tensors and the single-metric confidence maps of each group of depth image sequences respectively corresponding to at least two sharpness evaluation metrics;

determine, based on a voting result of the confidence voting, the single-source three-dimensional evaluation tensor and the single-source confidence map of each group of depth image sequences, so as to perform vector intra-source fusion on the single-metric evaluation vectors in the single-metric three-dimensional evaluation tensors corresponding to at least two sharpness evaluation metrics for each pixel position, and to perform confidence intra-source fusion on the single-metric confidences at pixel positions in the single-metric confidence maps corresponding to at least two sharpness evaluation metrics; and

wherein, in the single-metric three-dimensional evaluation tensors of any one group of depth image sequences respectively corresponding to at least two sharpness evaluation metrics,

a fusion weight of the single-metric evaluation vector at any one pixel position in the vector intra-source fusion is associated with the voting result; the single-source confidence at any one pixel position in the single-source confidence map of any one group of depth image sequences is associated, in the confidence intra-source fusion, with a voting value of a depth position that wins the voting; and the single-source confidence at any one pixel position in the single-source confidence map of any one group of depth image sequences is bound, in the confidence intra-source fusion, to the depth position that wins the voting.

22. The extended depth-of-field imaging device according to claim 21, wherein the image processing module is configured to:

assign, at each pixel position, one voting right for the confidence voting to each sharpness evaluation metric;

determine, by using the single-metric three-dimensional evaluation tensor, a voting range set by the confidence voting for each of at least two sharpness evaluation metrics at each pixel position, wherein, for the single-metric three-dimensional evaluation tensor of any one group of depth image sequences corresponding to any one sharpness evaluation metric, the voting range set by the confidence voting for the sharpness evaluation metric at each pixel position comprises: a depth position corresponding to a maximum sharpness evaluation value in the single-metric evaluation vector of the pixel position, and other depth positions in a position neighborhood of the depth position;

determine, by using the single-metric confidence map, a voting value contributed by the confidence voting for each of at least two sharpness evaluation metrics at each pixel position, wherein, for the single-metric confidence map of any one group of depth image sequences corresponding to any one sharpness evaluation metric: if the single-metric confidence at any one pixel position is greater than or equal to a preset voting value determination threshold, the voting value contributed by the confidence voting for the sharpness evaluation metric at the pixel position is a first voting value; if the single-metric confidence at any one pixel position is less than the preset voting value determination threshold, the voting value contributed by the confidence voting for the sharpness evaluation metric at the pixel position is a second voting value smaller than the first voting value.

23. The extended depth-of-field imaging device according to claim 21, wherein the image

processing module is configured to:

set, in the single-source confidence map of each group of depth image sequences, the single-source confidence at each pixel position as a highest voting value obtained by a depth position winning in the confidence voting; and

perform, in the single-source three-dimensional evaluation tensor of each group of depth image sequences, the vector intra-source fusion by using, as a weight, a voting value contributed by the pixel position to the depth position winning the voting, wherein, in the single-source evaluation vector at any one pixel position, the single-source sharpness evaluation value corresponding to any one depth position is a weighted value, the voting values contributed by the confidence voting at the pixel position to the depth position winning the voting as weights, of normalized results of the single-metric sharpness evaluation values corresponding to the depth position in the single-metric evaluation vectors of the single-metric three-dimensional evaluation tensors respectively corresponding to at least two sharpness evaluation metrics.

24. The extended depth-of-field imaging device according to claim 21, wherein the image processing module is configured to:

determine, based on the single-source confidence maps of the at least two groups of depth image sequences, a maximum single-source confidence among at least two single-source confidences at each pixel position;

based on a matching confidence level of the maximum single-source confidence at each pixel position among preset confidence levels, perform vector multi-source fusion on the single-source evaluation vectors at the pixel position in the single-source fusion three-dimensional evaluation tensors of the at least two groups of depth image sequences, and perform confidence multi-source fusion on the single-source confidences at the pixel position in the single-source confidence maps of the at least two groups of depth image sequences;

wherein the preset confidence levels comprise a high confidence level, a medium confidence level, and a low confidence level, and:

if the matching confidence level of the maximum single-source confidence is the high confidence level or the medium confidence level, the vector multi-source fusion is configured such that, for each pixel position, the multi-source sharpness evaluation value corresponding

to any one depth position in the multi-source evaluation vector, is determined by an average of single-source sharpness evaluation values corresponding to the depth position in the single-source evaluation vectors whose single-source confidences reach the matching confidence level;

if the matching confidence level of the maximum single-source confidence is the low confidence level, the vector multi-source fusion is configured such that, for each pixel position, the multi-source sharpness evaluation value corresponding to any one depth position in the multi-source evaluation vector is the single-source sharpness evaluation value corresponding to the depth position in the single-source evaluation vector having the maximum single-source confidence; and

the confidence multi-source fusion is configured to determine the multi-source confidence at the pixel position in the multi-source confidence map based on the matching confidence level at each pixel position.

25. The extended depth-of-field imaging device according to claim 24, wherein the image processing module is further configured to:

generate, based on the at least two groups of depth image sequences, a single-source highlight mask corresponding to each group of depth image sequences, wherein the single-source highlight mask corresponding to any one group of depth image sequences represents an overexposed region of the target object under the imaging exposure condition of the group of depth image sequences;

generate a multi-source highlight mask based on mask fusion of the single-source highlight masks respectively corresponding to the at least two groups of depth image sequences, wherein the multi-source highlight mask represents overexposed regions of the target object under different imaging exposure conditions, and the multi-source highlight mask is used to optimize the extended depth-of-field image generated based on the multi-source fusion;

wherein the mask fusion of the single-source highlight masks comprises:

determining whether each pixel position is located within the multi-source highlight mask based on a number of times the pixel position is present in the single-source highlight masks respectively corresponding to the at least two groups of depth image sequences;

and/or,

in a single-source highlight mask of a group of depth image sequences corresponding to the maximum single-source confidence at each pixel position, if the pixel position is located outside the single-source highlight mask of the group of depth image sequences and the multi-source confidence at the pixel position in the multi-source confidence map meets the high confidence level, prohibiting the pixel position from being placed in the multi-source highlight mask.

26. The extended depth-of-field imaging device according to claim 24, wherein the image processing module is configured to:

if the matching confidence level of the maximum single-source confidence at any one pixel position is the high confidence level or the medium confidence level, verify the matching confidence level at the pixel position based on a high-confidence-level matching number of single-source confidences at the pixel position that match the high confidence level and associated depth positions of the single-source confidences at the pixel position that match the high confidence level, wherein when verification of the matching confidence level at any one pixel position succeeds, the multi-source confidence at the pixel position is set to meet the matching confidence level at the pixel position, and when verification of the matching confidence level at any one pixel position fails, the multi-source confidence at the pixel position is set to meet a next lower confidence level relative to the matching confidence level at the pixel position; and

if the matching confidence level of the maximum single-source confidence at any one pixel position is the low confidence level, set the multi-source confidence at the pixel position according to an average of at least two single-source confidences at the pixel position.

27. The extended depth-of-field imaging device according to claim 26, wherein the image processing module is configured to:

if the high-confidence-level matching number at any one pixel position is greater than or equal to a confidence determination threshold and a position distribution span of the associated depth positions is less than or equal to a preset interval threshold, determine that verification of the matching confidence level at the pixel position succeeds; and

if the high-confidence-level matching number at any one pixel position is less than the

confidence determination threshold and/or the position distribution span of the associated depth positions is greater than the preset interval threshold, determine that verification of the matching confidence level at the pixel position fails.

28. The extended depth-of-field imaging device according to claim 27, wherein the image processing module is configured to:

determine, based on the single-source confidence maps of the at least two groups of depth image sequences, a single-source confidence sum of at least two single-source confidences at each pixel position; and

determine the confidence determination threshold adapted to the pixel position based on the single-source confidence sum of the single-source confidences at the pixel position.

29. An image generation method for an extended depth-of-field image, comprising:

acquiring at least two groups of depth image sequences obtained by imaging a target object by a photosensitive element of a movable imaging module under different imaging exposure conditions, wherein the imaging exposure condition under which the photosensitive element images the target object is associated with a light-reception state of the target object, and each group of the depth image sequences comprises a plurality of images captured by the photosensitive element with a same imaging exposure condition when the movable imaging module is located at a plurality of depth positions, respectively; and

generating the extended depth-of-field image of the target object based on the at least two groups of depth image sequences.

30. An image generation apparatus for an extended depth-of-field image, comprising:

an image acquisition module, configured to acquire at least two groups of depth image sequences obtained by imaging a target object by a photosensitive element of a movable imaging module under different imaging exposure conditions, wherein the imaging exposure condition under which the photosensitive element images the target object is associated with a light-reception state of the target object, and each group of the depth image sequences comprises a plurality of images captured by the photosensitive element with a same imaging

exposure condition when the movable imaging module is located at a plurality of depth positions, respectively;

an extended depth-of-field processing module, configured to generate the extended depth-of-field image of the target object based on at least two groups of the depth image sequences.

31. A non-transitory computer-readable storage medium, wherein the non-transitory computer-readable storage medium stores instructions which, when executed by a processor, cause the processor to perform the image generation method according to claim 29.