



(19)

Europäisches Patentamt
European Patent Office
Office européen des brevets



EP 2 793 187 A1

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication:
22.10.2014 Bulletin 2014/43

(51) Int Cl.:
G06T 3/40 (2006.01)

(21) Application number: **14164965.7**(22) Date of filing: **16.04.2014**

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB
GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO
PL PT RO RS SE SI SK SM TR
Designated Extension States:
BA ME

(30) Priority: 16.04.2013 US 201361812561 P
23.09.2013 US 201314034543

(71) Applicants:

- Disney Enterprises, Inc.
Burbank, CA 91521 (US)
- Eidgenössische Technische Hochschule Zürich
8092 Zürich (CH)

(72) Inventors:

- Richardt, Christian
Burbank, California 91521 (US)
- Zimmer, Henning
Burbank, California 91521 (US)
- Pritch, Yael
Burbank, California 91521 (US)
- Sorkine Hornung, Alexander
Burbank, California 91521 (US)

(74) Representative: Potter, Julian Mark
WP Thompson
55 Drury Lane
London WC2B 5SQ (GB)

(54) **Stereoscopic panoramas**

(57) Systems and methods to generate stereoscopic panoramas obtain images based on captured images. The obtained images may be processed and/or preprocessed, for example to compensate for perspective distortion caused by the non-ideal camera orientation during capturing, to reduce vertical parallax, to align adjacent images, and/or to reduce rotational and/or positional drift

between adjacent images. The obtained images may be used for interpolating virtual in-between images on the fly to reduce visible artifacts in the resulting panorama. Obtained and/or interpolated images (or image fragments) may be stitched together to form a stereoscopic panorama.

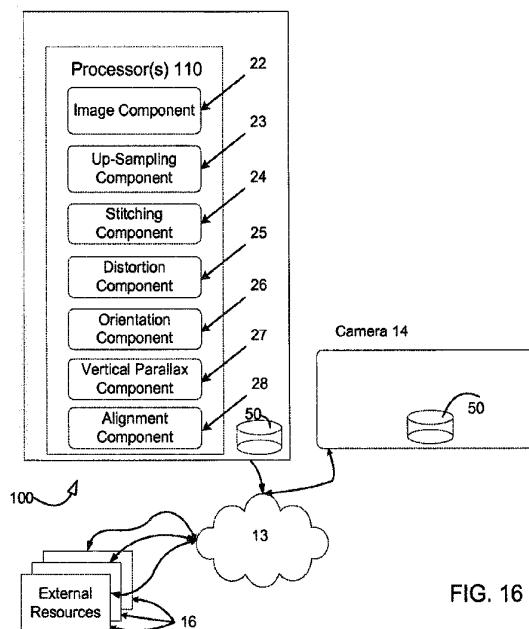


FIG. 16

Description**CROSS-REFERENCE TO RELATED APPLICATIONS**

5 [0001] The present Application claims priority of U.S. Provisional Patent Application No. 61/812,561, filed April 16, 2013, entitled "Stereoscopic Panoramas," which is expressly incorporated by reference herein.

FIELD OF THE DISCLOSURE

10 [0002] The present invention relates generally to generating panoramas, in particular stereoscopic panoramas, from (digitally and electronically) captured images.

BACKGROUND

15 [0003] Conventionally, obtaining artifact-free (or artifact-reduced), seamless panoramas requires stitching together densely (angular) sampled and/or captured images, e.g. significantly more than one image per degree. At high-definition resolutions, capturing thousands of images may be impractical and/or (computationally) costly.

SUMMARY

20 [0004] One aspect of the disclosure relates to a system configured to generate stereoscopic panoramas. One aspect of the disclosure relates to a method for generating stereoscopic panoramas. The system may include one or more cameras, a support structure, electronic storage, one or more processors configured to execute computer program components, and/or other components. The computer program components may include an image component, an up-sampling component, a stitching component, a distortion component, an orientation component, a vertical parallax component, an alignment component, and/or other components. As used herein, the term camera may be interpreted as any sensor, image sensor, still photography camera, video camera, and/or other image-producing device. The order of the operations of the computer program components is not intended to be limited by examples and embodiments described in this disclosure. Any order is contemplated within the scope of this disclosure.

25 [0005] The system may be configured to reduce image distortion and/or visual artifacts. The system may be configured to interpolate and/or synthesize images (or image portions/fragments) between captured images, in particular adjacent captured images. By virtue of such images, the angular density of the images available for stitching is increased. To accomplish image interpolation and/or synthesis, the optical flow between neighboring and/or adjacent pairs of images may need to be computed and/or determined. For the first image of such a pair, the optical flow is defined as a vector field where an individual pixel has an associated vector that points to the corresponding location of the individual pixel in the second image. Interpolation of in-between images may be accomplished by shifting pixels in the first and/or second image by a certain fraction of the flow vectors, and/or aggregating the shifted pixels to form an interpolated image.

30 [0006] Pre-computing a large number of in-between images that would be needed for artifact-free stitching may be impractical due to prohibitively large memory consumption. As used herein, the term "artifact-free" may be interpreted as "a reduction of artifacts that visually and/or visibly improves the result." By pre-computing and/or storing individual flow fields, interpolation and/or synthesis may be limited to (e.g. merely require) the pixels that are actually needed and/or used in the resulting image during panorama generation. In some implementations, the system may be configured to interactively generate a stereoscopic panorama having a variable interaxial camera distance or baseline, and/or an adjustable vergence point. In some implementations, users may modify the viewpoint (and/or other characteristics of the panorama) and/or navigate within the panorama interactively. As used herein, the term "interactively" may be interpreted as a real-time response to user input. For example, users may be able to zoom in to reveal more details or more features of a panorama.

35 [0007] The image component may be configured to obtain captured images and/or images based on captured images. The captured images may include digital and/or electronic images. Captured images may be processed and/or pre-processed prior to obtainment by the image component. Some or all of the captured images may form a set of adjacent and/or neighboring images. For example, the set may form a sequence. For example, adjacent images may partially overlap and/or include one or more of the same objects, though viewed and/or sampled from fields of view having a sufficient degree of overlap, and/or slightly different viewpoints and/or angles. As used herein, the term "adjacent" may be interpreted in a temporal and/or spatial sense. The term "adjacent" may be used interchangeably with the term "neighboring." In some implementations, a set of images may be captured using a circular camera trajectory. In some implementations, a set of images may be captured by a set of cameras, e.g. an array of cameras, which may capture images sequentially and/or simultaneously.

40 [0008] The up-sampling component may be configured to determine interpolated images between adjacent images.

In some implementations, the up-sampling component may be configured to determine interpolated image-fragments, image-portions, and/or image-strips between adjacent images. In some implementations, the up-sampling component may be configured to determine flow fields (and/or flow vectors) between related pixels in adjacent images. Flow fields may be used to compensate for transformations that images have undergone and/or process images that have undergone 5 transformations, including but not limited to a transformation of the pixel grid of an obtained image.

[0009] The stitching component may be configured to determine a stereoscopic panorama based on obtained images and interpolated images. The stitching component may be configured to blend pixels and/or image-strips from multiple images. Blending may be based on flow fields as determined by the up-sampling component.

[0010] The distortion component may be configured to determine images by compensating for distortion, including, 10 but not limited to, perspective distortion. By way of non-limiting example, the shape of the lens of the camera used to capture an image may cause perspective distortion.

[0011] The orientation component may be configured to determine images by compensating for a non-ideal capture setup. An ideal capture setup may include, by way of non-limiting example, a circular camera trajectory, a parallel principal (camera) axis, a viewing-direction that is perpendicular to the camera trajectory, and/or other capture conditions. In 15 some implementations, the orientation component may be configured to compensate for one or both of a non-circular camera trajectory during image capture and/or a non-parallel principal axis during image capture.

[0012] The vertical parallax component may be configured to determine images by re-projecting images onto a cylindrical surface and/or reducing vertical parallax.

[0013] The alignment component may be configured to determine images by projecting images into a planar and/or 20 two-dimensional surface. The alignment component may be configured to reduce rotational and/or positional drift between adjacent images. Operation of the alignment component may be based on structure-from-motion algorithms. Structure-from-motion algorithms may be used to estimate the three-dimensional position, location, and/or movement of the camera. In some implementations, structure-from-motion algorithms may use global bundle adjustment.

[0014] The system may include electronic storage. The electronic storage may comprise non-transitory storage media 25 that electronically stores information. The electronic storage may be configured to store captured images, obtained images, pre-processed images, processed images, *etc.*

[0015] These and other features, and characteristics of the present technology, as well as the methods of operation and functions of the related elements of structure and the combination of parts and economies of manufacture, will 30 become more apparent upon consideration of the following description and the appended claims with reference to the accompanying drawings, all of which form a part of this specification, wherein like reference numerals designate corresponding parts in the various figures. It is to be expressly understood, however, that the drawings are for the purpose of illustration and description only and are not intended as a definition of the limits of the invention. As used in the specification and in the claims, the singular form of "a", "an", and "the" include plural referents unless the context clearly dictates otherwise.

35

BRIEF DESCRIPTION OF THE DRAWINGS

[0016]

40 FIG. 1 illustrates a stereoscopic panorama. FIGs 1A-1H illustrate details or close-ups of the panorama of FIG. 1 for known methods of panorama production and methods in accordance with one or more implementations.

FIG. 2 illustrates examples of a hand-held capture setup and a mounted capture setup in accordance with one or 45 more implementations.

FIG. 3 illustrates examples or artifacts in captured images and the step-wise reduction of such artifacts. FIGs. 3A-3L illustrate captured images that are in turn processed to accomplish lens undistortion, orientation stabilization, and parallax compensation in accordance with one or more implementations.

50 FIG. 4 illustrates principles for achieving an omni-stereo effect in accordance with one or more implementations.

FIG. 5 illustrates aliasing artifacts related to FIG. 4.

FIG. 6 illustrates a resolution to the aliasing artifacts of FIG. 5.

55 FIG. 7 illustrates a table comparing rotational and vertical drift for different approaches.

FIG. 8 illustrates visual artifacts that may result from existing blending approaches.

FIG. 9-9A illustrate the stereoscopic panorama of FIG. 1 and its associated net flow diagram.

FIGs. 10-11 illustrate examples of rotational and vertical drift under different approaches.

5 FIGs 12, 12A-12C illustrate examples having strong parallax under different (alignment) approaches.

FIGs 13, 13A-13D illustrate examples having thin vertical structures under different approaches.

10 FIG. 14, 14A-14D illustrate examples of flow-based blending using under-sampled input under different approaches.

FIG. 15, 15A-15D illustrate examples of circular and linear panoramas in accordance with one or more implementations.

15 FIG. 16 illustrates a system 100 configured to generate stereoscopic panoramas, in accordance with one or more implementations.

FIG. 17 illustrates a method for generating stereoscopic panoramas, in accordance with one or more embodiments

DETAILED DESCRIPTION

20 [0017] Many of the figures illustrate one or more different types of visual artifacts that may result from different approaches and/or procedures to the generation of stereoscopic panoramas. The visual artifacts in any particular figure are not limited to the described artifacts, but may include additional artifacts as may be apparent to knowledgeable practitioners in relevant fields of technology. Some of the figures include stereoscopic (anaglyph) images, which inherently and expectantly may look blurry, in particular when reproduced in black and white. However, knowledgeable practitioners will be able to distinguish any relevant details, including any visual artifacts described in this disclosure. Detailed line-drawings are included for additional clarity and to aid the reader's understanding.

25 [0018] FIG. 16 illustrates a system 100 configured to generate stereoscopic panoramas. System 100 may be configured to co-operate and/or communicate with one or more cameras 14. System 100 may include electronic storage 50, one or more processors 110, and/or other components. Processor 110 may be configured to execute one or more computer program components. The computer program components may be configured to enable a user to interface with system 100 and/or external resources 16, and/or provide other functionality described herein. System 100 may include communication lines, or ports to enable the exchange of information with a network and/or other computing platforms. System 100 may include a plurality of hardware, software, and/or firmware components operating together to provide the functionality described herein. For example, system 100 may be implemented by a cloud of computing platforms operating together.

30 [0019] External resources 16 may include sources of information, hosts and/or providers of (image) storage outside of system 100, external entities participating with system 100, and/or other resources. In some implementations, some or all of the functionality attributed herein to external resources 16 may be provided by resources included in system 100.

35 [0020] Electronic storage 50 of system 100 may be configured to store different information than electronic storage 50 of camera 14. System 100, camera 14, and/or external resources 16 may be operatively linked via one or more electronic communication links. For example, such electronic communication links may be established, at least in part, via a network 13 such as the Internet and/or other networks. It will be appreciated that this is not intended to be limiting, and that the scope of this disclosure includes implementations in which operative links and/or connections may be 40 implemented via some other communication media.

45 [0021] System 100 may be configured to execute one or more computer program components. The computer program components may include one or more of an image component 22, an up-sampling component 23, a stitching component 24, a distortion component 25, an orientation component 26, a vertical parallax component 27, an alignment component 28, and/or other components.

50 [0022] High-resolution stereoscopic panoramas may have visual artifacts. This disclosure describes corrective steps and a compact representation for the input images in order to determine a highly accurate approximation of the required ray space. By virtue of flow-based up-sampling (e.g. by up-sampling component 23) of the available input rays (interchangeably referred to as captured images or input images), known aliasing issues, such as stitching artifacts, may be effectively reduced and/or resolved. The required rays (and/or interpolated image-fragments) may be generated on-the-fly to match the desired output resolution, even for small numbers of input images. Up-sampling may be performed in real-time, thus enabling direct interactive control over the desired stereoscopic depth effect. The disclosure enables the generation of stereoscopic panoramas at high output resolutions that are virtually free of various artifacts, including but not limited to seams, stereo discontinuities, vertical parallax, mono-/stereoscopic shape distortions, and/or other artifacts.

The described features and/or techniques may be applied to other types of multi-perspective panoramas, such as linear panoramas.

[0023] There is a strong consumer interest in a more immersive experience of content, such as 3D photographs, television and cinema. Environmental content may be captured through panoramas. By way of non-limiting example,

5 FIG. 1 illustrates a stereoscopic panorama with a gate or fence in the foreground (near the bottom of the image) and various buildings in the background. As depicted in FIG. 1, the panorama represents a red-cyan anaglyph. The square highlighted sections illustrates one or more visual artifacts and are labeled with reference numbers to line-drawings (from FIG. 1A to FIG. 1H) that illustrates additional details. FIGs. 1A-1B-1C-1D illustrate close-ups of visual artifacts in 2D. FIGs. 1E-1F-1G-1H illustrate close-ups of vertical parallax (using a representation of stereoscopic images).

10 [0024] FIG. 1A illustrates, at least, aliasing and stitching artifacts (as may be visible using known methods of panorama generation) for a section of FIG. 1 that depicts a fence or gate in which a brick or stone column has been embedded. FIG. 1C illustrates the same section as FIG. 1A, in which the visual artifacts are reduced and/or removed in comparison to FIG. 1A.

15 [0025] FIG. 1B illustrates, at least, duplication and stitching artifacts (as may be visible using known methods of panorama generation) for a section of FIG. 1 that depicts a building and at least three framed windows. Elements in this section are duplicated and shifted in FIG. 1B. FIG. 1D illustrates the same section as FIG. 1B, in which the visual artifacts are reduced and/or removed in comparison to FIG. 1B.

20 [0026] FIG. 1E is a stereoscopic representation in which the image for one eye uses solid lines and the image for the other eye uses dotted lines. FIG. 1E illustrates, at least, aliasing and distortion artifacts (as may be visible using known methods of stereoscopic panorama generation) for a section of FIG. 1 that depicts a fence or gate that includes diagonal structures. FIG. 1F illustrates the same section as FIG. 1E, in which the visual artifacts are reduced and/or removed in comparison to FIG. 1E.

25 [0027] FIG. 1G is a stereoscopic representation in which the image for one eye uses solid lines and the image for the other eye uses dotted lines. FIG. 1G illustrates, at least, orientation and distortion artifacts (as may be visible using known methods of stereoscopic panorama generation) for a section of FIG. 1 that depicts a building. FIG. 1H illustrates the same section as FIG. 1G, in which the visual artifacts are reduced and/or removed in comparison to FIG. 1G.

30 [0028] Automatic tools for stitching panoramas from multiple (input) images may be available, even in consumer cameras. For circular 360° panoramas, a user may use, expect to use, and/or be required to use a common camera center for all images. Using a common center minimizes stitching artifacts. This may be achieved by rotating the camera around its optical center. However, such panoramas inherently lack parallax and therefore cannot be experienced stereoscopically.

35 [0029] To overcome this limitation, motion parallax may be explicitly captured by taking images with varying centers of projection, e.g. by mounting the camera on a rotating arm. By way of non-limiting example, such a capture setup is illustrated in the right-hand section of FIG. 2, in which the dotted arrow indicates camera movement. A stereoscopic panorama may be created by stitching together specific strips from the captured input views. While successfully introducing parallax, this strategy may suffer from a number of unresolved practical issues that cause disturbing artifacts such as visible seams or vertical parallax in the panorama. Image-quality issues in 2D may lead to considerable misperceptions and even become intolerable when viewed stereoscopically.

40 [0030] One reason for these artifacts is that in practice a camera can capture light rays only at quite a limited spatio-angular resolution, i.e. a finite set of images with finite resolution. This insufficiently-dense sampling of the scene manifests itself as visible discontinuities in the (generated) output panorama in areas where neighboring output pixels have been synthesized from different input views having strong parallax between them. Capturing more and higher resolution images may not be a practically feasible solution, since this would result in orders of magnitude more data while not fully resolving those issues. Secondly, while the optical centers of all camera positions should ideally lie on a perfect 45 circle to create panoramas, in practical acquisition scenarios, and especially for hand-held capture (as illustrated in the left-hand section of FIG. 2, in which the dotted arrow indicates camera movement), this may be difficult to achieve.

45 [0031] By virtue of this disclosure, automated, high-quality, virtually artifact-free stereoscopic panorama generation may be provided through system 100. System 100 may be configured to compensate for deviations from an ideal capture setup, e.g. through operations of orientation component 26. Specifically, the captured images may be corrected to resolve 50 issues caused by perspective distortion and vertical parallax (e.g. by vertical parallax component 27). System 100 may be configured to obtain alignment of the captured images with minimal drift, e.g. through operations of alignment component 28. System 100 may be configured to reduce visible seams caused by the truncations and duplications of objects. The set of captured and corrected light rays (or images) may be up-sampled using optical flow-based interpolation techniques (e.g. by up-sampling component 23). In effect, a (nearly and/or virtually) continuous representation of the 55 ray space may be achieved. System 100 may be configured to sample from this representation, and with real-time control over the resulting stereo effect.

[0032] The disclosed techniques may be used for any multi-perspective imaging method based on stitching, like x-slit, pushbroom, or general linear cameras. Multi-perspective approaches combine images taken from different viewpoints

to create mosaics that cannot be obtained from a single center of projection. For example, pushbroom images may be stitched from images taken by a linearly moving camera. This may be well suited for capturing buildings along a street. However, existing approaches cannot capture both a 360° view of the scene and stereoscopic depth. For panoramas the camera may be moved along a circular trajectory with either a tangential or orthogonal camera viewing direction.

5 The latter may be referred to as omni-stereo panoramas. These approaches capture a range of viewpoints for the same panorama, thus in theory allowing stereoscopic renderings with varying baseline.

[0033] System 100 may be configured to stitch the images into a panorama, responsive to alignment of the images relative to each other. Such alignment may amount to estimation of the camera motion. This may be accomplished with a purely image-based method that estimates a parameterized motion model (e.g. a 2D homography) from image correspondences. In some implementations, estimation may use image registration techniques. Image-based methods may lack knowledge of the scene geometry (e.g. depth) and only compute the optimal 2D alignment. This may cause artifacts if the scene has a complex depth structure. In some implementations, the scene depth of the panorama may be estimated and used to compute the ego motion of the camera, *i.e.* the image motion scaled by the depth. With this information, system 100 may be configured to compute an optimal (possibly curved) shape of each strip to be pasted.

10 15 However, estimating the shape of strips as well as the depth may be computationally expensive, while not resolving all problems.

[0034] To obtain high-quality alignment, system 100 may be configured to leverage structure-from-motion algorithms to estimate the full 3D camera position, location, and/or movement. Such an approach may allow for global optimization of camera poses.

20 [0035] During stitching of multi-perspective images, parallax may lead to disturbing seams, *i.e.* discontinuities between the stitched strips. One way to alleviate this problem may be to blend between the strips using strategies like simple linear (alpha), pyramid-based, or gradient-domain blending. These strategies effectively attempt to hide the discontinuities and thus only work well for mild seams. In the presence of significant discontinuities caused by strong parallax, the results may include noticeable artifacts. In the context of omni-stereo panoramas, concise control over the resulting 25 output parallax may be required in order to achieve proper stereoscopic viewing. While the above blending approaches might be applicable for monoscopic stitching, in stereoscopic 3D the resulting inconsistencies may become more noticeable and/or unacceptable.

30 [0036] A reduction of artifacts may be accomplished by using more images to get a denser angular sampling of the scene, resulting in thinner strips and smaller discontinuities. This can be achieved by leveraging depth or optical flow information to interpolate novel virtual views. For practical considerations all interpolated images cannot be stored, in particular for high spatial resolutions. Furthermore, denser angular sampling may still leave some artifacts for thin vertical structures as they may be missed in depth or flow estimates. In such cases, system 100 may automatically use an alternative approach, including but not limited to linear blending, which at these small locations allows to remedy the stitching problems and degrades gracefully. In some implementations, system 100 may be configured to combine multiple 35 blending techniques (simultaneously and/or subsequently) responsive to a determination that doing so may reduce artifacts.

[0037] Optical flow may also be used for improving hand-held capture of 2D panoramas, to remove the (in this case) undesired motion parallax.

40 [0038] Omni-stereo panorama generation may be based on creating two sets of light rays: one for the left eye and one for the right eye, ideally with all rays and/or camera directions tangential to a circle. By way of non-limiting example, FIG. 4 illustrates the rays required to generate such a panorama. In practice, systems usually capture an image set with a camera moving along a circular trajectory with its principal axis parallel to the plane spanned by the camera trajectory, as depicted in FIG. 2. An omni-stereo panorama may be created by stitching, for each eye, specific vertical strips from the aligned images, such that the ray geometry as depicted in FIG. 4 may be approximated. By way of illustration, FIG. 45 5 illustrates duplication and truncation artifacts caused by aliasing. By way of illustration, FIG. 6 illustrates flow-based up-sampling to synthesize required rays and/or image-fragments.

50 [0039] This approximation to the desired ray geometry typically suffers from inaccuracies of the capture setup and limited angular sampling, resulting in the previously mentioned artifacts. By using a specific transformation and alignment approach (including, but not limited to, employing camera orientation correction, cylindrical image re-projection, and optimized homography matching) visual artifacts may be reduced.

[0040] In practice the angular sampling may be sparse, *e.g.* using a limited number of input images. This may further deteriorate the approximation quality and may lead to aliasing artifacts (*e.g.* seams, truncation, and/or duplication).

55 [0041] Input correction may include optical undistortion to better approximate a pinhole model, correction of camera orientation, vertical parallax compensation, compact representation via 2D alignment, and/or other types of corrections and/or improvements. Input correction may be accomplished and/or performed by one or more components, including, but not limited to, stitching component 25, vertical parallax component 27, and/or other components.

[0042] For capturing stereoscopic panoramas, it may be beneficial to use wide-angle lenses to capture images with significant overlap and a large vertical field of view. Due to the optical distortion introduced by those lenses, approximation

of the desired ray geometry may include converting images to correspond to a pinhole camera model. Rational lens distortion models may provide better results than the standard polynomial approach targeting simpler radial and tangential lens distortion.

[0043] Any deviation from an ideal capture setup (e.g. circular trajectory and parallel principal axis) may lead to visible shape distortions in a stereoscopic output panorama (e.g. tilted buildings, vertical parallax). System 100 may be configured to correct the orientation and viewing direction of each input view to approximate an ideal setup, for example by orientation component 26.

[0044] The camera poses may be estimated using general structure-from-motion with global bundle adjustment. For purely circular motions, restricted models could be used in principle, such as enforcing the camera positions to lie on a circle. However, a more general approach enables creation of omni-stereo panoramas from hand-held input and/or other camera trajectories, including, but not limited to, a straight line. Inherent scale ambiguity may be removed by scaling coordinates so that the camera's trajectory conforms to the actual physical dimensions.

[0045] Individual camera coordinate frames may be rotated towards the idealized setup with a single consistent up-direction and consistent viewing directions that are perpendicular to the camera trajectory. By rotating around the optical center (e.g. using a homography) the transformation may be scene independent and hence does not introduce inaccuracies. Let $P=KR[I|-C]$ be the decomposition of a camera's projection matrix into intrinsics K , rotation R and the camera center C . Further let $R=[x \ y \ z]$ represent the left, up and viewing direction of that camera, respectively. System 100 may be configured to compute a consistent up direction "u" across multiple images. For omni-stereo panoramas, we obtain "u" by fitting a circle to all camera centers and using the normal vector n of the plane the circle lies in: $u=n$. For other trajectories (e.g. linear as for pushbroom panoramas), system 100 may be configured to compute "u" as the mean up direction of the individual camera coordinate systems. The mean up direction may also be used to disambiguate the normal direction in the omni-stereo case, by choosing the direction that is closer to the mean up direction. Given the new up direction, the corrected coordinate frame may be computed as follows: the new viewing direction becomes $v = u \times x$, and the new left direction is $w = v \times u$ (here we assume that the vectors are normalized before each step). The rotation that needs to be applied to a camera then is $R'=[w \ u \ v]^T R^{-1}$. This rotation is equivalent to the 2D homography $H=KR'K^{-1}$ applied to the image and achieves the desired idealized capture setup.

[0046] By way of non-limiting example, FIG. 3 illustrates the effect of, at least, this transformation and alignment, for a scene depicting a street with buildings on either side and a lamppost on the left side. The images in the top row of FIG. 3 have been captured as the camera moved from a right-most position (FIG. 3C), to a center position (FIG. 3B), and finally to a left-most position (FIG. 3A). In other words, FIGs. 3A-3B-3C form the captured, or "raw" images. Each subsequent row below this top row illustrates one or more particular types of artifact-reduction based on the image directly above it. For example, the center image in the top row is processed to produce the center image in the row below it, then further processed to produce the center image in the row below it, and finally processed again to produce the center image in the bottom row.

[0047] FIG. 3A illustrates the captured image at the left-most position. FIG. 3A illustrates, at least, lens distortion that has warped the image. FIG. 3B illustrates the captured image at the center position. FIG. 3C illustrates the captured image at the right-most position. FIG. 3C illustrates, at least, lens distortion that has warped the image. By virtue of performing the process referred to as "lens undistortion" (explained elsewhere in this disclosure), FIGs. 3A-3B-3C are converted/processed/transformed into FIGs. 3D-3E-3F, respectively.

[0048] FIG. 3D illustrates the "lens undistorted" image at the left-most position, for which the warping artifact has been reduced and/or removed. FIG. 3E illustrates the "lens undistorted" image at the center position. FIG. 3F illustrates the "lens undistorted" image at the right-most position, for which the warping artifact has been reduced and/or removed. Note that the building on the right-hand side is tilted at -3.9, 0.0 and 3.3 degrees in FIGs. 3D-3E-3F, respectively. By virtue of performing the process referred to as "camera correction," "orientation stabilization," and/or "perspective undistortion" (explained elsewhere in this disclosure), FIGs. 3D-3E-3F are converted/processed/transformed into FIGs. 3G-3H-3I, respectively.

[0049] FIG. 3G illustrates the "orientation stabilized" image at the left-most position, for which the tilting artifact has been reduced. FIG. 3H illustrates the "orientation stabilized" image at the center position. FIG. 3I illustrates the "orientation stabilized" image at the right-most position, for which the tilting artifact has been reduced. Note that the building on the right-hand side is tilted at (merely) 0.6, 0.9 and 0.8 degrees in FIGs. 3G-3H-3I, respectively. The vertical line that connects the top of the lampposts in FIGs. 3G-3H-3I has been added to indicate vertical parallax between images. By virtue of "parallax compensation" (explained elsewhere in this disclosure), FIGs. 3G-3H-3I are converted/processed/transformed into FIGs. 3J-3K-3L, respectively.

[0050] FIG. 3J illustrates the "parallax compensated" image at the left-most position, for which the vertical parallax has been reduced. FIG. 3K illustrates the "parallax compensated" image at the center position. FIG. 3L illustrates the "parallax compensated" image at the right-most position, for which the vertical parallax has been reduced. The vertical line that connects the top of the lampposts in FIGs. 3J-3K-3L has been added to indicate the reduction of vertical parallax between images.

[0051] Lens undistortion (e.g. by distortion component 25) may restore rectilinear projection and straight lines. Correction of the camera orientations (e.g. by orientation component 26) may remove perspective distortion. Cylindrical projection (e.g. by vertical parallax component 27) may compensate for vertical parallax. In the standard pinhole model with a planar imaging surface, objects near the image border are larger than near the image center. As a consequence, for non-linear camera trajectories, motion parallax between two input views may include a horizontal and a vertical component. While the horizontal component may be desirable for constructing a stereoscopic output image, the vertical component may need to be eliminated in order to allow for proper stereoscopic viewing.

[0052] By re-projecting individual ones of the undistorted, orientation-corrected input images onto a cylindrical imaging surface, the vertical parallax may be effectively removed. The cylinder may be defined to be concentric to the circle fitted to the camera centers (as described elsewhere herein), with the cylinder's axis orthogonal to the circle plane and a specific radius. The cylinder radius may determine the amount of horizontal scale changes for objects at different distances. In practice, setting the radius to the approximate distance of salient objects to the camera may provide good results. By virtue of the structure-from-motion (SfM) computation (described elsewhere), system 100 may be configured to provide various opportunities for a user to define those distances interactively and/or automatically from the scene geometry.

[0053] To efficiently project an image onto a cylinder, a pixel grid may need to be established on the cylinder at the desired output resolution. Individual pixels may be projected onto the pinhole camera's imaging plane to sample the corresponding output color. Specifically, for individual ones of the images, the extent of the image on the cylinder may be approximated by tracing rays from pixels on the image border through the camera center and intersecting these rays with the cylinder.

[0054] In some implementations, the re-oriented, parallax-compensated images may be used for synthesizing an omni-stereo panorama. In some implementations, another processing step, transformation, may provide beneficial results. The remaining relative transformations between the images may be mostly simple horizontal translations without any rotational component. By projecting the images back into a planar 2D setting that is compatible with existing methods for panorama generation the process of stitching together images may be simplified. The alignment information about images may be encoded via homographies. For pairs of consecutive images, the reconstructed camera geometry may be leveraged to calculate the homography induced by a plane tangent to the cylinder halfway between the two camera centers in order to minimize distortions. For general camera trajectories, the plane may be positioned at a fixed distance in front of the midpoint of the two camera centers, with the plane normal halfway between the viewing directions of the two cameras. The plane-induced homography H_p (in normalized camera coordinates) may be decomposed to remove 5 of its 8 degrees of freedom to obtain a Euclidean transform. Specifically, if the chain of transformations may be represented as:

$$H_p = \begin{bmatrix} A & t \\ v^T & v \end{bmatrix} = \underbrace{\begin{bmatrix} sR & t \\ 0^T & 1 \end{bmatrix}}_{\text{similarity}} \underbrace{\begin{bmatrix} K & 0 \\ 0^T & 1 \end{bmatrix}}_{\text{affinity}} \underbrace{\begin{bmatrix} I & 0 \\ v^T & v \end{bmatrix}}_{\text{projectivity}}, \quad (1)$$

(Note that the transformations may be represented using similarity, affinity, and projectivity.)

[0055] By decomposing $sRK = A - tv^T$ using QR decomposition the rotation R may be obtained. The decomposition yields an orthonormal matrix Q^* and an upper-triangular matrix R^* . R may be obtained by inverting the columns of Q^* with negative diagonal entries to restrict the rotation to less than $\pm 90^\circ$. Compared to purely image-based alignment techniques, the correction and alignment steps as described herein may reduce the accumulation of positional errors (drift) in panoramas.

[0056] In Table 1 (FIG. 7), both the rotational and vertical drift corresponding to different approaches are compared. The approaches described herein may reduce drift in all or most cases by at least two orders to an unnoticeable 0.1° . As depicted in Table 1, rotational drift between both ends of a panorama, and remaining vertical drift after cancellation of rotational drift, may be compared between an image-based (IB) approach and the approach described herein, which includes virtually no drift.

[0057] Stitching an omni-stereo panorama may be accomplished by extracting specific strips from each input image - dependent on the desired stereoscopic output disparities - and by combining the strips into a left and right output view. The omni-stereo effect may be achieved by collecting rays that are tangent to a common viewing circle, as illustrated in FIG. 4.

[0058] Proper capturing and/or sampling of input rays and/or images for the generation of a stereo panorama may depend on the targeted output resolution. In some implementations, strips from input views or images may project to less than a pixel's width in the output image, thereby avoiding aliasing artifacts and deviation of rays from the desired ray projection. The deviation angles may be defined as the angular difference between the ideal ray and the ray that is used for stitching (see angles α_1, α_2 in FIG. 4). The output quality may depend on the angular resolution of the capture setup, *i.e.* the number of input views rather than their spatial resolution. With a coarser angular resolution the deviation angles grow (approximately inversely proportional to the angular resolution) and stitching artifacts such as discontinuities, duplications, and truncations of objects may become apparent.

[0059] Artifacts may be reduced through various forms of blending. FIG. 8 illustrates that some blending approaches may obscure artifacts, to some extent, without addressing the problem at its core. From left to right, FIG. 8 illustrates the results for "nearest neighbor" blending, linear blending, pyramid-based blending, and flow-based blending (which is explained in more detail elsewhere in this disclosure). Capturing enough input images for an angular resolution of less than 0.05° may lead to prohibitive memory requirements, orders of magnitude higher than the approach described in this disclosure. By virtue of the disclosed techniques, angular resolutions from $0.1^\circ, 0.5^\circ, 1^\circ, 2^\circ, 3^\circ, 4^\circ, 8^\circ$, and/or other angular resolutions may result in a high-quality panorama that is virtually artifact-free.

[0060] In order to characterize the aliasing artifacts, consider FIG. 5. Given two images I_K and I_L , the rays may need to be collected at an angle β to the principal axis. Let E and G be the intersections of these projected rays with the cylinder. Filling the strip EG in the panorama will in general require additional nearby rays to compensate for the relatively coarse angular resolution between input images. An object at distance d_{far} from the rotation center may be further away than the cylindrical projection. The section $C D$ (in the tree) will be duplicated in the stitched output since its projections c' and d' as well as c'' and d'' from images I_K and I_L , respectively, to the cylinder are visible in both the green and the red strip. On the other hand, objects in the range $A B$ at distance d_{near} may appear truncated in the final panorama.

[0061] These aliasing artifacts may be resolved by generating the missing rays using optical flow-based up-sampling, as illustrated in FIG. 6, *e.g.* by up-sampling component 23. A point P at distance d_{far} is projected to p' and p'' from images I_K and I_L , respectively, to the cylinder. The optical flow vector F at point p' maps to point p'' as $p' + F(p') = p''$. To avoid stitching artifacts, the intermediate point \tilde{p} between p' and p'' may be interpolated, effectively synthesizing missing rays

at the virtual camera location \tilde{I}_M . The same concept may apply to points closer than the cylindrical projection surface, as well as to other camera trajectories like the ones used for street view panoramas.

[0062] In some implementations, system 100 may be configured to use pairwise optical flow. For a simplified notation, the flow-based ray up-sampling may be described using the corrected images. However, system 100 may be configured to compute the flow directly on the input images, which thereby undergo the same correction transformations.

[0063] Denote the flow fields between pairs of adjacent images k and l by $F_{k \rightarrow l}$. This provides image correspondences, *i.e.* $I_k(x) \approx I_l(x + F_{k \rightarrow l}(x))$, where both images and the flow field are defined on the same pixel grid. However, this assumption may be violated by the alignment described elsewhere herein, which aligns all images to the same global coordinate frame, with a different pixel grid for each image.

[0064] System 100 may be configured to compensate for the global alignment. This compensation may be performed prior to using the flow fields to interpolate missing rays. The flow field $F_{k \rightarrow l}$ is defined on the pixel grid of I_k , and consequently it is sampled in the coordinate frame of image k , at $x_k = A_{0 \rightarrow k} x$, where $A_{0 \rightarrow k}$ transforms from global coordinates to the local coordinate system of I_k . The flow field may also encompass the Euclidean motion which is already accounted for by aligning the images in global space. This motion is given by $A_{0 \rightarrow l} x - A_{0 \rightarrow k} x$, which leads to the definition of the motion-compensated net flow as

$$F^*_{k \rightarrow l}(x) = F_{k \rightarrow l}(A_{0 \rightarrow k} x) - (A_{0 \rightarrow l} x - A_{0 \rightarrow k} x). \quad (2)$$

[0065] An example for the resulting net flow is shown in FIG. 9 (and additionally in FIG. 9A). The top-half of FIG. 9 illustrates a similar panorama as depicted in FIG. 1. The bottom half depicts the computer net flow for the image in the top half. Objects at greater distance from the camera, such as the buildings in FIG. 9, may have little or no net flow, as depicted by white or near-white sections (labeled 90) in the bottom half of FIG. 9 and in FIG. 9A. Objects at a medium distance from the camera, such as certain sections of the fence or gate in FIG. 9, may have a small net flow, as depicted by lightly-shaded sections (labeled 91) in the bottom half of FIG. 9 and in FIG. 9A. Objects close to the camera, such as certain parts of the fence or gate in FIG. 9, may have a (relatively) large net flow, as depicted by more darkly shaded sections (labeled 92) in the bottom half of FIG. 9 and in FIG. 9A.

[0066] The desired interpolation to synthesize in-between light rays on the fly may be achieved by warping corresponding pixels by a fraction of the flow, depending on the horizontal angular interpolation factor η between two images.

Linearly blending between the two warped images k and l dependent on η may produce the flow-based blending result as

$$5 \quad \mathbf{S}_k^{\text{Flow}}(\mathbf{x}) = (1 - \eta) \cdot \mathbf{I}_k(\mathbf{A}_{0 \rightarrow k} \mathbf{x} + \eta \cdot \mathbf{F}_{k \rightarrow l}^*(\mathbf{x})) \\ + \eta \cdot \mathbf{I}_l(\mathbf{A}_{0 \rightarrow l} \mathbf{x} + (1 - \eta) \cdot \mathbf{F}_{l \rightarrow k}^*(\mathbf{x})). \quad (3)$$

10 [0067] This type of blending may degrade gracefully in cases where the flow estimation shows inaccuracies.

[0068] The described techniques may require querying the two images at some warped sub-pixel locations of the form $\mathbf{x} + \eta \times \mathbf{F}^*$, which can easily and efficiently be implemented on a graphics processing unit (GPU) using texture lookups. The flow-based blending may be accomplished without computational overhead during rendering. In some implementations, the flow fields may be pre-computed.

15 [0069] By virtue of the described techniques in this disclosure, system 100 may be configured to stitch the panoramas in real-time at full screen resolution, which gives the user the freedom to adapt important stereo viewing parameters like interaxial distance or vergence on the fly. Furthermore, as illustrated by comparing FIG. 10 and FIG. 11, system 100 may be configured to obtain clean and linear results, e.g. by removing rotational and vertical drift.

20 [0070] FIG. 12 illustrates an example with very strong parallax caused by a person standing close to the camera. Here, it becomes clear that purely image-based (alignment) strategies lead to severe shape distortions, whereas the techniques described in this disclosure produce a high-quality result without distortions. The images in the top row have been produced using image-based alignment of images captured at three positions: the right-most position (FIG. 12C), the center position (FIG. 12B), and the left-most position (FIG. 12A). The images in the bottom row of FIG. 12 have been 25 produced using the alignment described in this disclosure. FIGs. 12A-12B-12C are line-drawings that represent the left-most, center, and right-most column in FIG. 12, respectively. The distortion and parallax artifacts in the top half of each of FIGs 12-A-12B-12C is clear, especially in comparison with the bottom half of the same figures.

30 [0071] FIG. 13 illustrates an example with thin (vertical) structures and linear blending. Note that these results are actually not stereo panoramas, but linear pushbroom images as the input was captured from a car driving along the street. Note the fence in the foreground. FIG. 13 illustrates that the disclosed techniques may be applied for this type of input. The top-left image (FIG. 13A) illustrates the result for minimal aspect distortion (MAD) mosaicing of the captured input (using crossed-slit projection). Note the visual artifact that is emphasized with an oval shape. The bottom-left (FIG. 13C) illustrates the depth (camera distance) for the top-left image, in which nearby objects are light or lighter and faraway objects are dark or darker. In FIG. 13C, nearby objects are labeled 60, the furthest objects/sections are labeled 69, and 35 objects/sections in-between are correspondingly labeled numerically in-between 60 and 69. The top-right image (FIG. 13B) illustrates the result for applying the techniques described in this disclosure. A similar oval shape emphasizes a reduction in the visual artifact for a thin vertical structure. FIG. 13D illustrates the net flow for image 13B. The net flow at or below dotted line 70 ranges between a small net flow (indicated by light shading and/or a label of 71) to a medium net flow (indicated by medium shading and/or a label of 72) to a relatively larger net flow (indicated by darker shading 40 and/or a label of 73). The net flow above dotted line 70 has the opposite direction as the net flow below dotted line 70. The net flow above dotted line 70 ranges between a small net flow (indicated by light shading and/or a label of 75) to a medium net flow (indicated by medium shading and/or a label of 76) to a relatively larger net flow (indicated by darker shading and/or a label of 77).

45 [0072] FIG. 15 illustrates examples of circular and linear stereoscopic panoramas. Stereoscopic panoramas corrected and stitched using the described techniques herein are shown as red-cyan anaglyph images. The top-left section of FIG. 15 is FIG. 15A. The top-left section of FIG. 15 is FIG. 15B. The bottom-left section of FIG. 15 is FIG. 15C. The bottom-right section of FIG. 15 is FIG. 15D (a linear panorama).

50 [0073] FIG. 14 illustrates flow-based blending (bottom row) using under-sampled input. As shown, other blending methods (including nearest neighbor blending in the top row and linear blending in the center row) may fail (e.g. show visual artifacts, including but not limited to aliasing, duplication, and/or blending artifacts) in these scenarios. The left-most column (FIG. 14A) represents results from images captures at 1 degree intervals. The columns next to the left-most column, from left to right, represent 2, 4, and 8 degree intervals, respectively (depicted by FIGs. 14B-14C-14D respectively). Visual artifacts are worse for larger intervals, so gradually get worse from the left-most to the right-most column.

55 [0074] Referring to FIG. 16, electronic storage 50 may comprise electronic storage media that electronically stores information. The functionality of electronic storage 50 may be similar between different instantiations, even if the stored information may be different. Electronic storage 50 may be configured to store captured images, pre-processed images, pre-processed information, interpolated images, generated panoramas, and/or other information. The electronic storage

media of electronic storage 50 may include one or both of system storage that is provided integrally (i.e., substantially non-removable) and/or removable storage that is removably connectable via, for example, a port (e.g., a USB port, a firewire port, etc.) or a drive (e.g., a disk drive, etc.). Electronic storage 50 may include one or more of optically readable storage media (e.g., optical disks, etc.), magnetically readable storage media (e.g., magnetic tape, magnetic hard drive, floppy drive, etc.), electrical charge-based storage media (e.g., EEPROM, RAM, etc.), solid-state storage media (e.g., flash drive, etc.), and/or other electronically readable storage media. Electronic storage 50 may include one or more virtual storage resources (e.g., cloud storage, a virtual private network, and/or other virtual storage resources). Electronic storage 50 may store software algorithms, information determined by processor(s) 110, information received from components of system 100, and/or other information that enables system 100 to function as described herein.

[0075] Processor(s) 110 may be configured to provide information processing capabilities in system 100. The functionality of processor(s) 110 may be similar between different instantiations, even if the processing capabilities may be different. Processor(s) 110 may include one or more of a digital processor, an analog processor, a digital circuit designed to process information, an analog circuit designed to process information, a physical processor, and/or other mechanisms for electronically processing information. Although processor(s) 110 is shown in FIG. 16 as a single entity, this is for illustrative purposes only. In some embodiments, processor(s) 110 may include a plurality of processing units. These processing units may be physically located within the same device, or processor(s) 110 may represent processing functionality of a plurality of devices operating in coordination. The processor(s) 110 may be configured to execute components 22-28, and/or other components, modules, and/or code devices. Processor 110 may be configured to execute components 22-28, and/or other components by software; hardware; firmware; some combination of software, hardware, and/or firmware; and/or other mechanisms for configuring processing capabilities on processor(s) 110.

[0076] It should be appreciated that although components 22-28 are illustrated in FIG. 16 as being co-located within a processing unit, one or more of components 22-28 may be located remotely from the other components. The description of the functionality provided by the different components 22-28 described herein is for illustrative purposes, and is not intended to be limiting, as any of components 22-28 may provide more or less functionality than is described. For example, one or more of components 22-28 may be eliminated, and some or all of its functionality may be provided by other ones of components 22-28. As another example, processor 110 may be configured to execute one or more additional components that may perform some or all of the functionality attributed herein to one of components 22-28. The order in which operations and/or functionality attributed to different components is ordered and/or organized is not intended to be limited in any way, and any such order is envisioned within the scope of this disclosure.

[0077] FIG. 17 illustrates a method 200 for generating stereoscopic panoramas, in accordance with one or more embodiments. The operations of method 200 presented below are intended to be illustrative. In some embodiments, method 200 may be accomplished with one or more additional operations not described, and/or without one or more of the operations discussed. Additionally, the order in which the operations of method 200 are illustrated in FIG. 17 and described below is not intended to be limiting.

[0078] In some embodiments, method 200 may be implemented in one or more processing devices (e.g., a digital processor, an analog processor, a controller, a digital circuit designed to process information, an analog circuit designed to process information, and/or other mechanisms for electronically processing information). The one or more processing devices may include one or more devices executing some or all of the operations of method 200 in response to instructions stored electronically on an electronic storage medium. The one or more processing devices may include one or more devices configured through hardware, firmware, and/or software to be specifically designed for execution of one or more of the operations of method 200.

[0079] At an operation 202, images based on captured images are obtained. At least some of the captures images form a set of adjacent images. In some embodiments, operation 202 is performed by an image component the same as or similar to image component 22 (shown in FIG. 16 and described herein).

[0080] At an operation 204, interpolated images between adjacent images are determined. In some embodiments, operation 204 is performed by an up-sampling component the same as or similar to up-sampling component 23 (shown in FIG. 16 and described herein).

[0081] At an operation 206, a stereoscopic panorama is determined based on obtained images and interpolated images. In some embodiments, operation 206 is performed by a stitching component the same as or similar to stitching component 24 (shown in FIG. 16 and described herein).

[0082] The present invention, in accordance with one or more various implementations, is described above in detail with reference to the accompanying figures. The drawings are provided for purposes of illustration only and merely depict exemplary implementations. These drawings are provided to facilitate the reader's understanding of the systems and methods described herein and shall not be considered limiting of the breadth, scope, or applicability of any of the claims.

[0083] Although the systems and methods set forth herein are described in terms of various exemplary implementations and implementations, it should be understood that the various features, aspects and functionality described in one or more of the individual implementations are not limited in their applicability to the particular implementation with which they are described, but instead can be applied, alone or in various combinations, to one or more of the other implemen-

tations, whether or not such implementations are described and whether or not such features are presented as being a part of a described implementation. Thus, the breadth and scope of the present disclosure should not be limited by any of the above-described exemplary implementations.

[0084] Terms and phrases used in this document, and variations thereof, unless otherwise expressly stated, should be construed as open ended as opposed to limiting. As examples of the foregoing: the term "including" should be read as meaning "including, without limitation" or the like; the term "example" is used to provide exemplary instances of the item in discussion, not an exhaustive or limiting list thereof; the terms "a" or "an" should be read as meaning "at least one," "one or more" or the like; and adjectives such as "conventional," "traditional," "normal," "standard," "known" and terms of similar meaning should not be construed as limiting the item described to a given time period or to an item available as of a given time. Likewise, where this document refers to technologies that would be apparent or known to one of ordinary skill in the art, such technologies encompass those apparent or known to the skilled artisan now or at any time in the future.

[0085] The presence of broadening words and phrases such as "one or more," "at least," "but not limited to" or other like phrases in some instances shall not be read to mean that the narrower case is intended or required in instances where such broadening phrases may be absent.

[0086] Additionally, the various implementations set forth herein are described in terms of exemplary block diagrams, flow charts, and/or other illustrations. As will become apparent to one of ordinary skill in the art after reading this document, the illustrated implementations and their various alternatives can be implemented without confinement to the illustrated examples. For example, block diagrams and their accompanying description should not be construed as mandating a particular architecture or configuration.

[0087] Although the present technology has been described in detail for the purpose of illustration based on what is currently considered to be the most practical and preferred implementations, it is to be understood that such detail is solely for that purpose and that the technology is not limited to the disclosed implementations, but, on the contrary, is intended to cover modifications and equivalent arrangements that are within the spirit and scope of the appended claims. For example, it is to be understood that the present technology contemplates that, to the extent possible, one or more features of any implementation can be combined with one or more features of any other implementation. It is to be understood that the present technology contemplates that one or more features of any claim can be dependent on and/or combined with one or more features of any other claim.

Claims

1. A system configured to generate stereoscopic panoramas, the system comprising:

35 one or more physical processors configured to execute computer program modules, the computer program modules comprising:

an image module configured to obtain images based on captured images, wherein at least some of the captured images form a sequence of adjacent images;

40 an up-sampling module configured to determine interpolated images between adjacent images; and

a stitching module configured to determine a stereoscopic panorama based on obtained images and interpolated images.

2. The system of claim 1, wherein the up-sampling module is configured to determine the interpolated images by determining flow fields between related pixels in adjacent images.

3. The system of claim 2, wherein the up-sampling module is configured to determine the flow fields by compensating for transformations of pixel grids for the obtained images; and/or

50 wherein the stitching module is configured to determine the stereoscopic panorama based on determining vertical strips of the obtained images and further based on blending the vertical strips, wherein blending is based on the determined flow fields.

4. The system of any preceding claim, wherein the obtained images are based on compensated images, the system further comprising:

55 a distortion module configured to determine the compensated images based on the captured images, wherein the distortion module is configured to determine the compensated images by compensating for perspective distortion caused by a shape of a lens of a camera used to capture the captured images.

5. The system of claim 4, wherein the obtained images are based on orientation-corrected images, the system further comprising:
an orientation module configured to determine the orientation-corrected images based on the compensated images, wherein the orientation module is configured to determine the orientation-corrected images by compensating for one or both of a non-circular camera trajectory and/or a non-parallel principal axis used during image capture.

6. The system of claim 5, wherein the obtained images are based on re-projected images, the system further comprising:
a vertical parallax module configured to determine the re-projected images based on the orientation-corrected images, wherein the vertical parallax module is configured to determine the re-projected images by re-projecting orientation-corrected images onto a cylindrical imaging surface and reducing vertical parallax.

7. The system of claim 6, wherein the obtained images are based on aligned images, the system further comprising:
an alignment module configured to determine the aligned images based on alignment of the re-projected images and/or vertical fragments/strips thereof, wherein the alignment module is configured to determine the aligned images by projecting the re-projected images into a planar two-dimensional surface such that rotational and/or positional drift between adjacent images is reduced, wherein alignment is based on structure-from-motion.

8. The system of claim 7, wherein alignment by the alignment module causes a transformation of pixel grids of the aligned images, and wherein the up-sampling module is configured to determine the flow fields by compensating for the transformation caused by the alignment module.

9. The system of any preceding claim, further comprising physical storage media configured to store the captured images, optionally further comprising a camera configured to capture images, wherein the camera is configured to be carried and/or supported by the support structure, wherein a viewing direction of the camera is either tangential or orthogonal to the circular path; and/or
further comprising a support structure configured to carry and/or support a camera, wherein the support structure is configured to move along a circular path such that the camera is moved along the circular path.

10. A computer-implemented method for generating stereoscopic panoramas, the method being implemented in a computer system that includes one or more physical processors and physical storage media, the method comprising:
obtaining images based on captured images, wherein at least some of the captured images form a sequence of adjacent images;
determining interpolated images between adjacent images; and
determining a stereoscopic panorama based on obtained images and interpolated images.

11. The computer-implemented method of claim 10, wherein determining the interpolated images includes determining flow fields between related pixels in adjacent images.

12. The computer-implemented method of claim 11, wherein determining the flow fields includes compensating for transformations of pixel grids for the obtained images; and/or
wherein determining the stereoscopic panorama is based on determining vertical strips of the obtained images and further based on blending the vertical strips, wherein blending is based on the determined flow fields.

13. The computer-implemented method of any of claims 10 to 12, wherein the obtained images are based on compensated images, the method further comprising:
determining the compensated images based on the captured images by compensating for perspective distortion caused by a shape of a lens of a camera used to capture the captured images.

14. The computer-implemented method of claim 13, wherein the obtained images are based on orientation-corrected images, the method further comprising:
determining the orientation-corrected images based on the compensated images by compensating for one or

both of a non-circular camera trajectory and/or a non-parallel principal axis used during image capture; optionally wherein the obtained images are based on re-projected images, the method further comprising:

5 determining the re-projected images based on the orientation-corrected images by re-projecting orientation-corrected images onto a cylindrical imaging surface, thereby reducing vertical parallax.

15. The computer-implemented method of claim 14, wherein the obtained images are based on aligned images, the method further comprising:

10 determining the aligned images based on alignment of the re-projected images and/or vertical fragments/strips thereof, wherein the alignment is accomplished by projecting the re-projected images into a planar two-dimensional surface such that one or both of rotational and/or positional drift between adjacent images is reduced; optionally

15 wherein the alignment causes a transformation of pixel grids of the aligned images, and wherein determining the flow fields is accomplished by compensating for the transformation caused by the alignment.

16. The computer-implemented method of claim 15, further comprising:

20 capturing images, by a camera, wherein a viewing direction of the camera is either tangential or orthogonal to a circular path, and wherein the obtained images are based on the images captured by the camera.

25

30

35

40

45

50

55

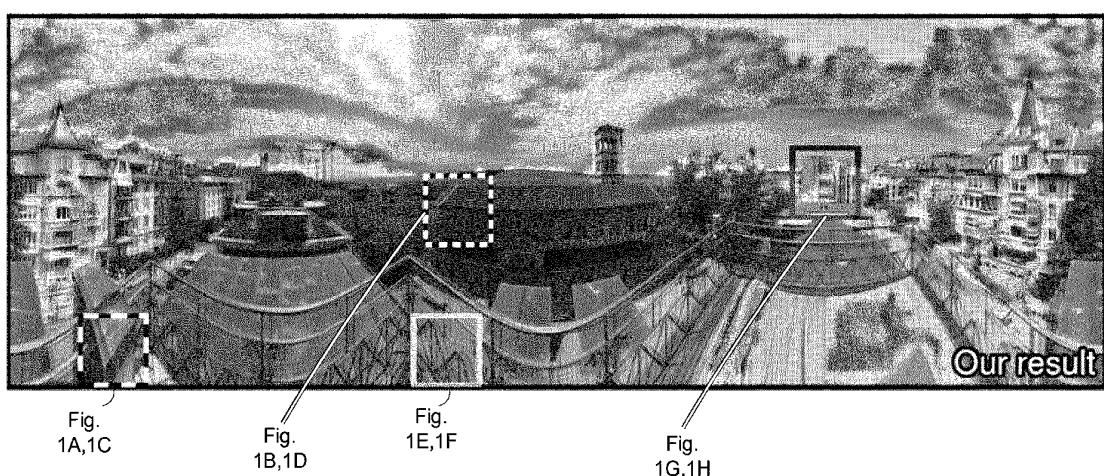


FIG. 1

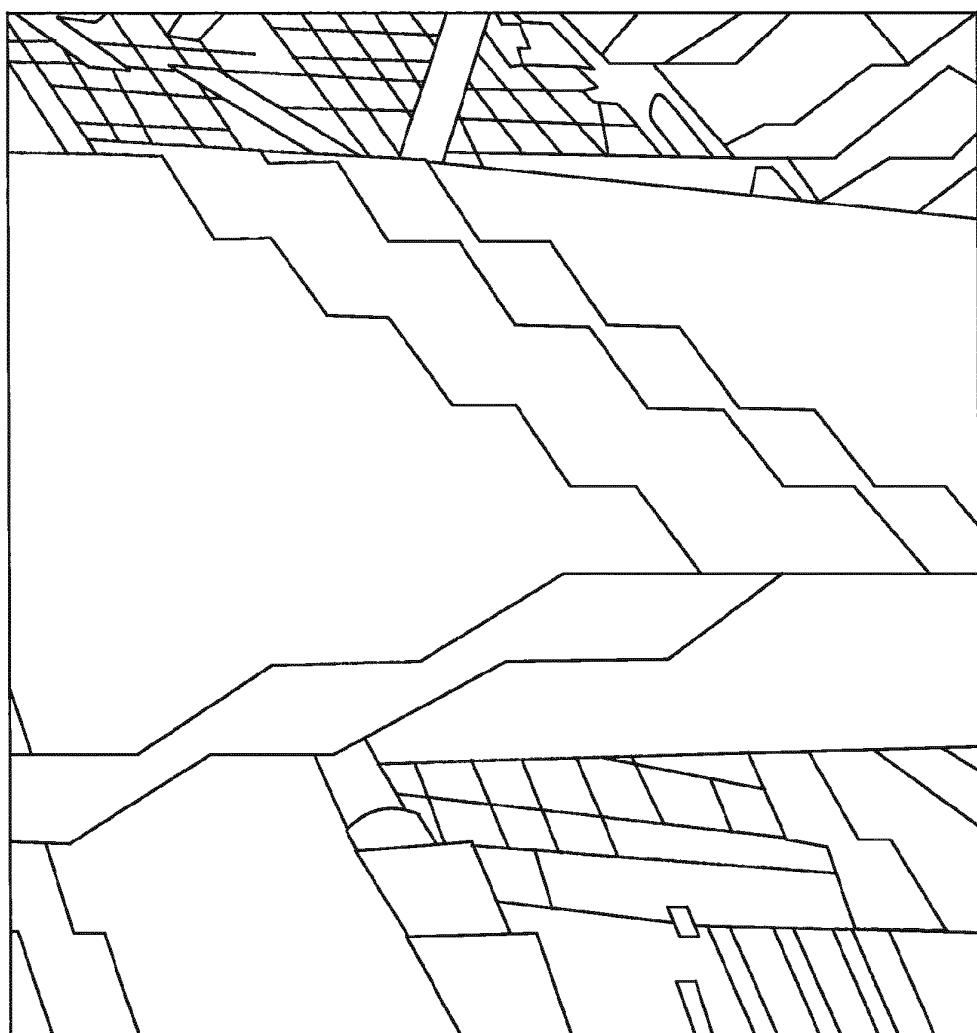


FIG. 1A

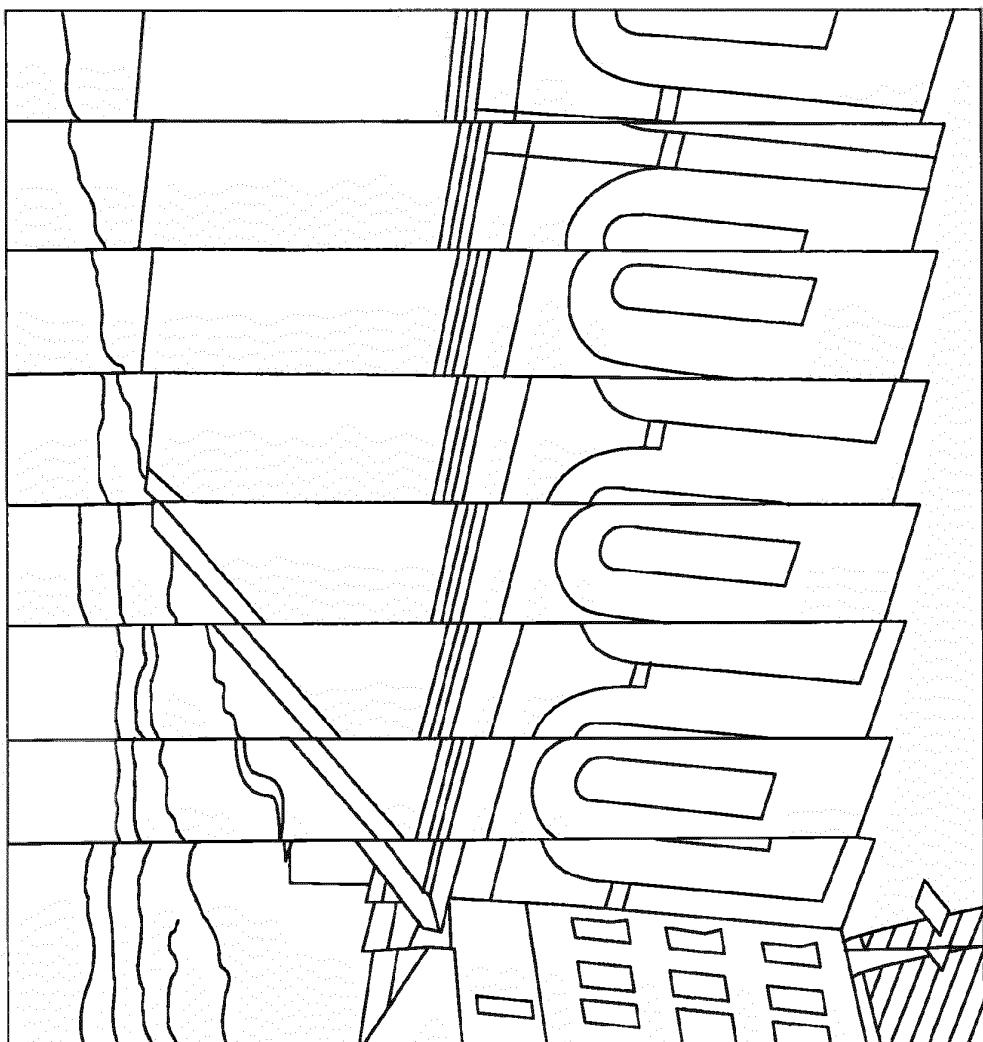


FIG. 1B

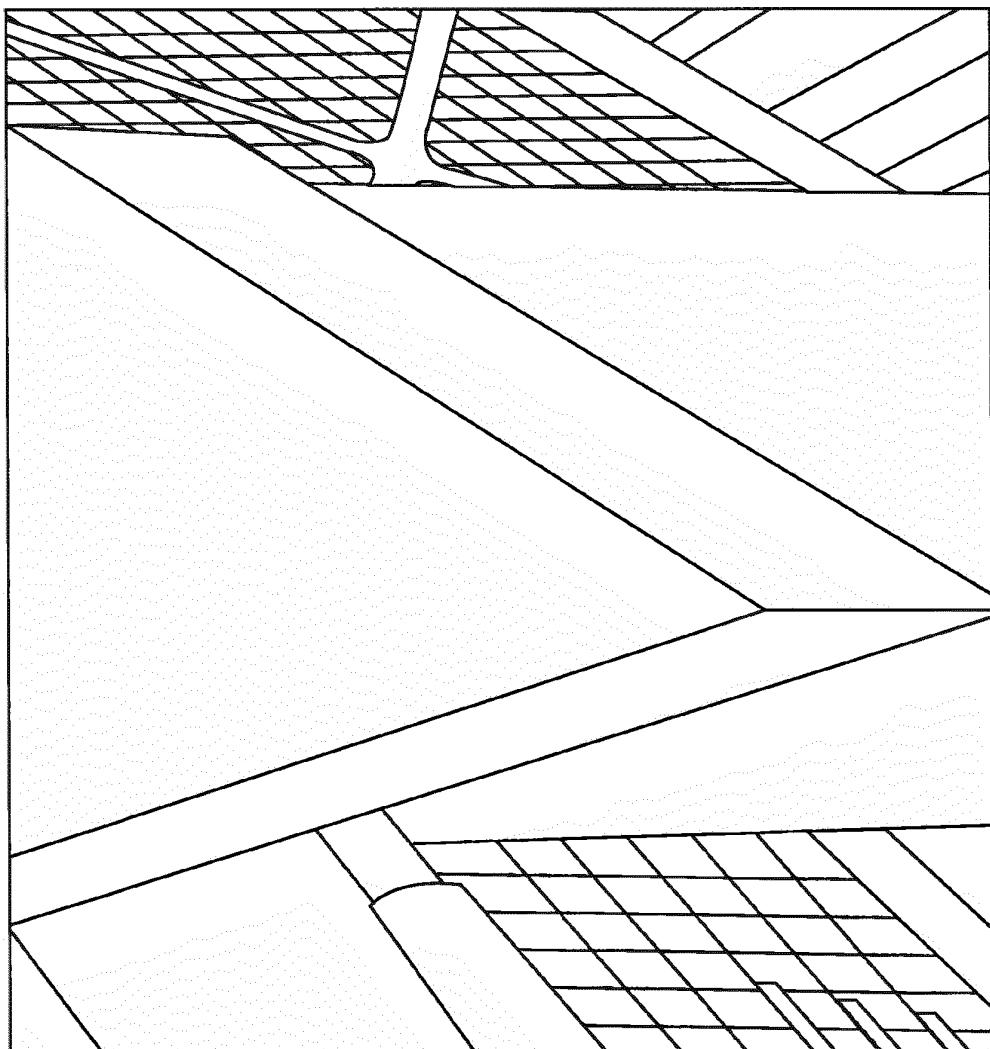


FIG. 1C

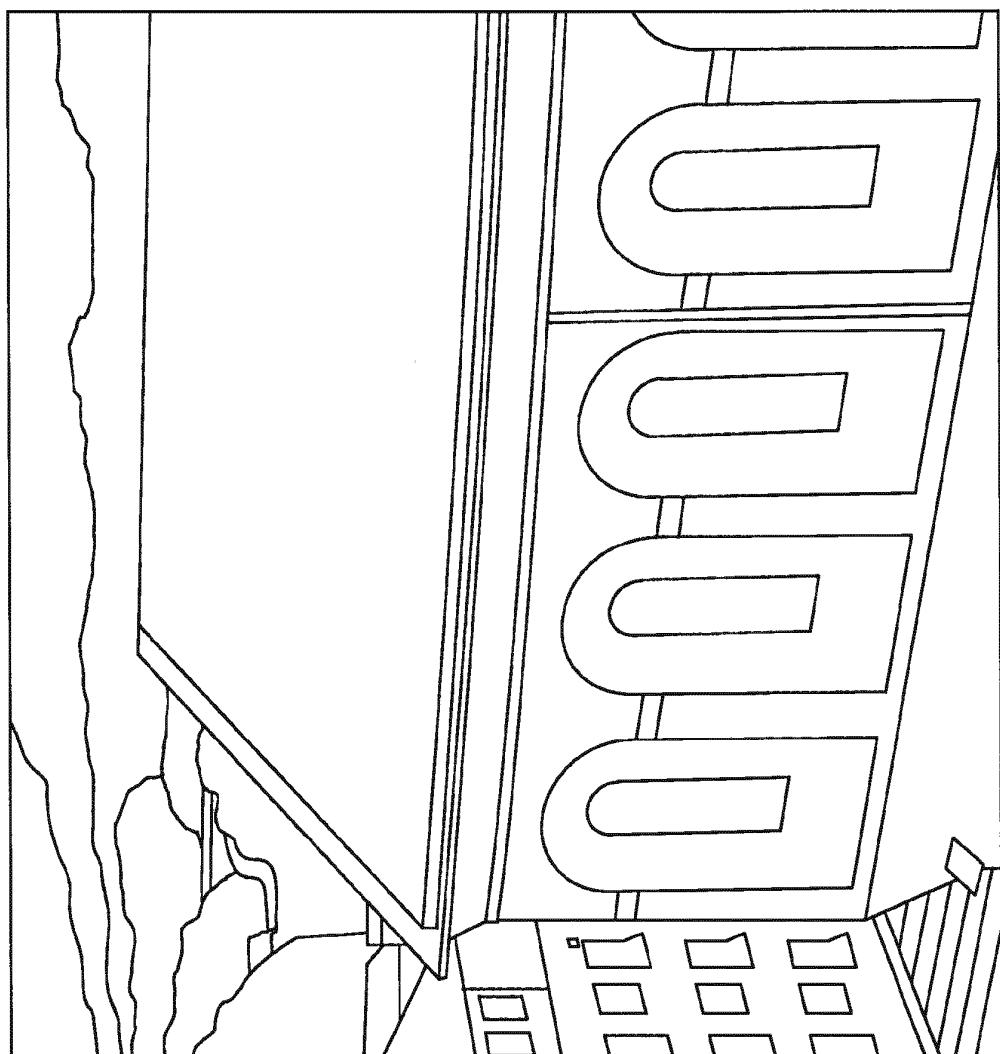


FIG. 1D

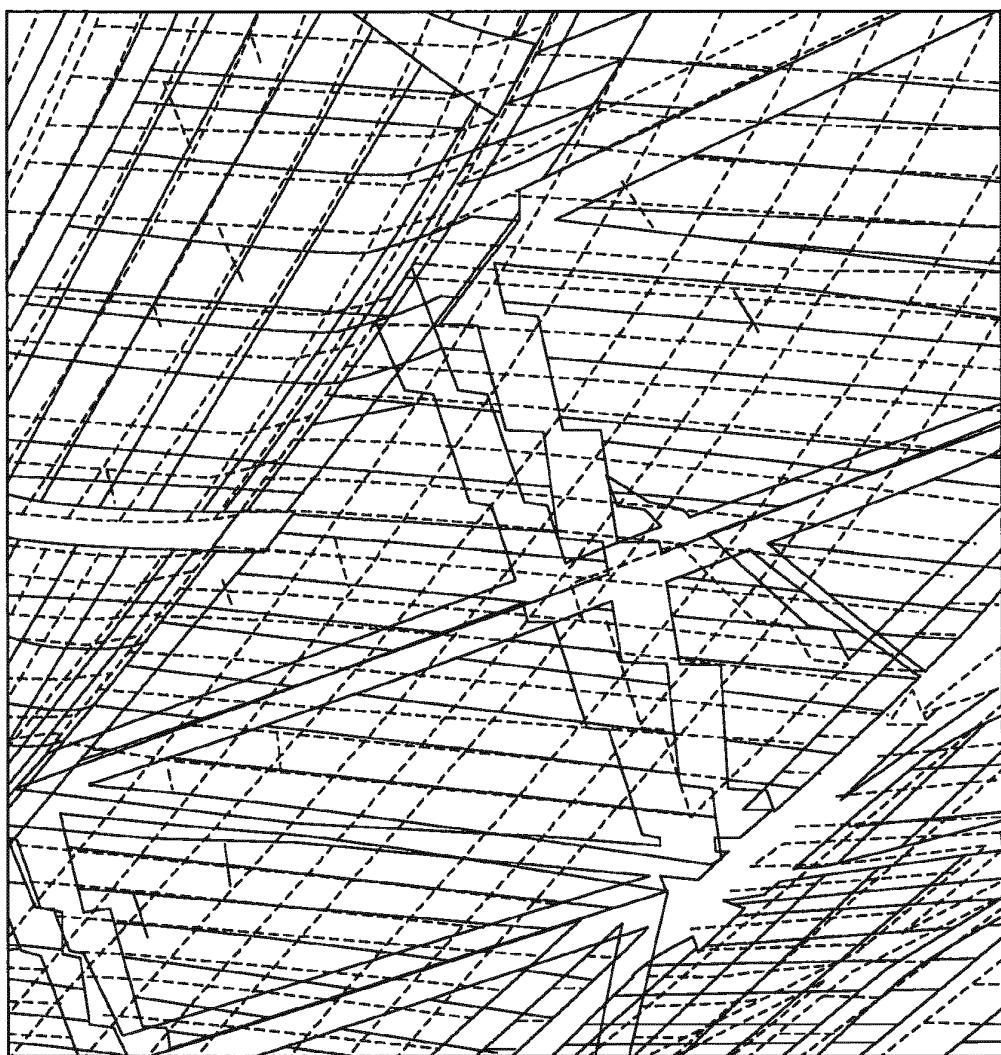


FIG. 1E

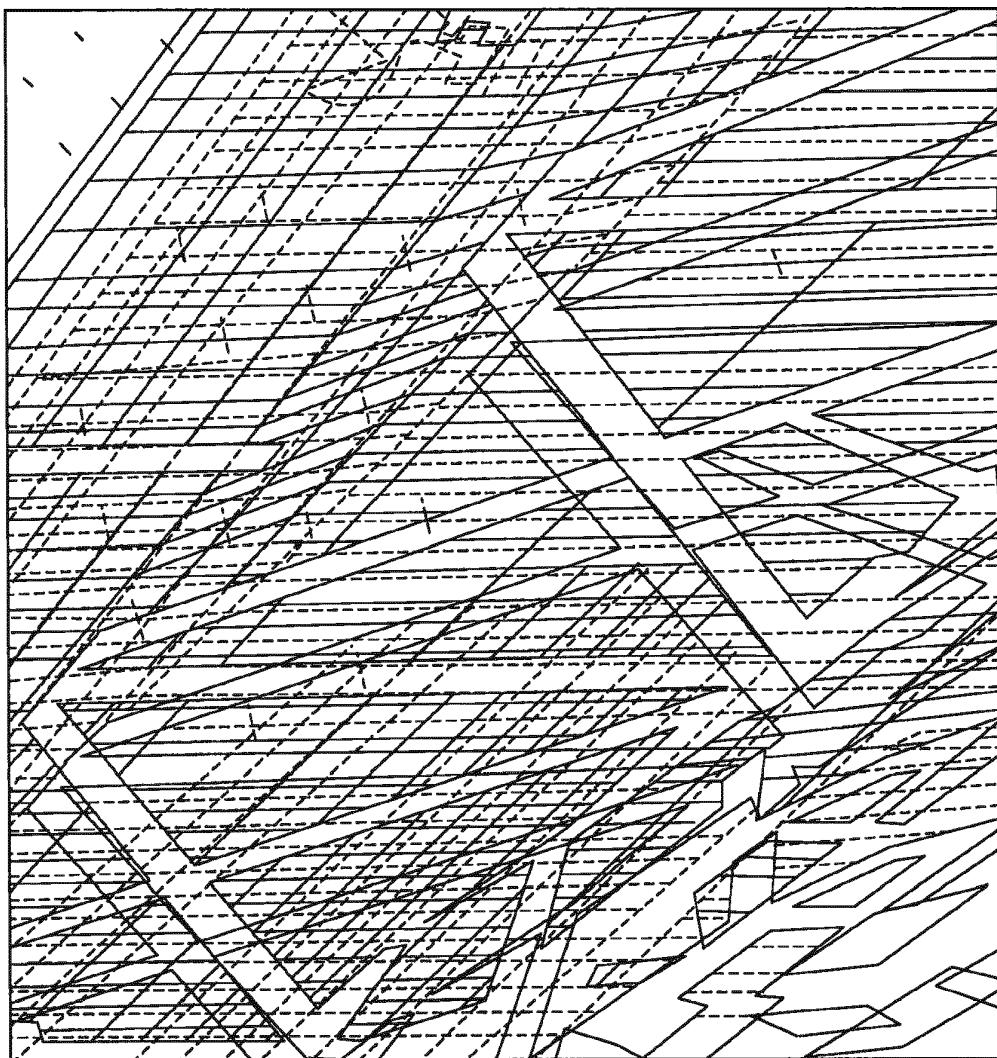


FIG. 1F

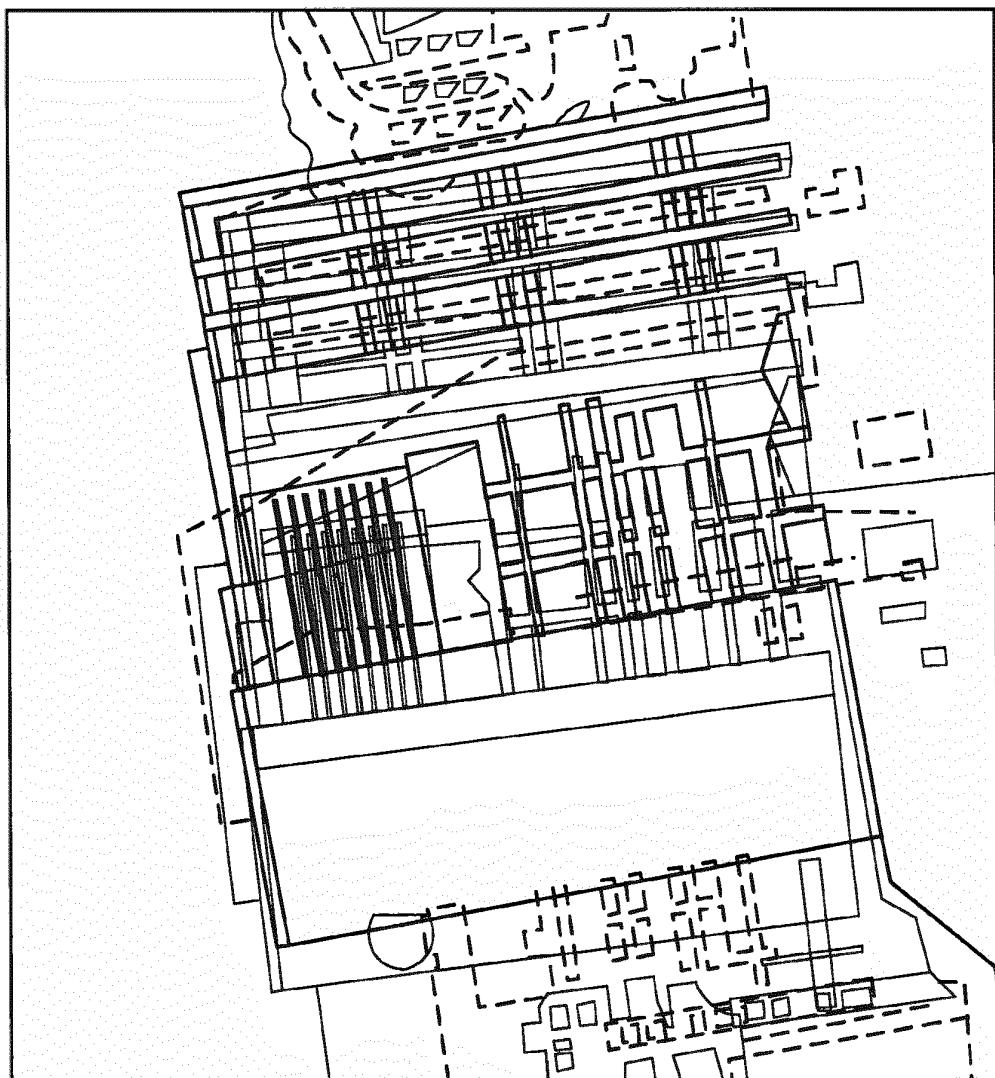


FIG. 1G

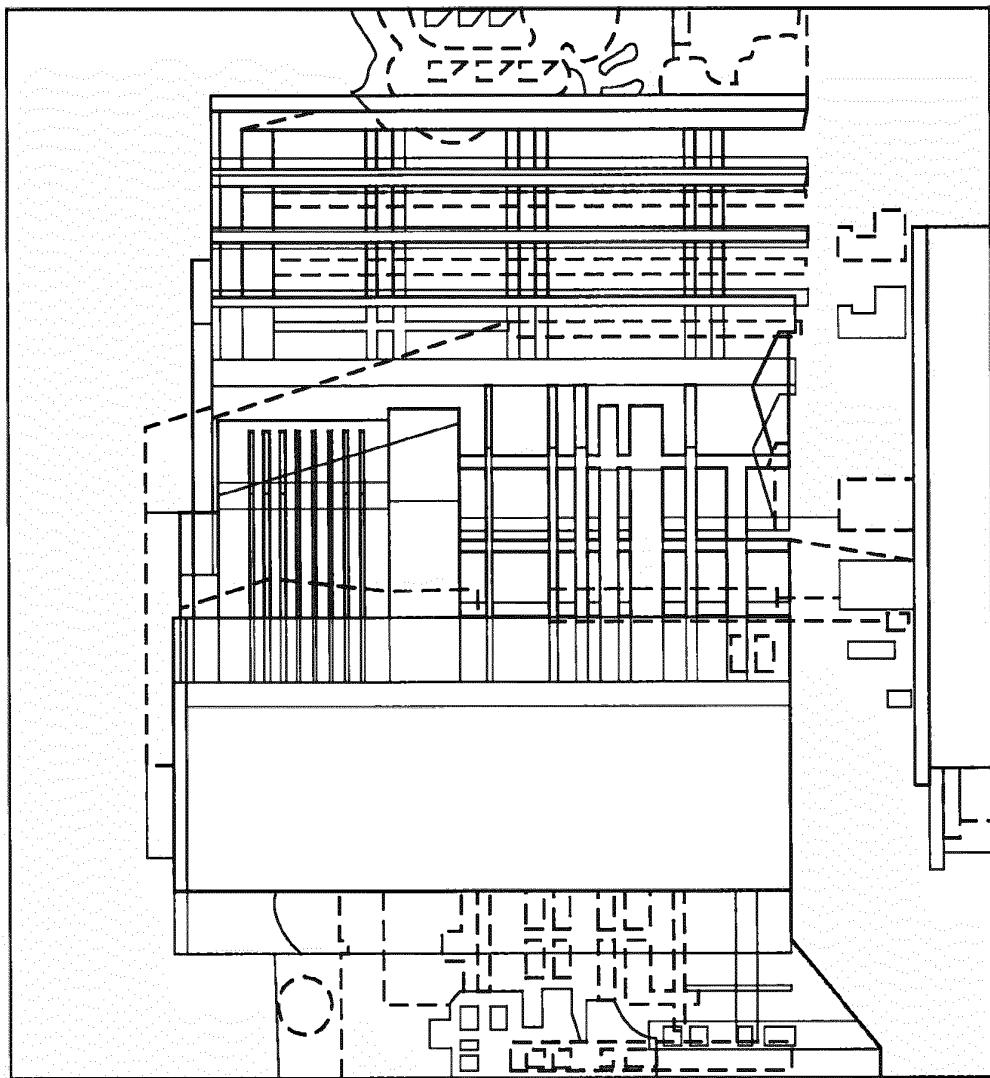


FIG. 1H

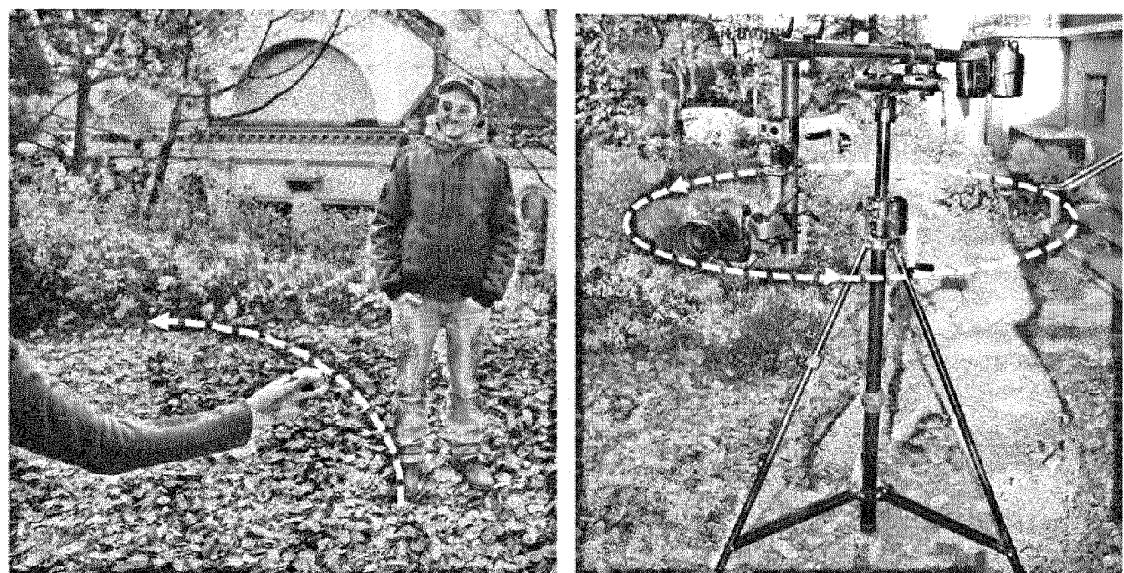


FIG. 2



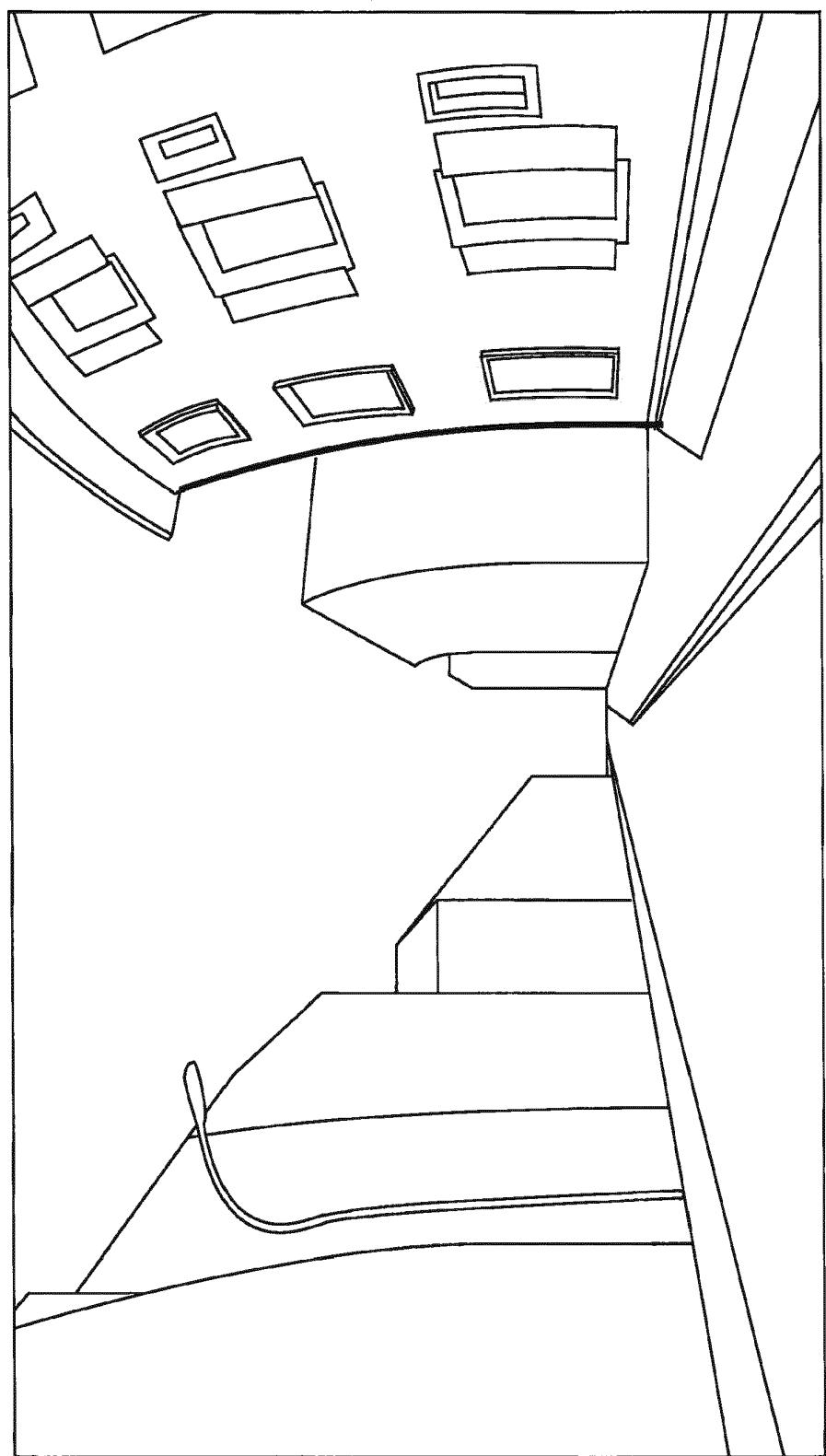


FIG. 3A

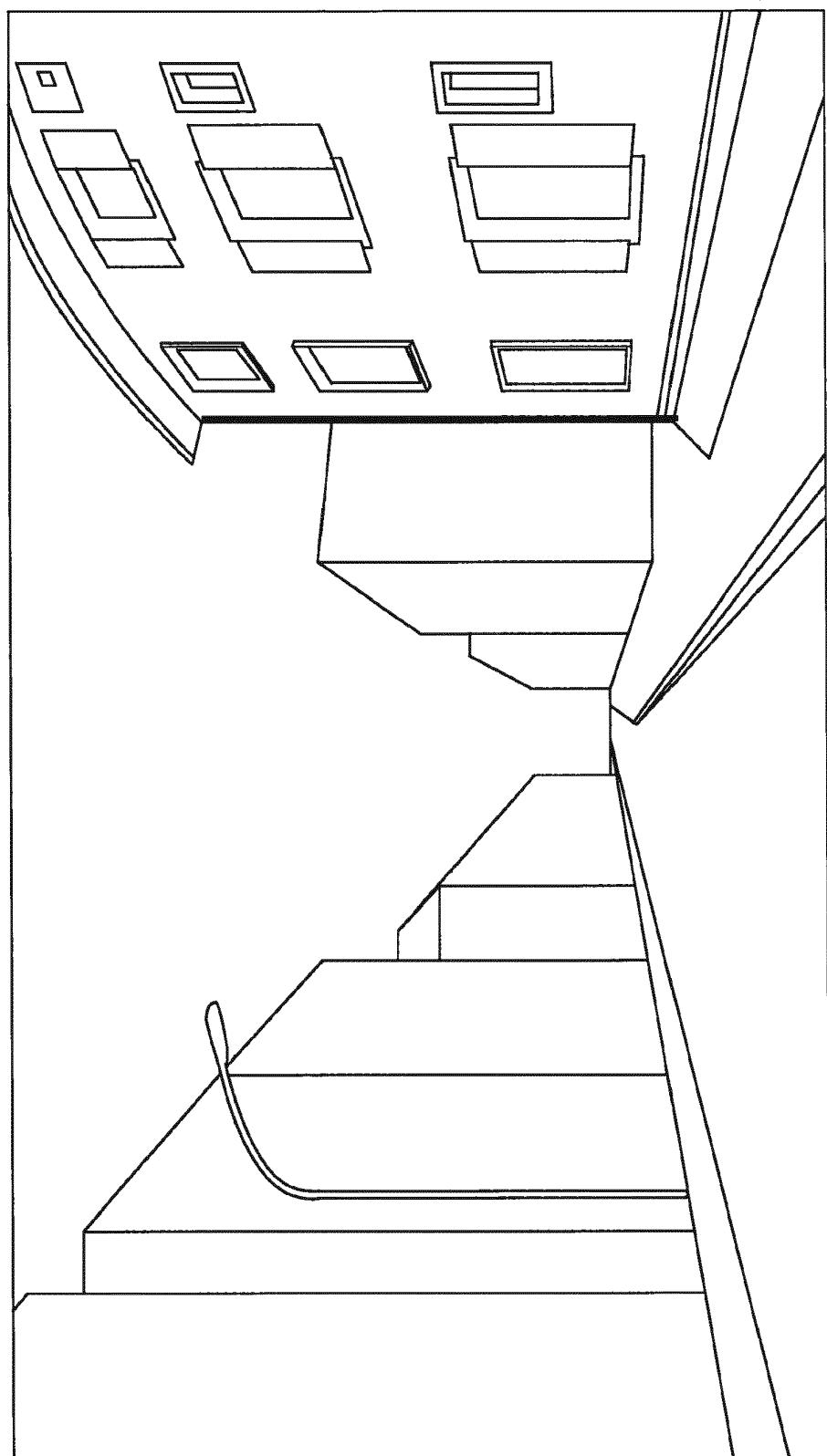


FIG. 3B

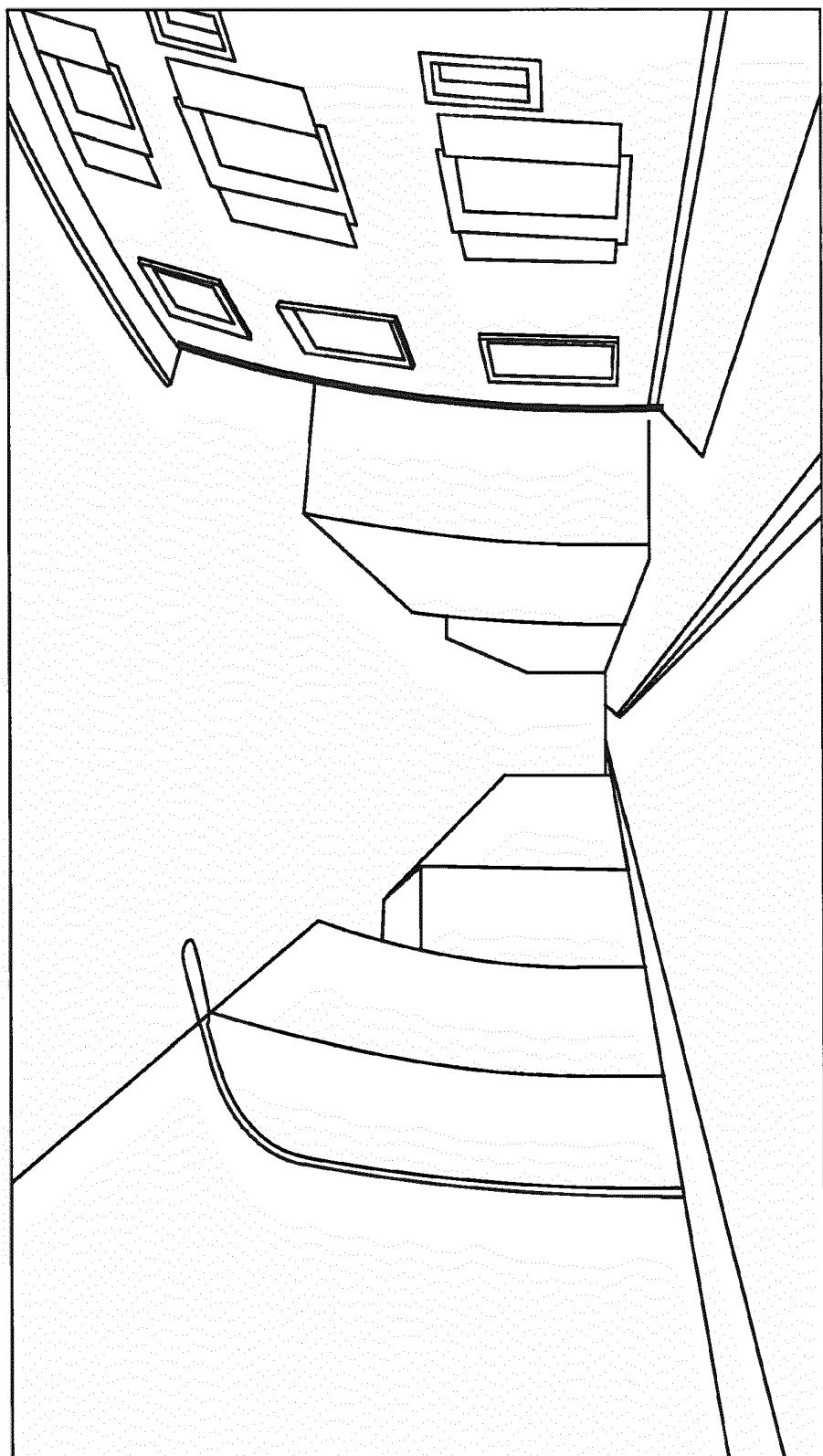


FIG. 3C

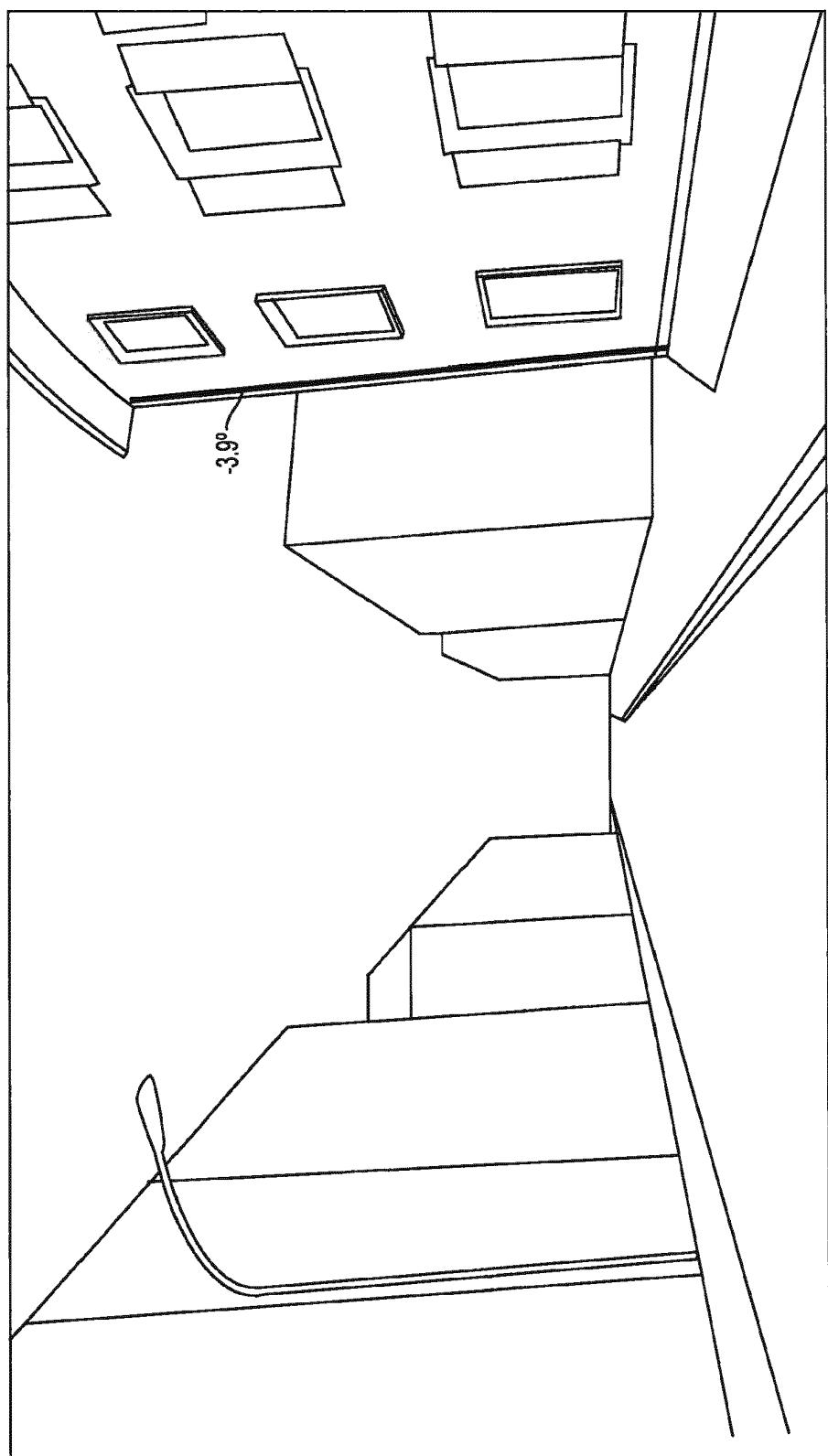


FIG. 3D

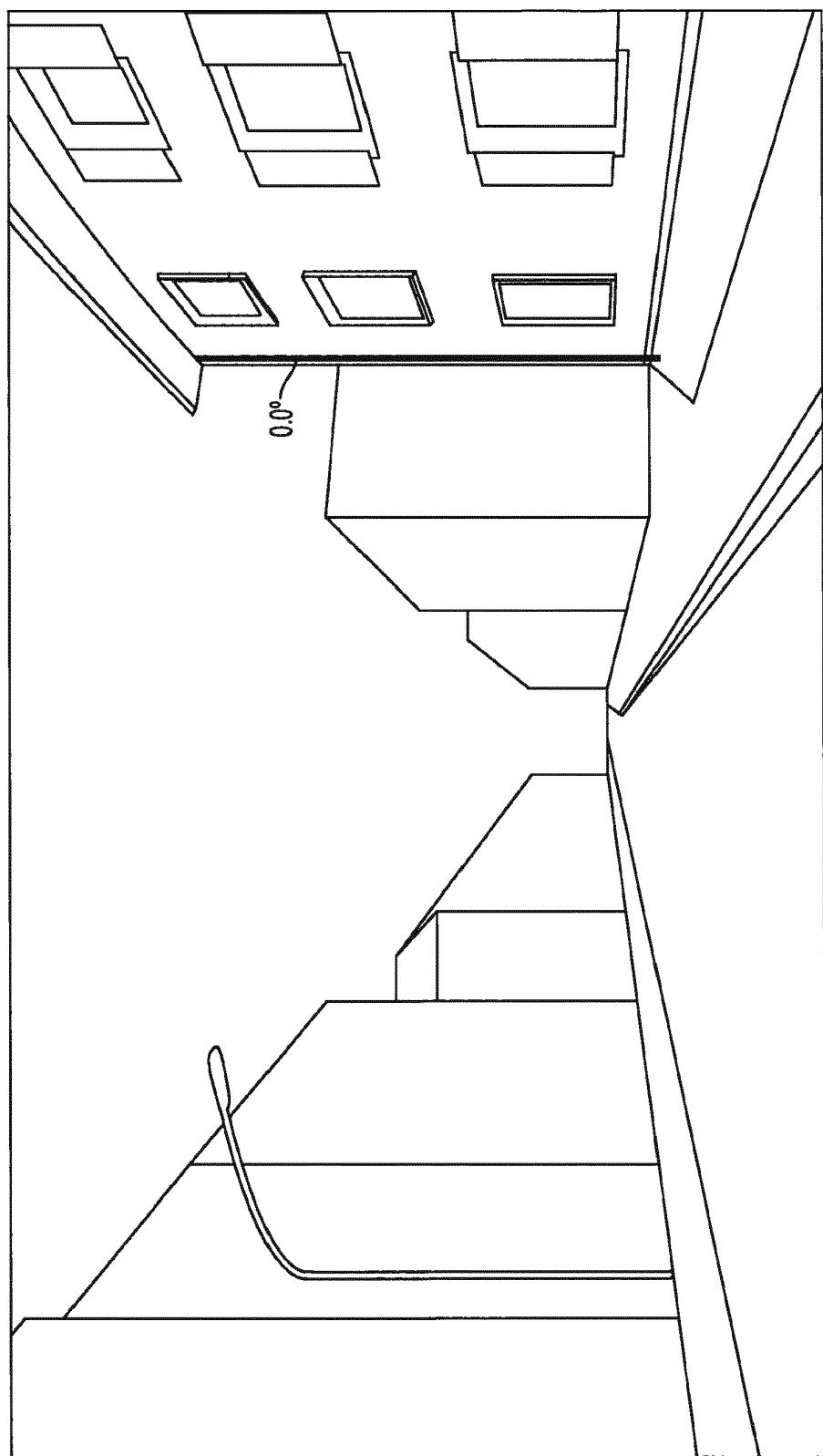


FIG. 3E

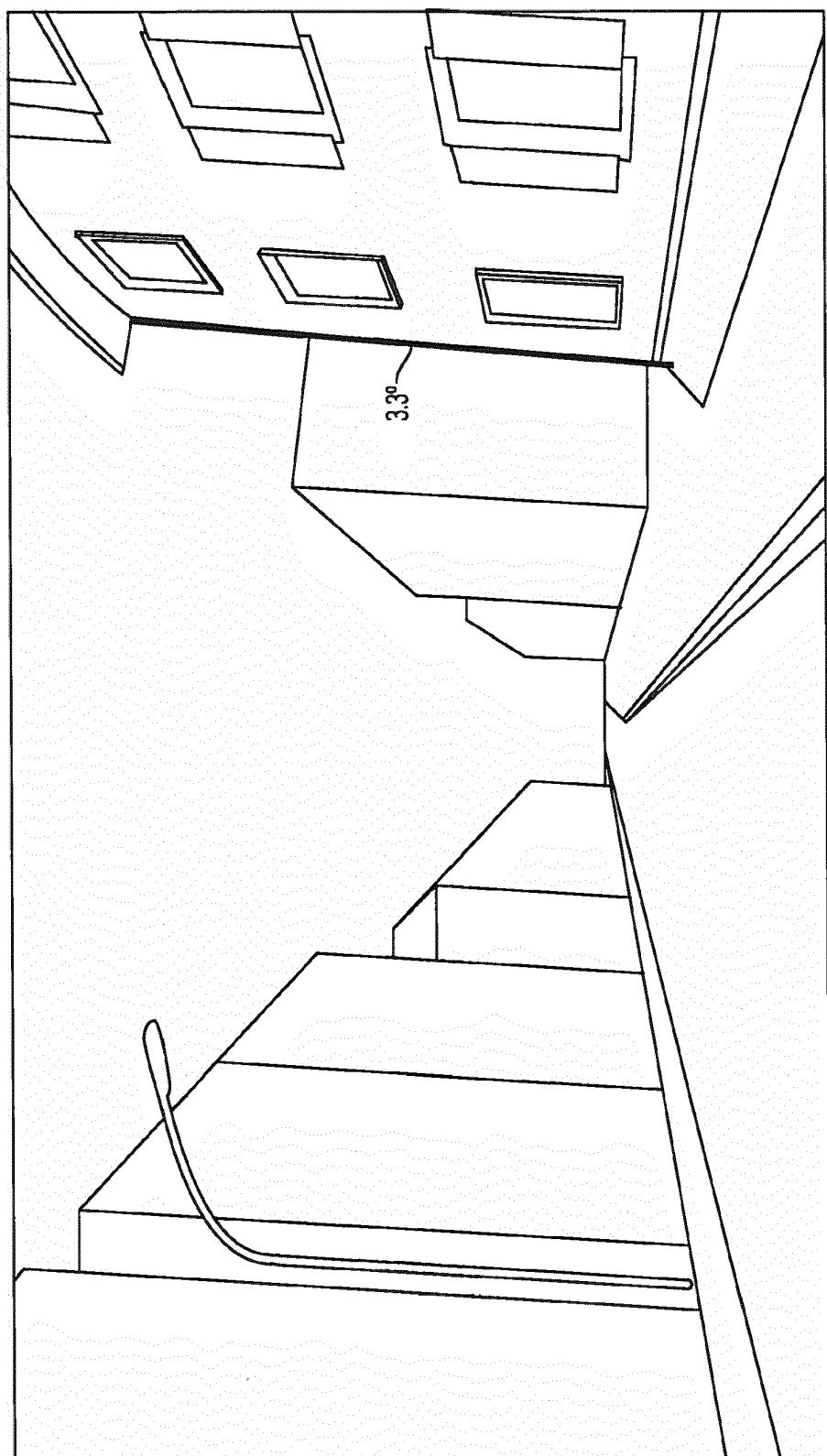


FIG. 3F

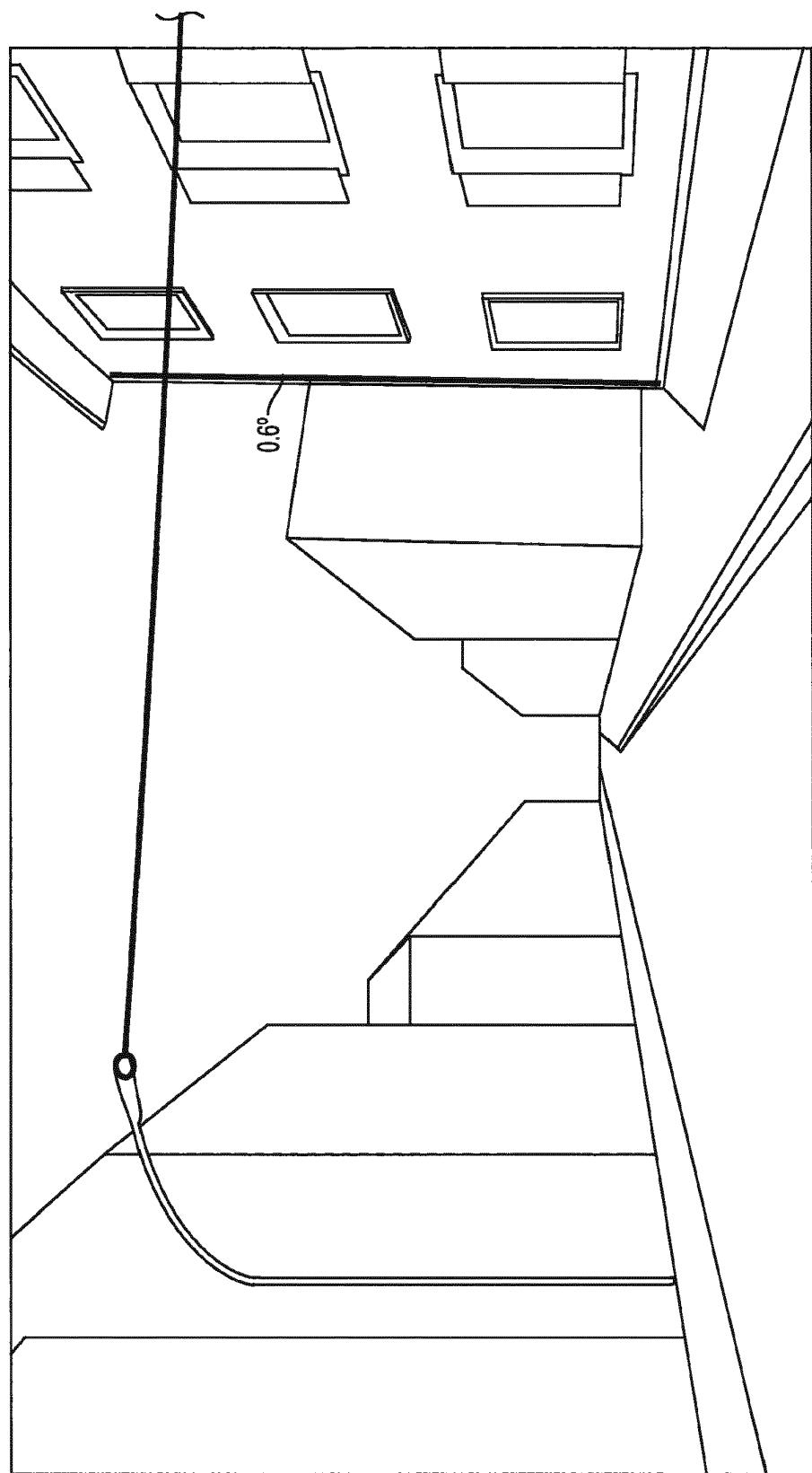


FIG. 3G

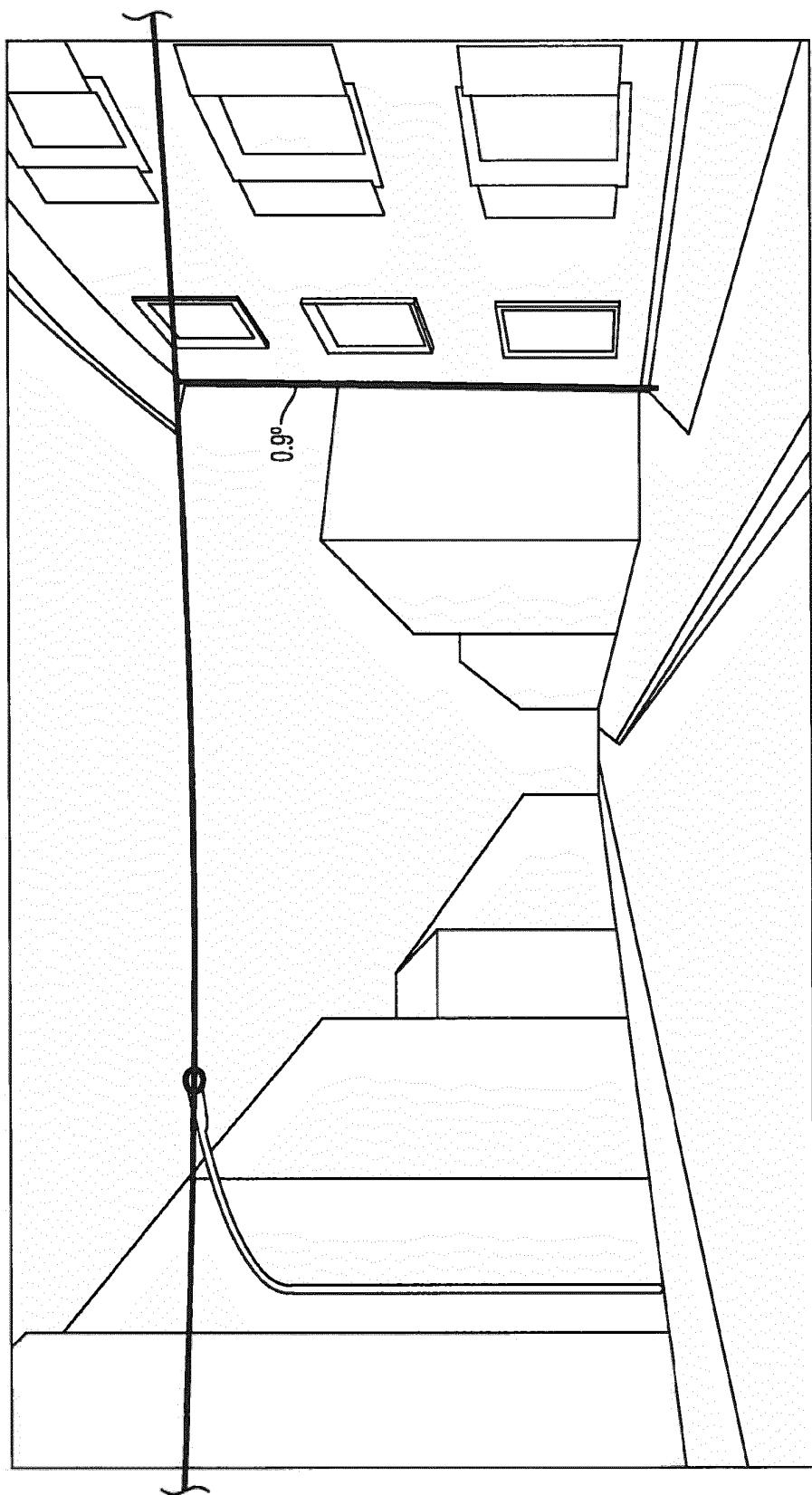


FIG. 3H

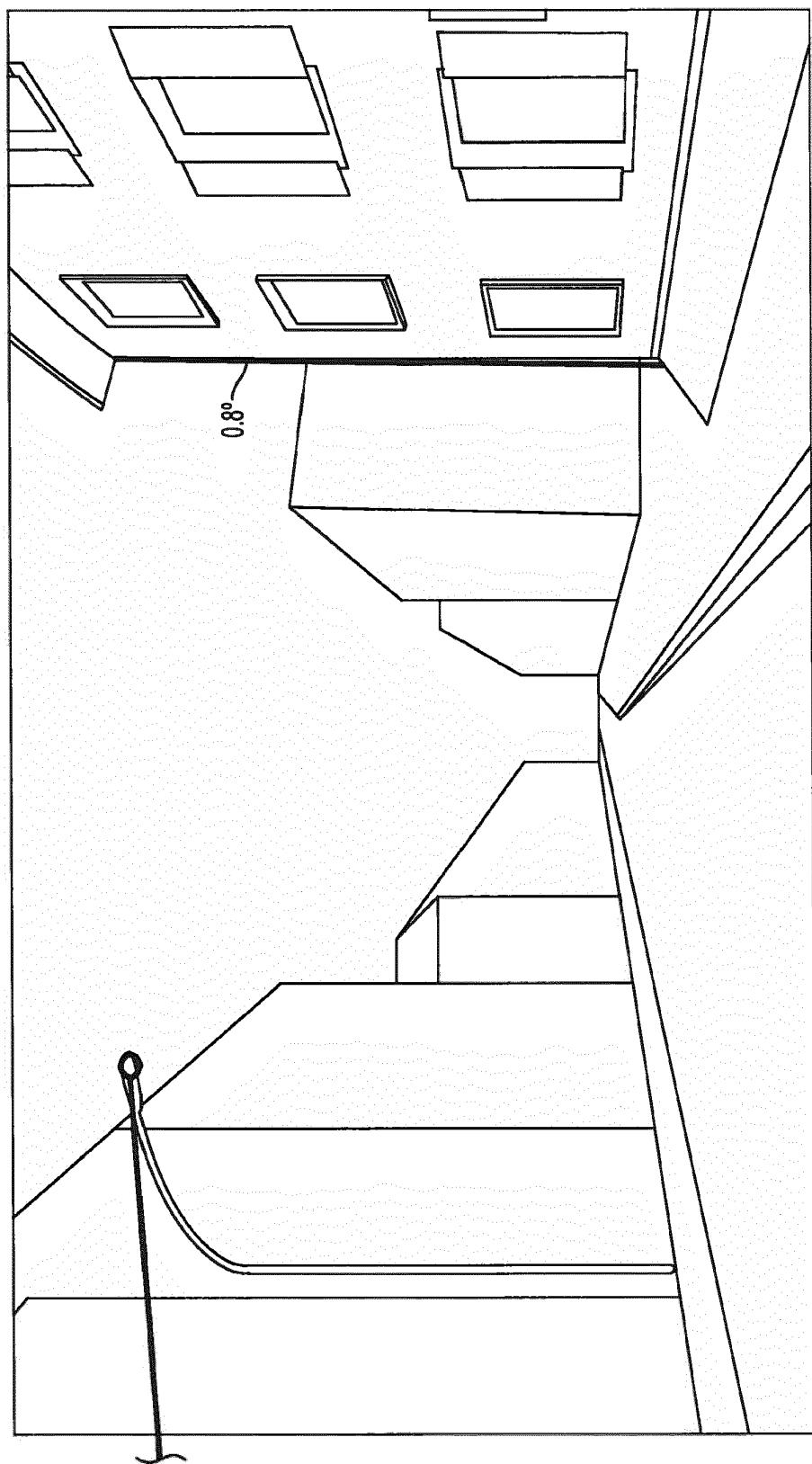


FIG. 31

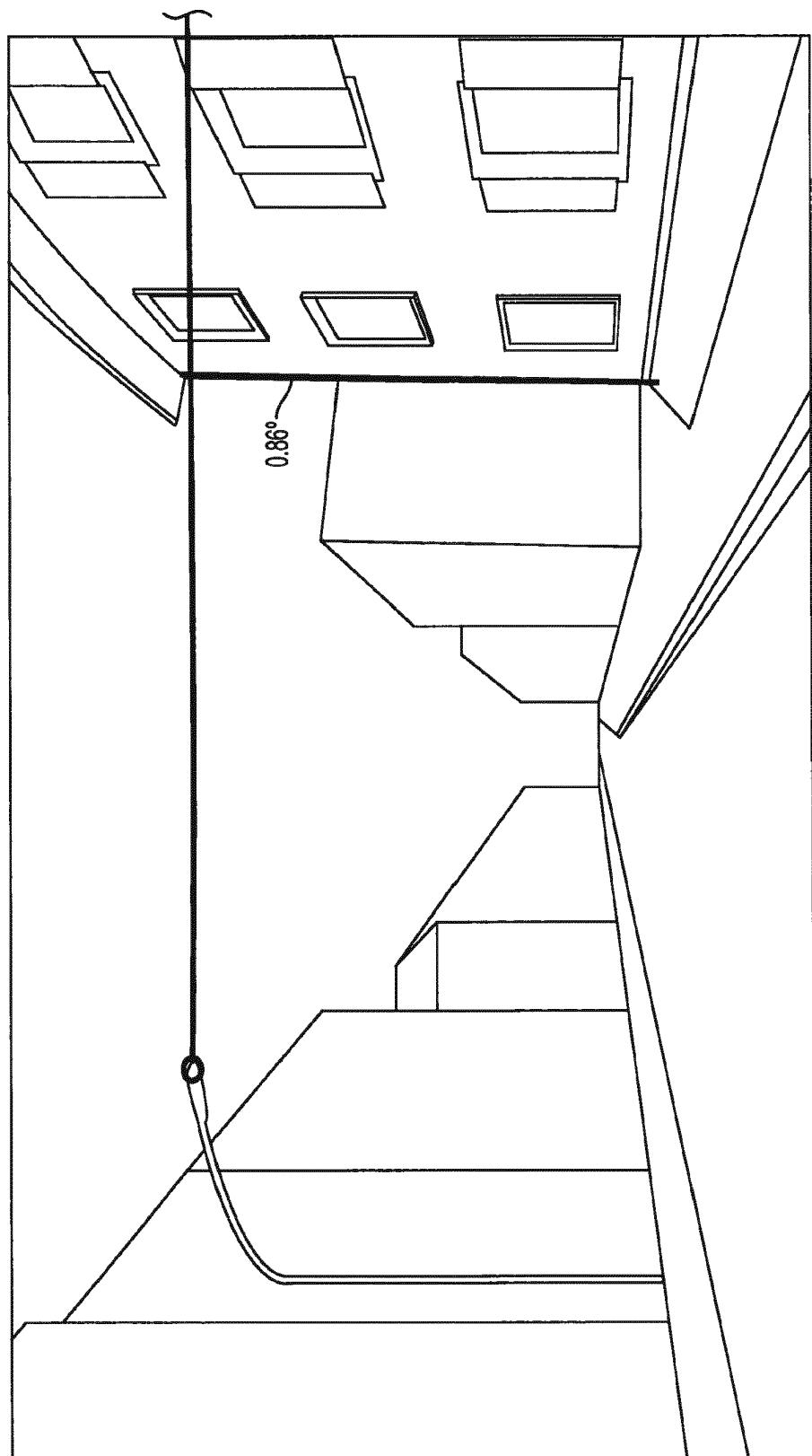


FIG. 3J

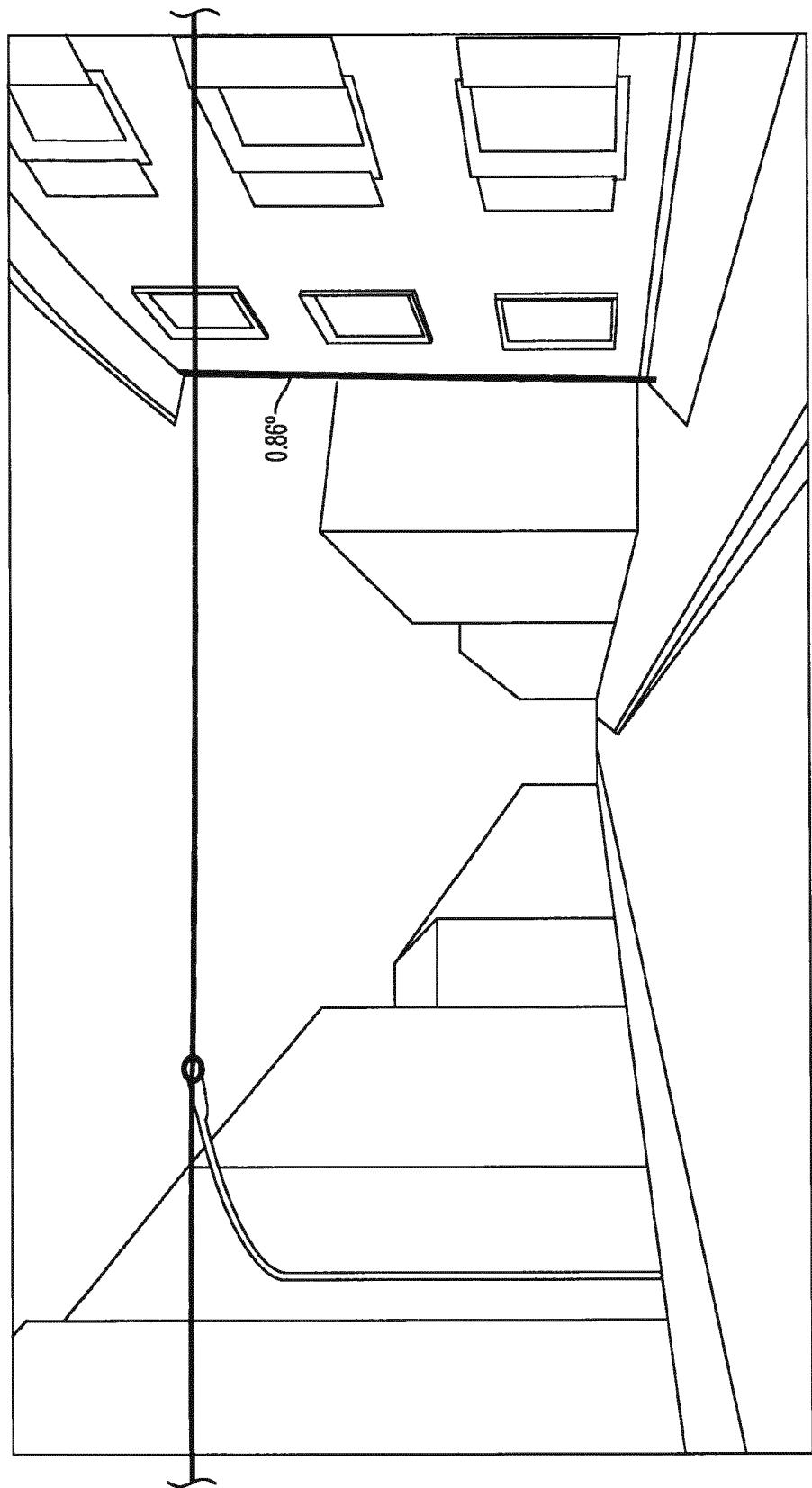


FIG. 3K

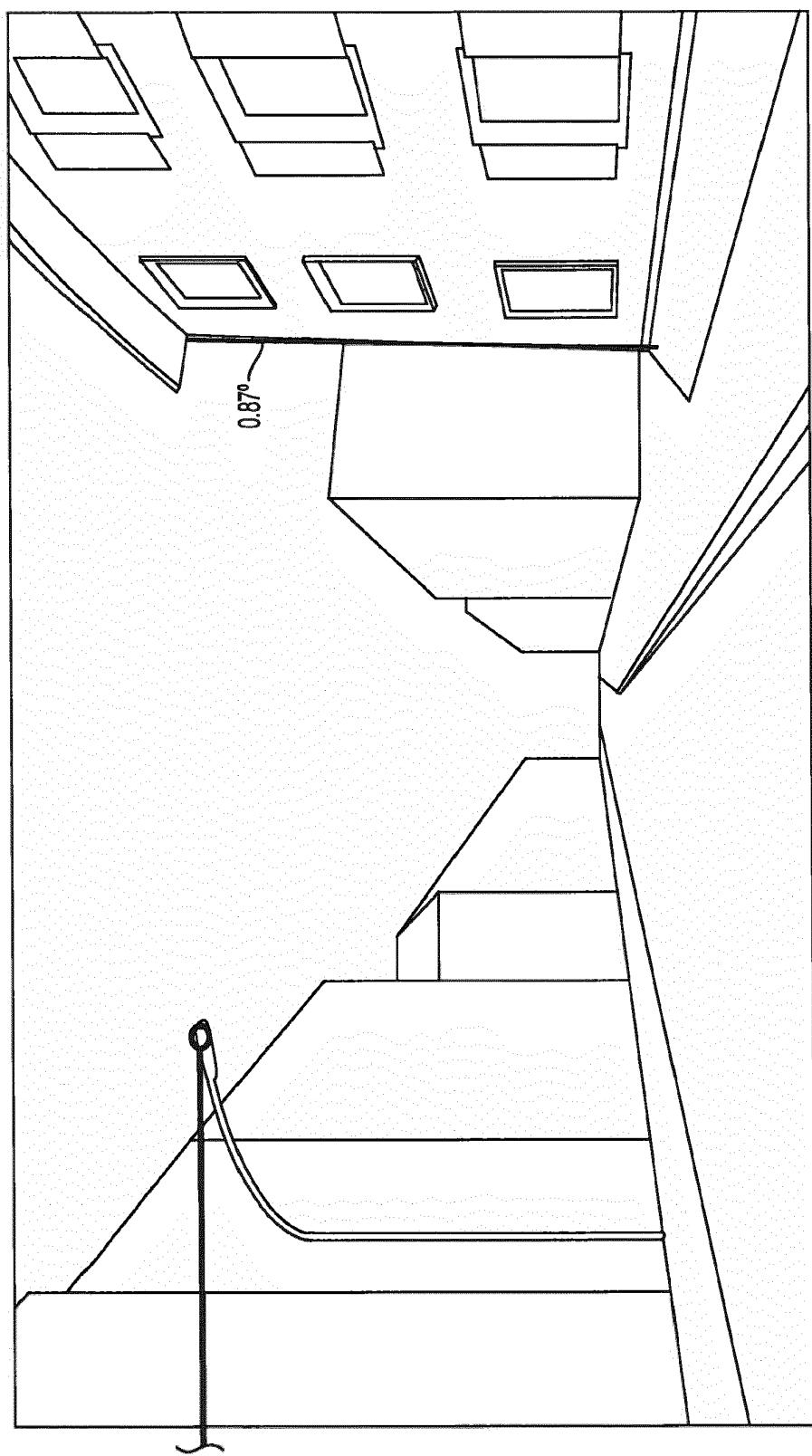


FIG. 3L

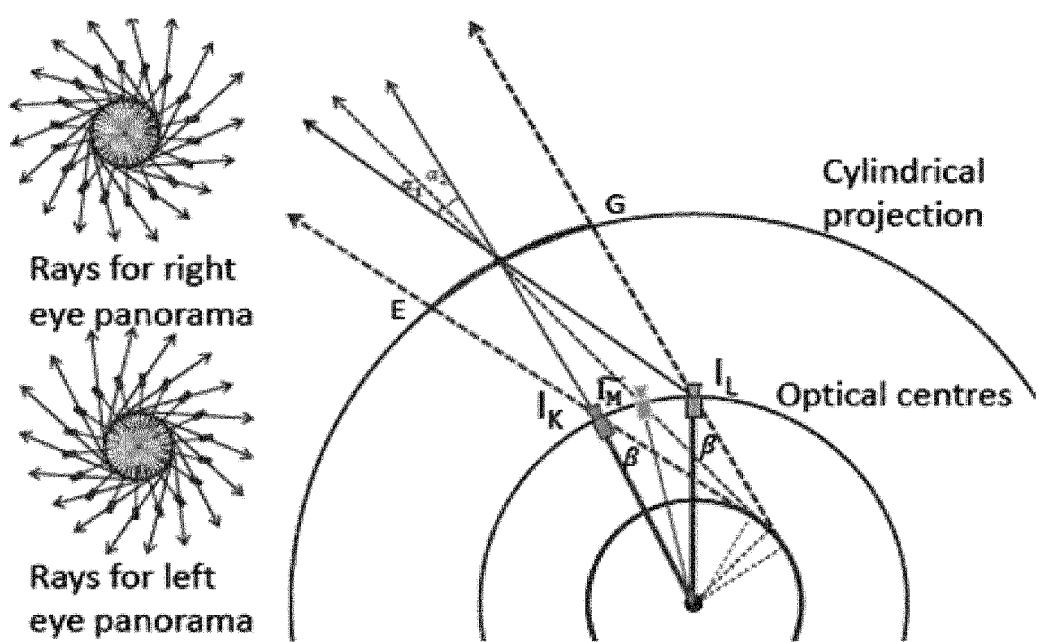


FIG. 4

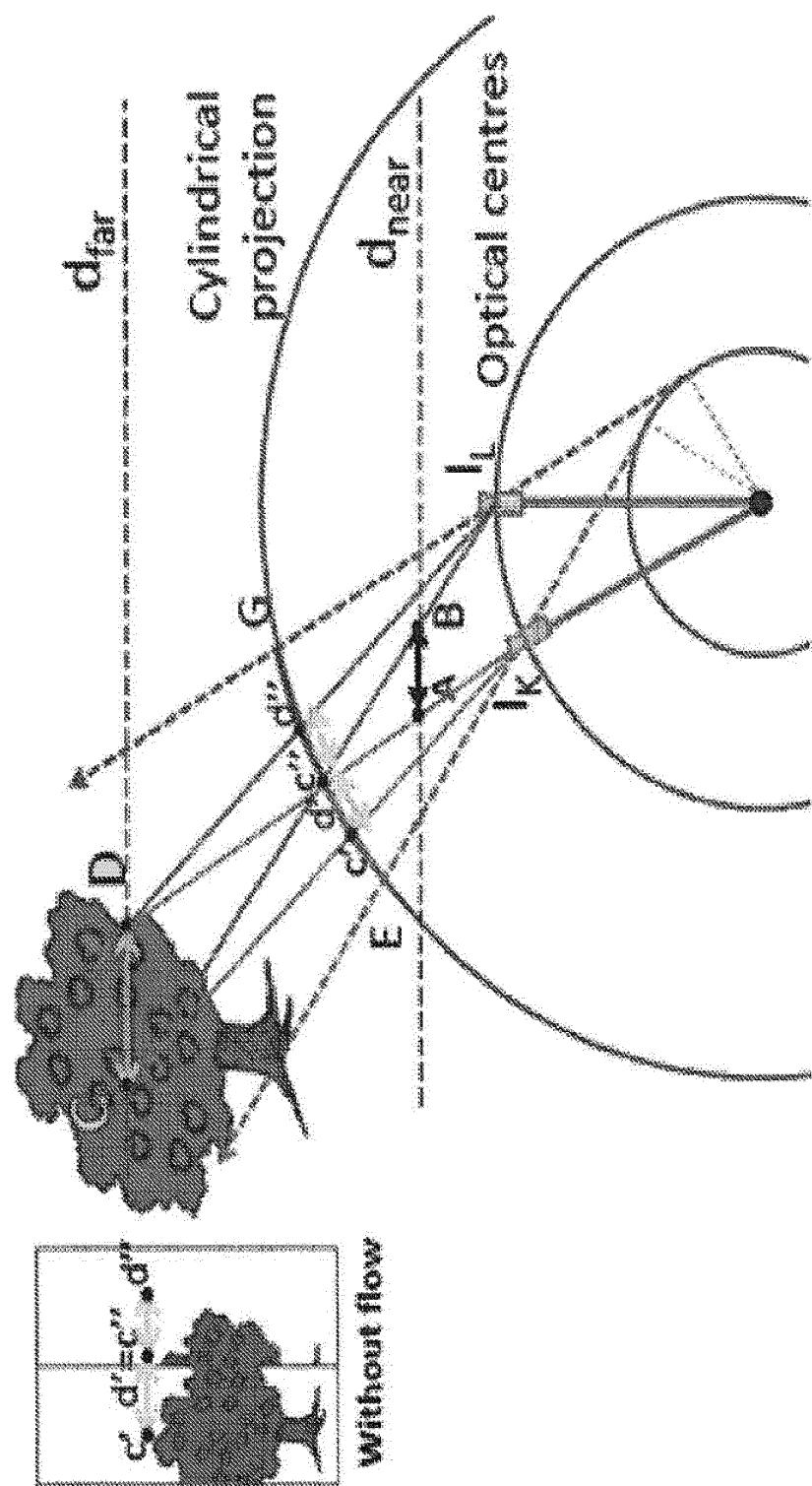


FIG. 5

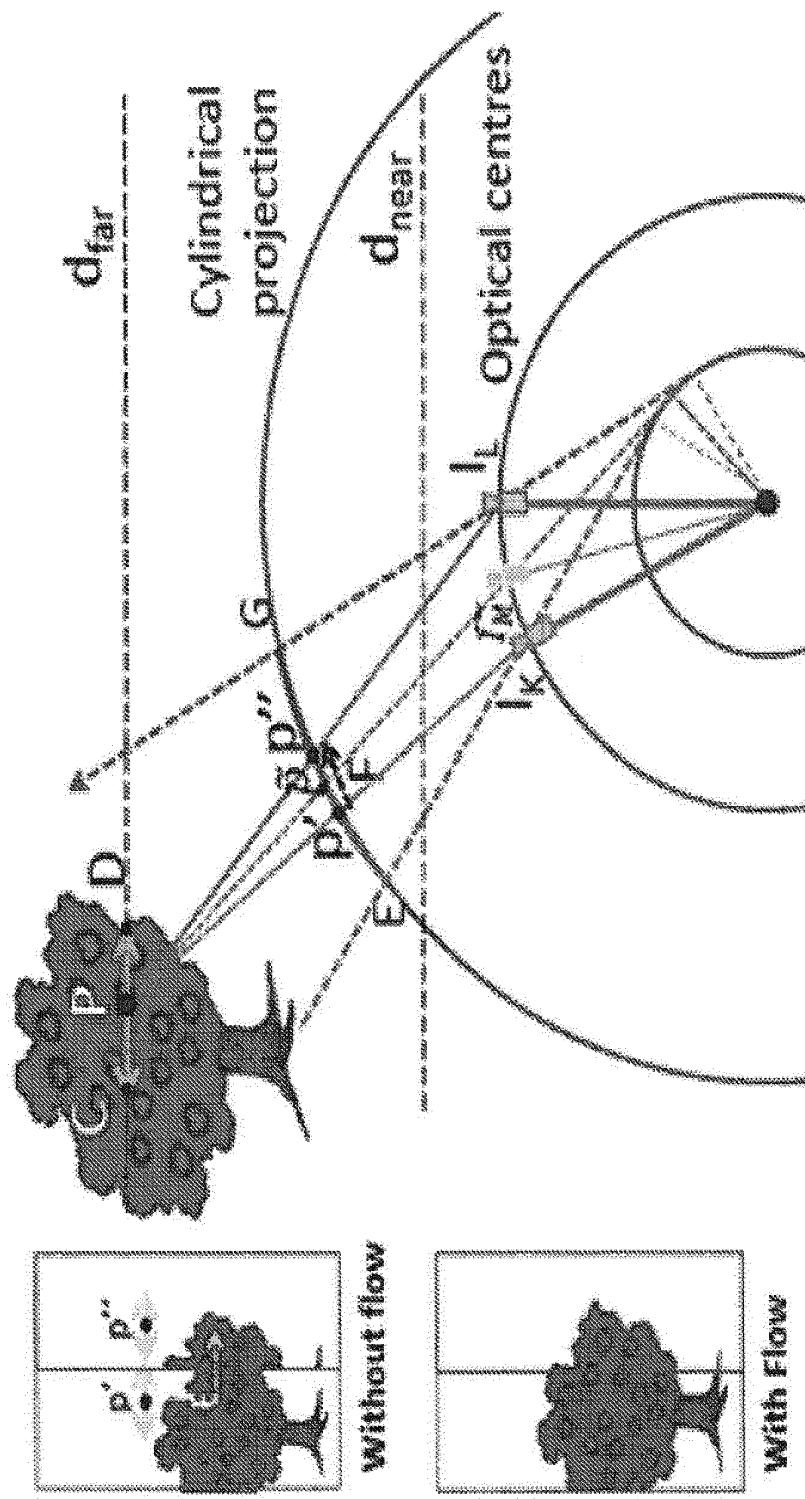


FIG. 6

Dataset	rotational drift		vertical drift	
	IB	Ours	IB	Ours
rooftop	98.7°	0.0043°	1.57°	0.0002°
street	82.7°	0.0154°	3.07°	0.0113°
mountain	66.8°	0.0600°	6.26°	0.0100°

FIG. 7

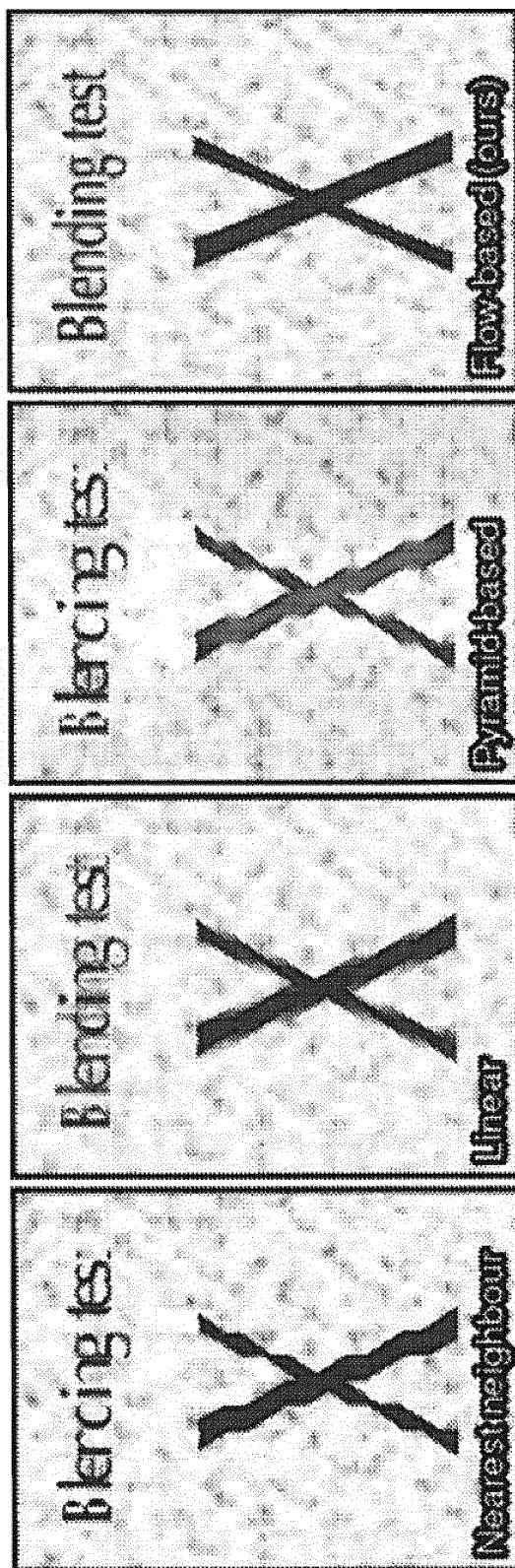


FIG. 8

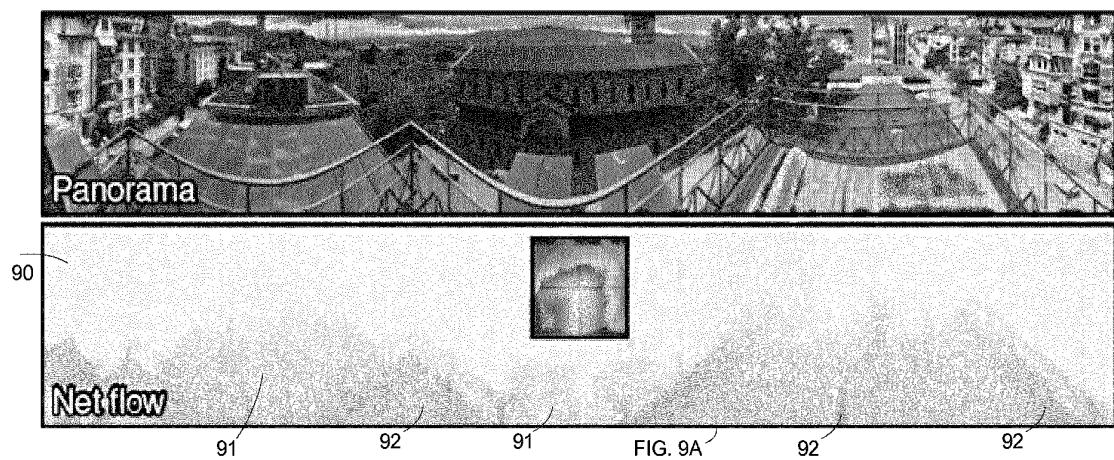


FIG. 9

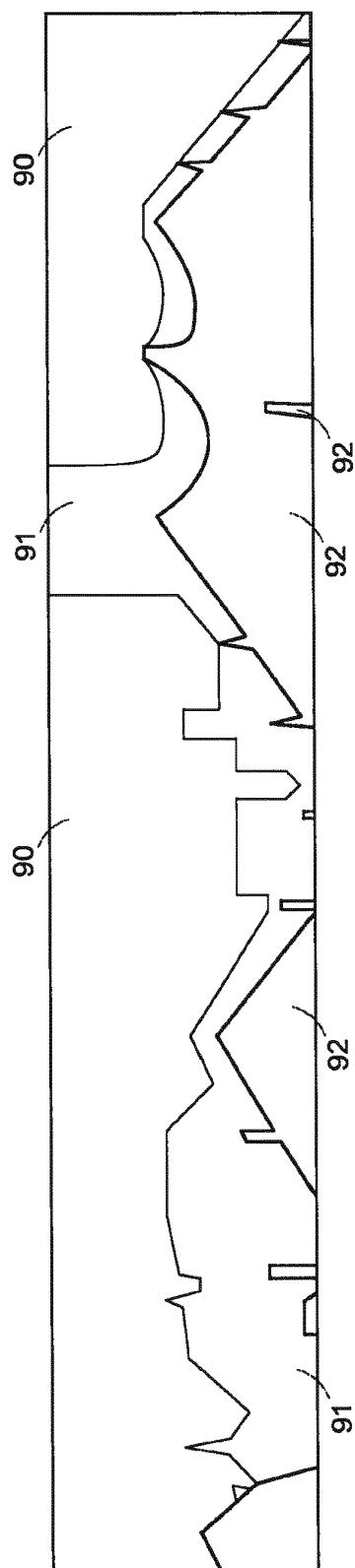


FIG. 9A

EP 2 793 187 A1



FIG. 10

EP 2 793 187 A1

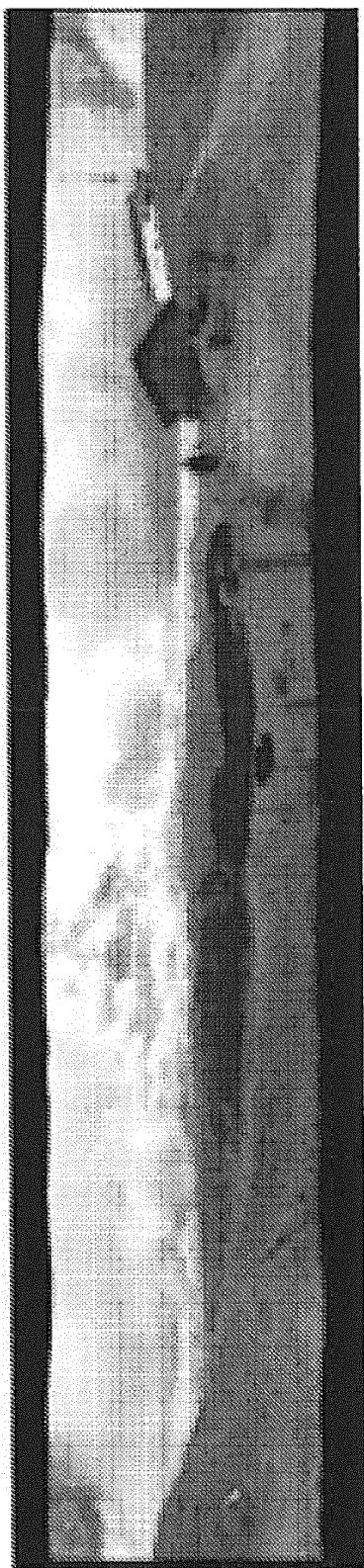


FIG. 11



FIG. 12

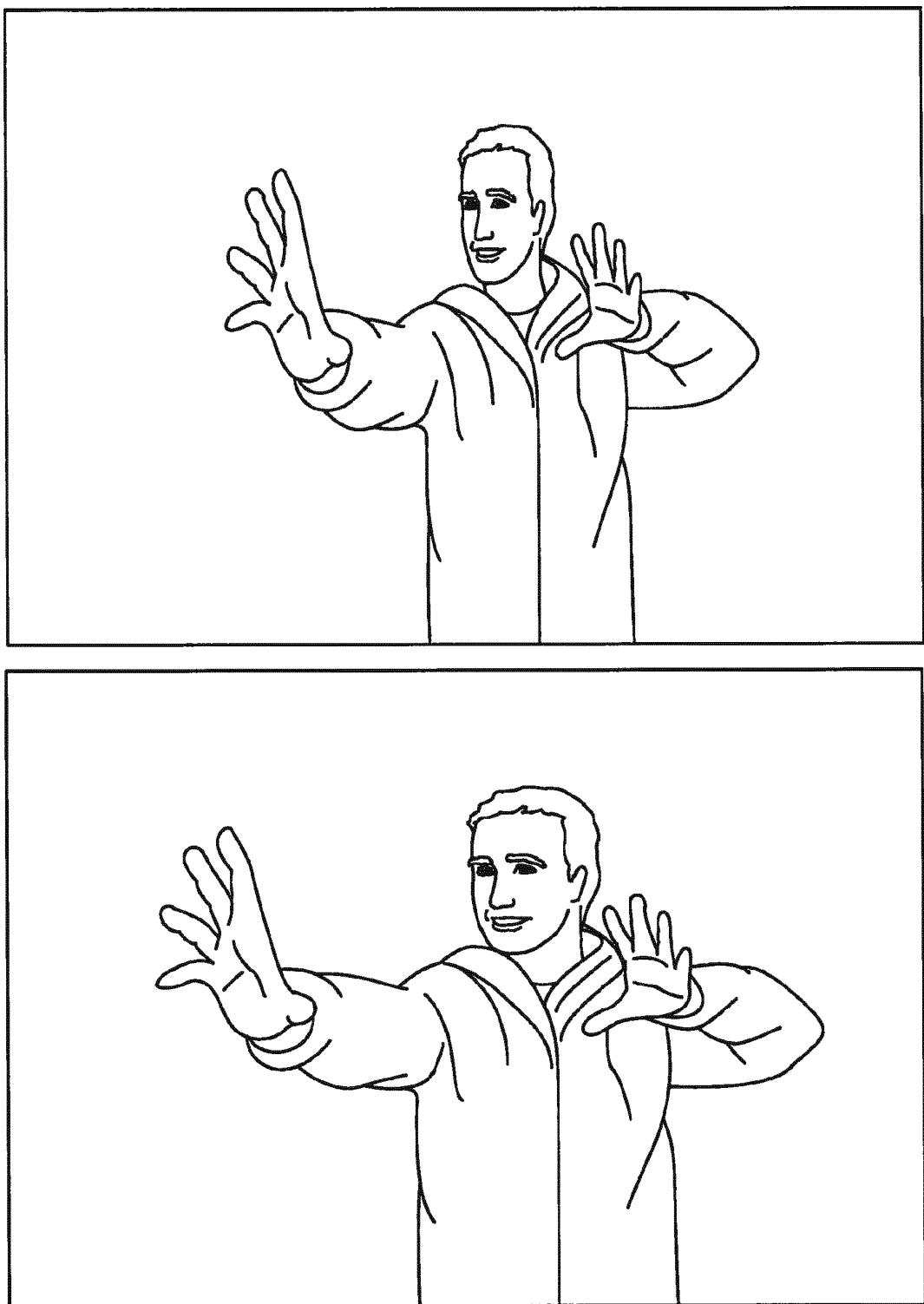


FIG. 12A

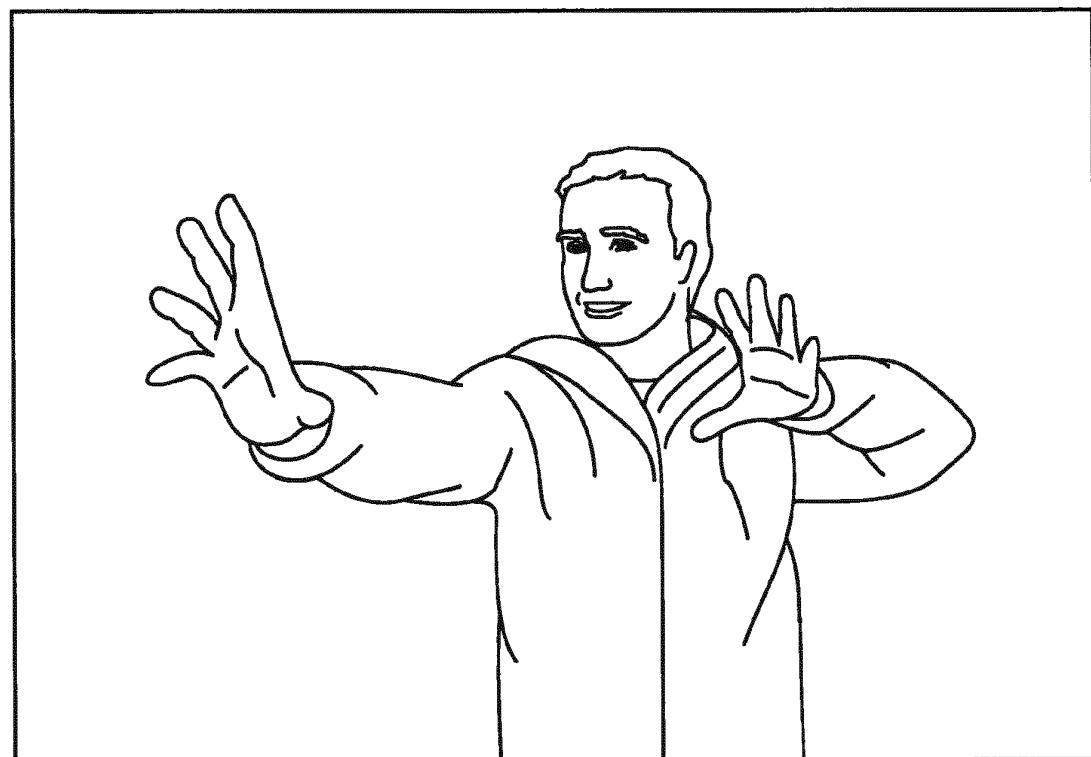
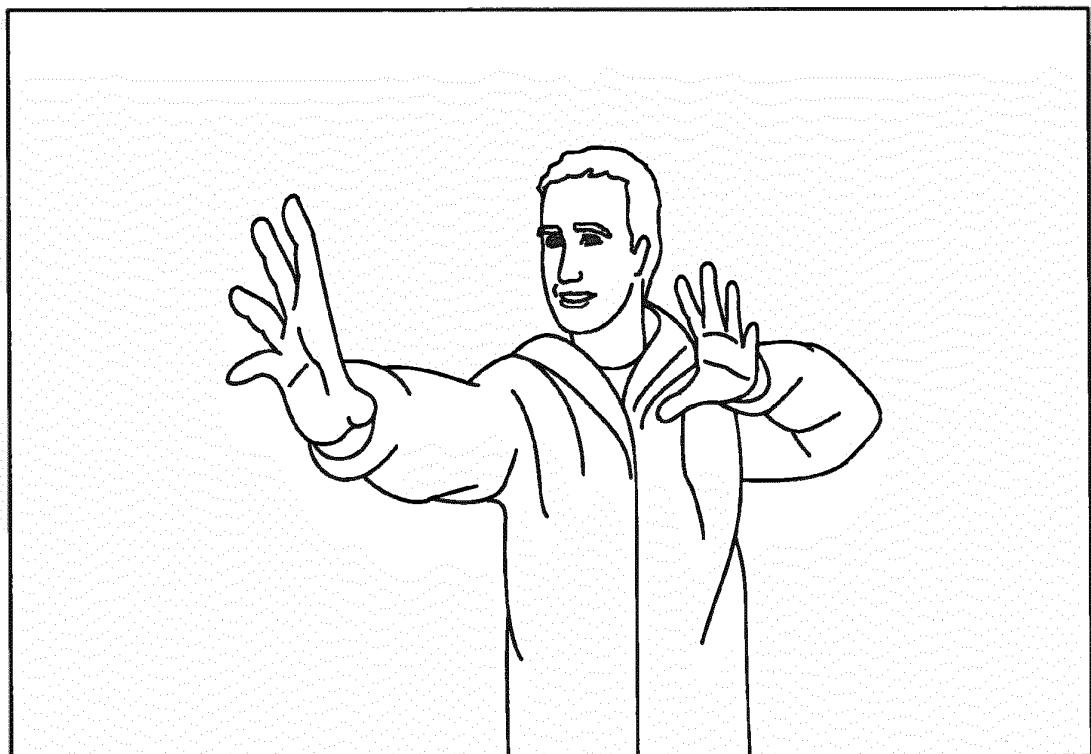


FIG. 12B

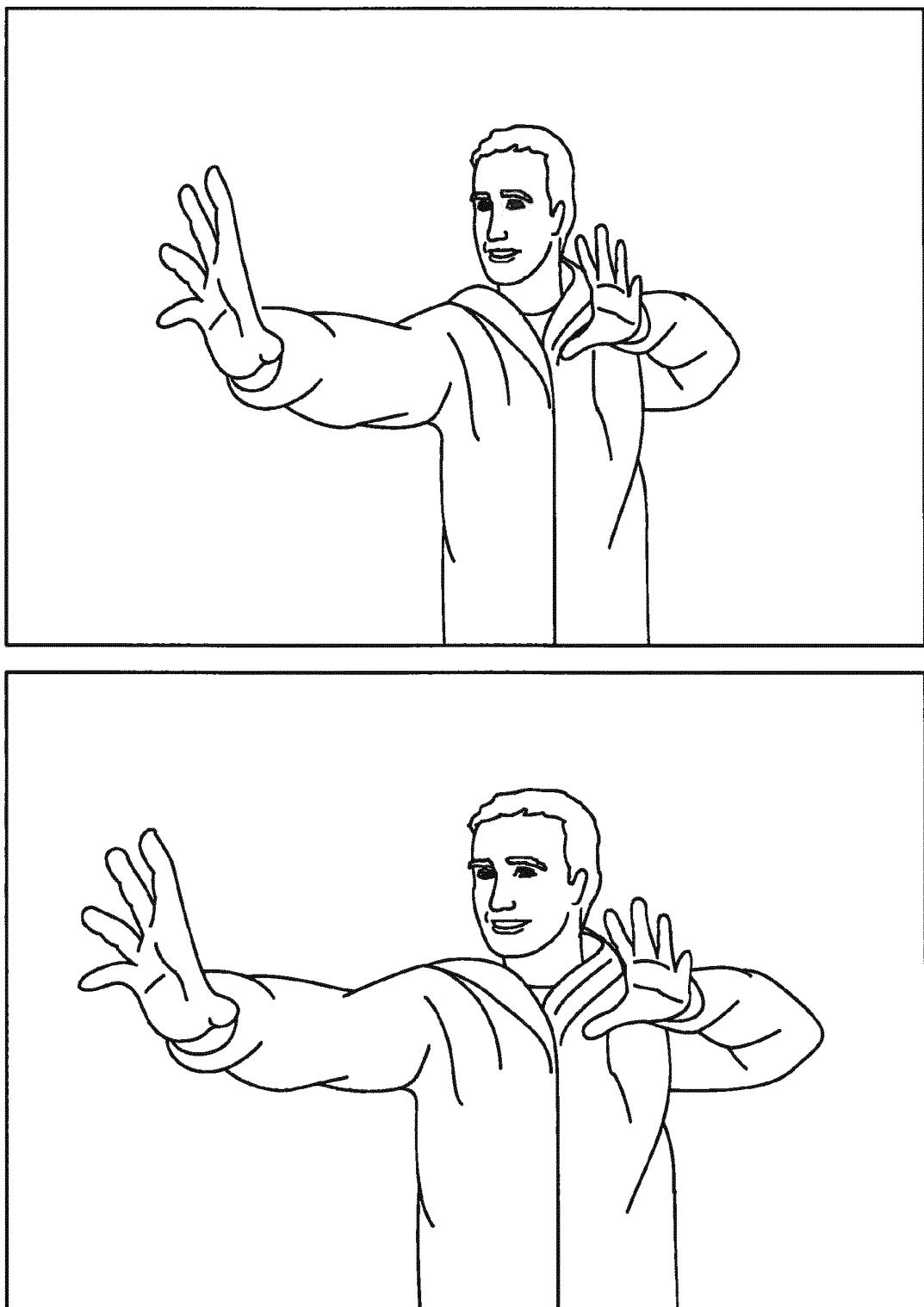


FIG. 12C

Fig. 13A ↗



Fig. 13B ↗

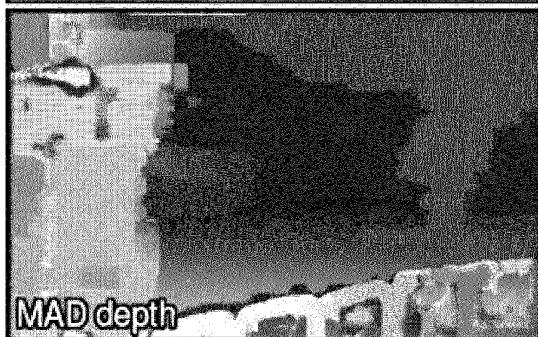
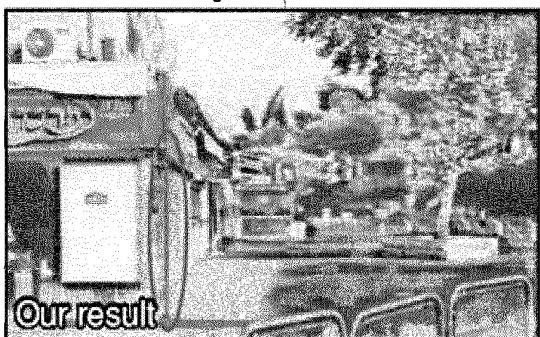


Fig. 13C ↗



Fig. 13D ↗

FIG. 13

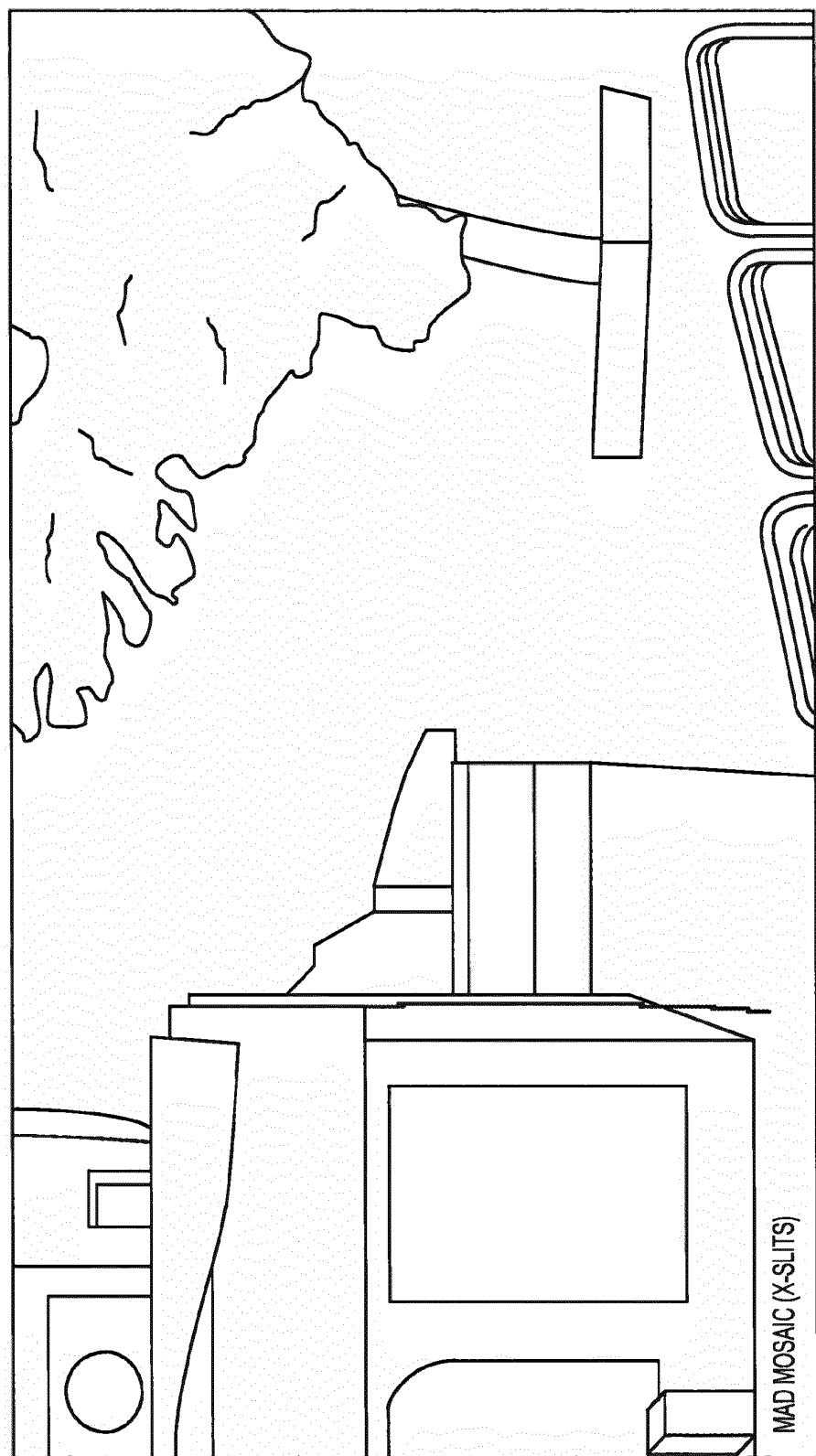


FIG. 13A

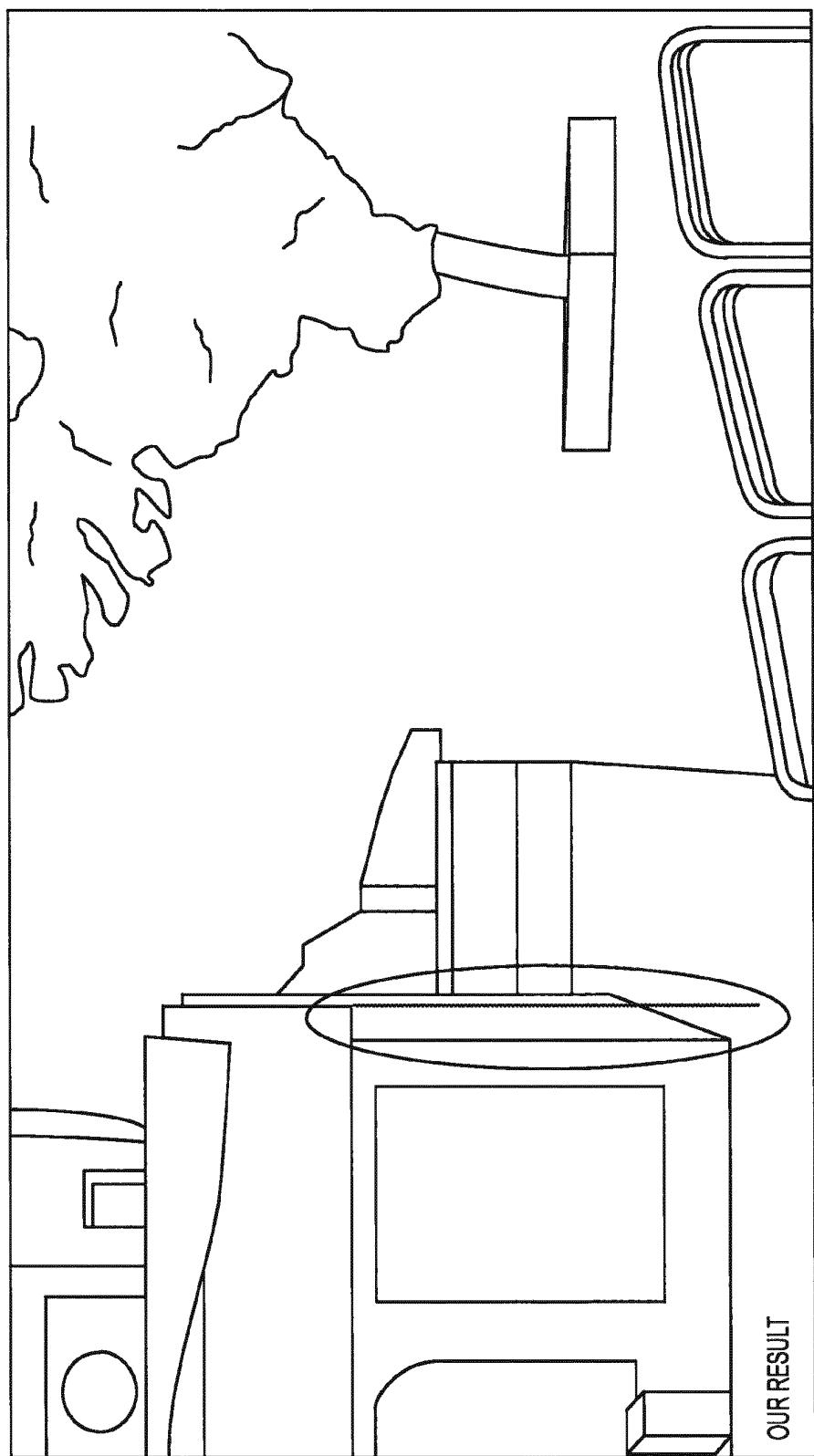
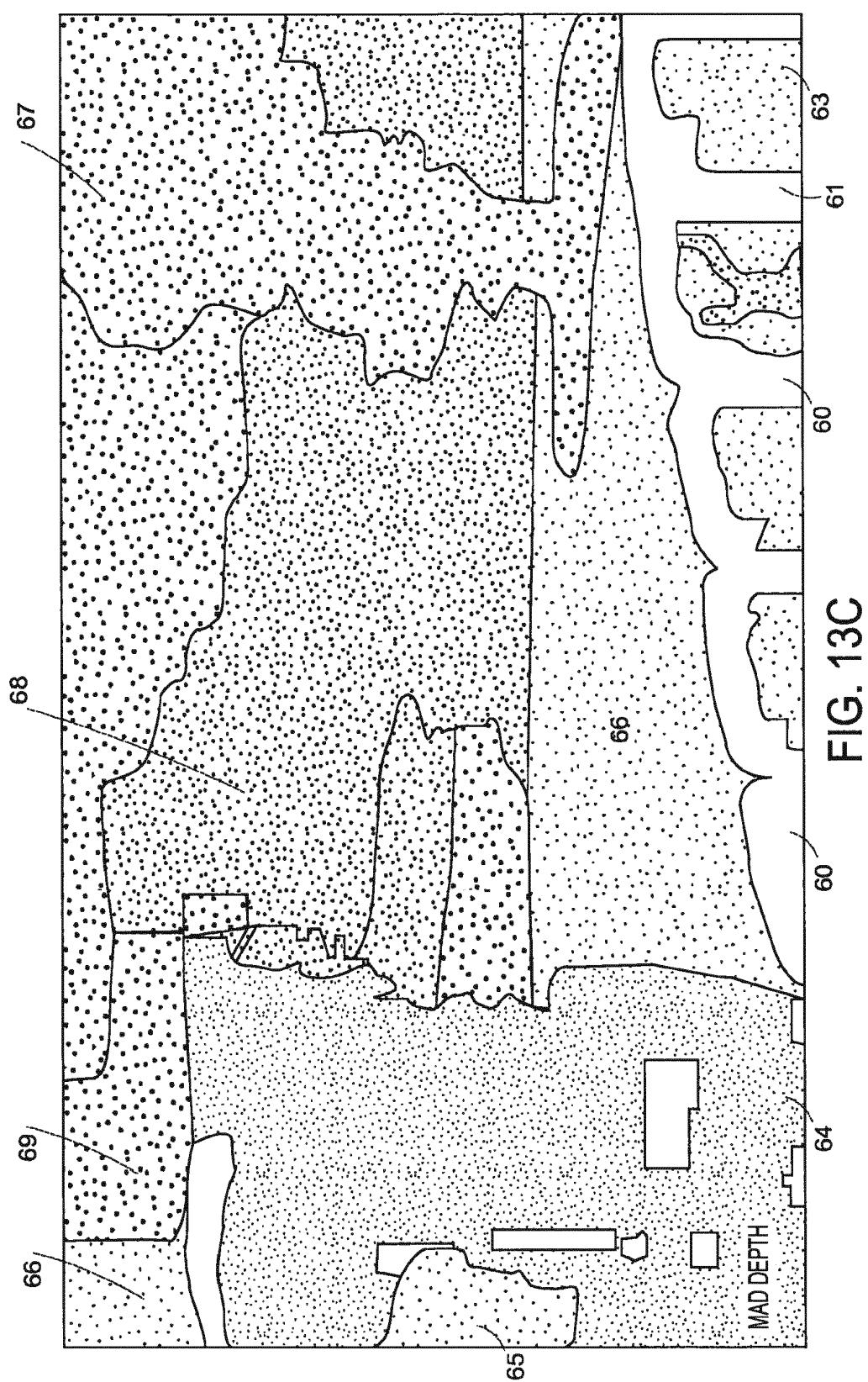


FIG. 13B



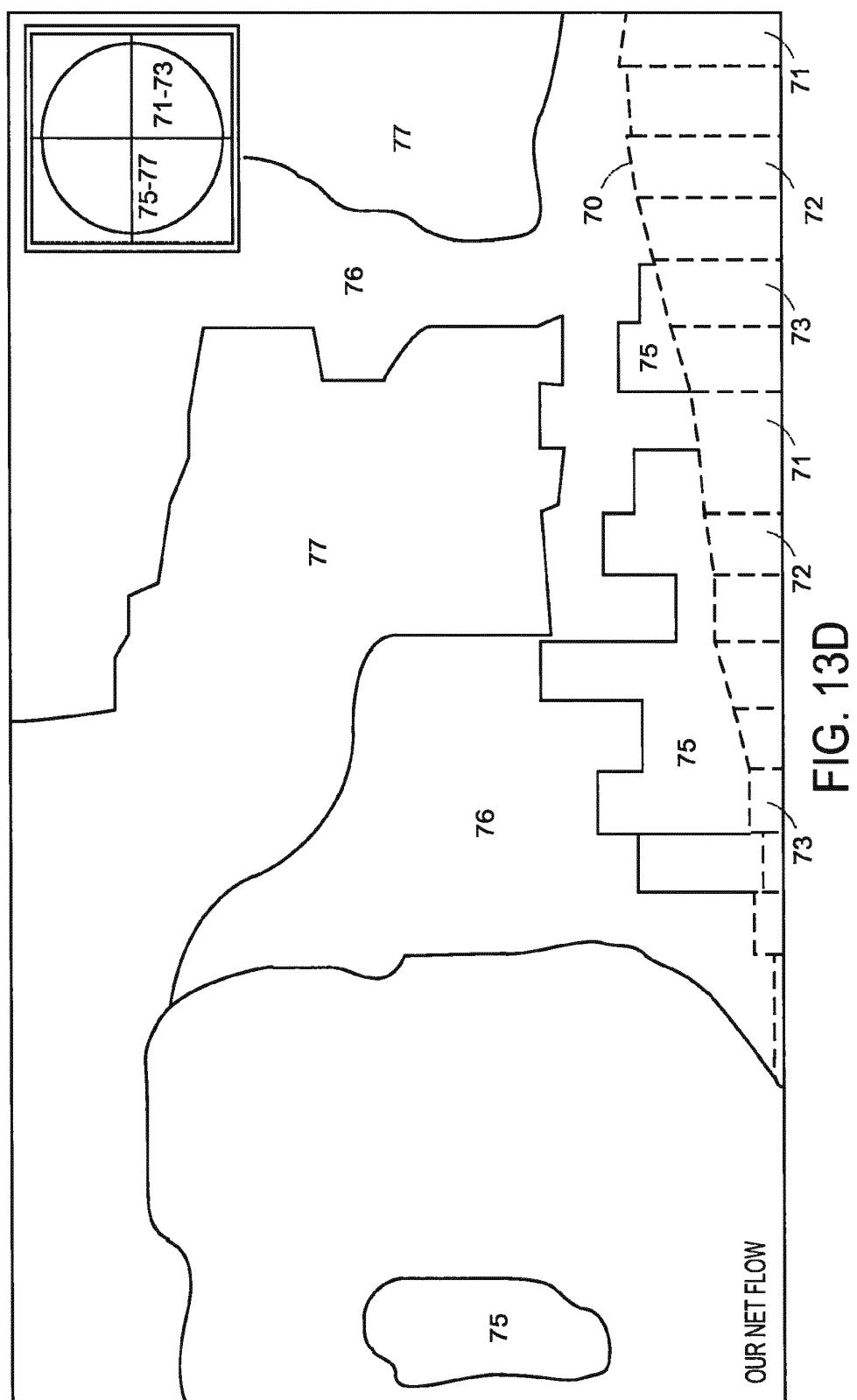


FIG. 13D



Fig.
14A

Fig.
14B

Fig.
14C

Fig.
14D

FIG. 14

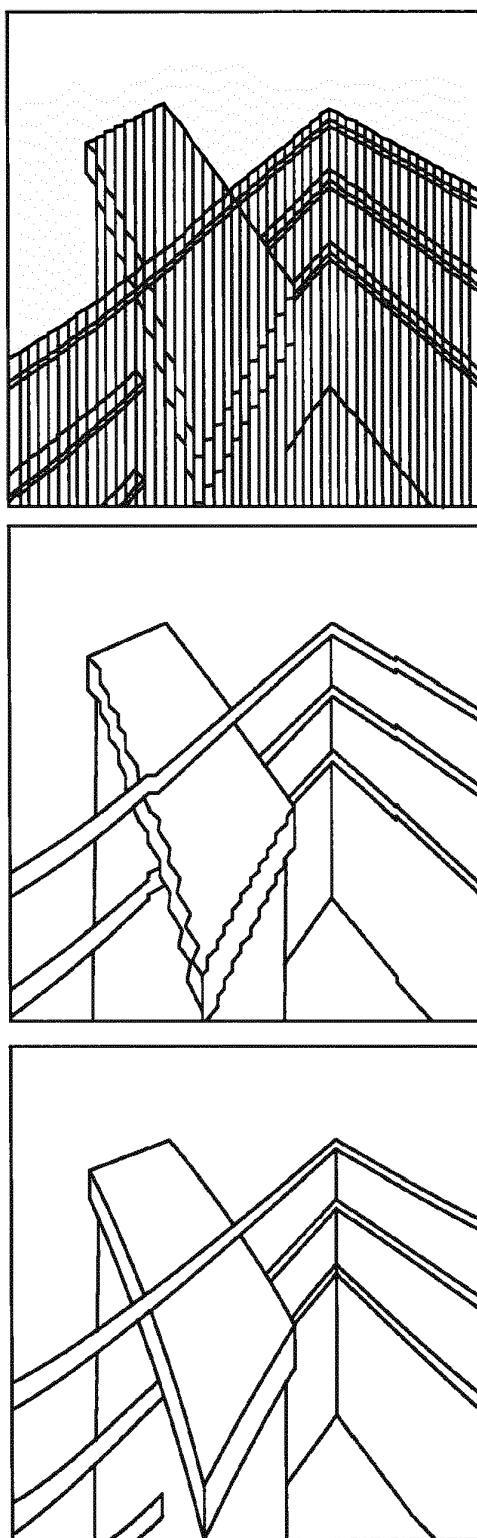


FIG. 14A

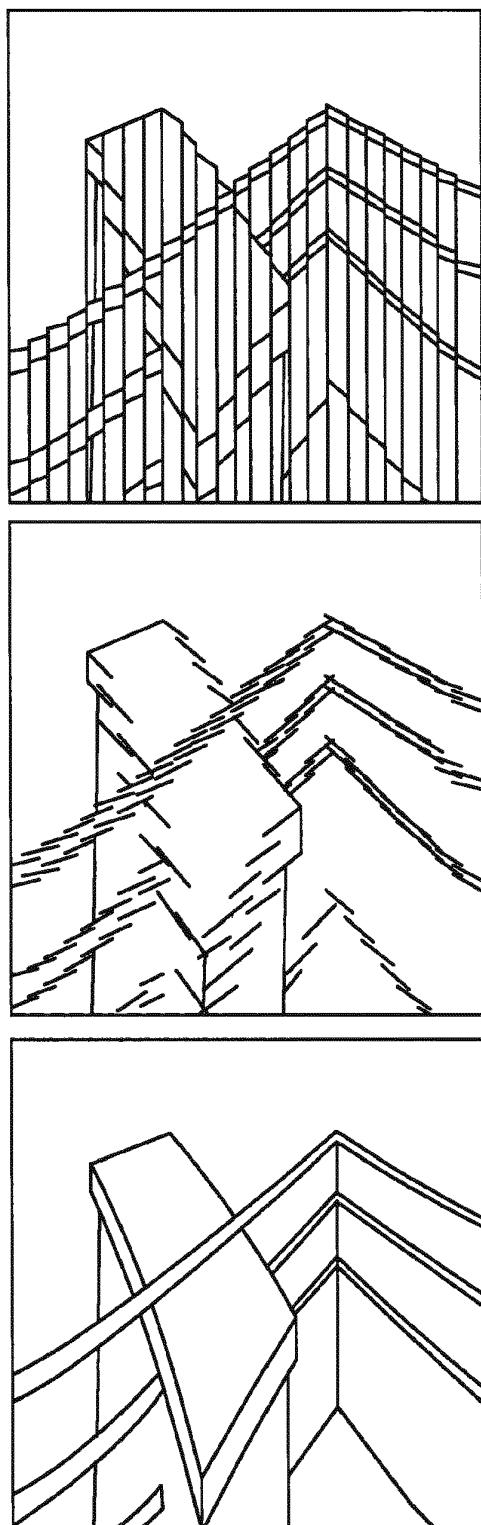


FIG. 14B

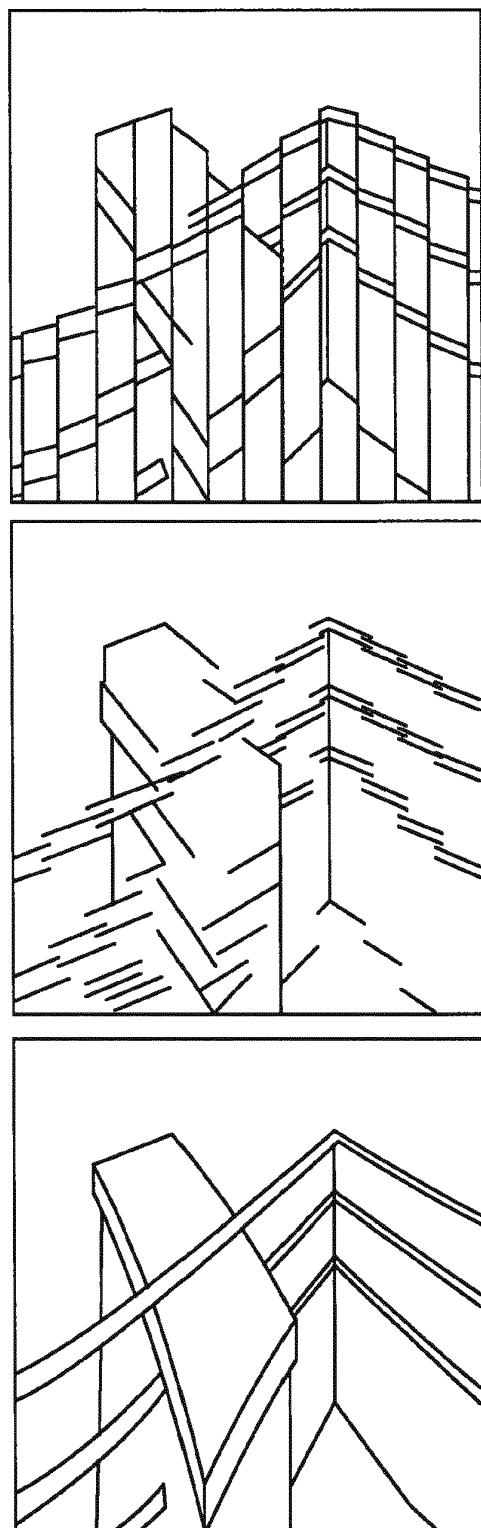


FIG. 14C

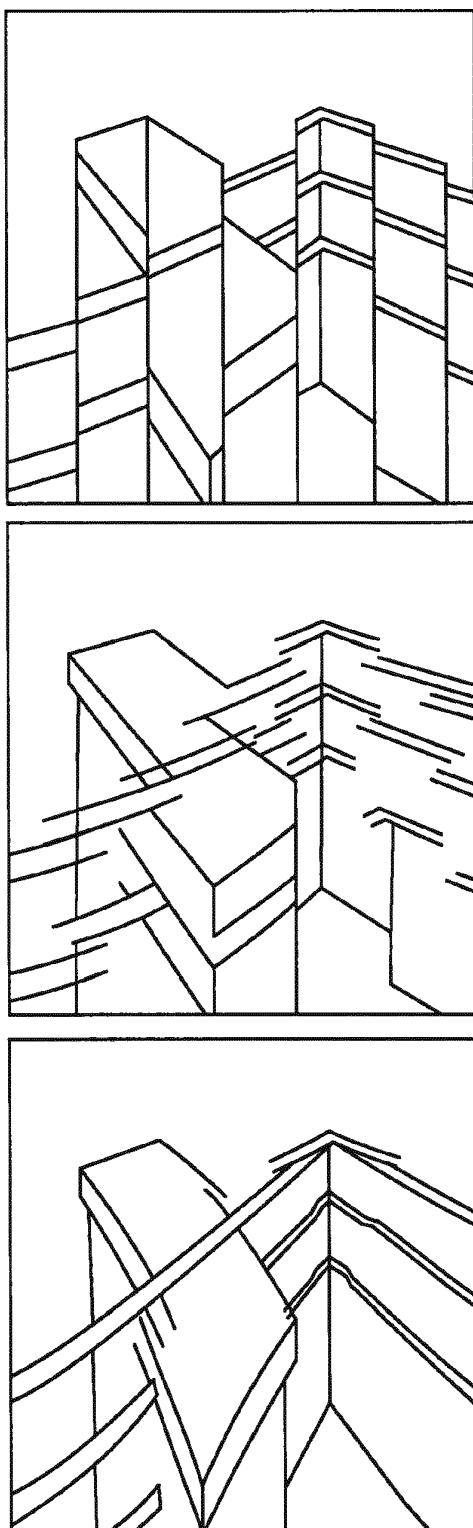
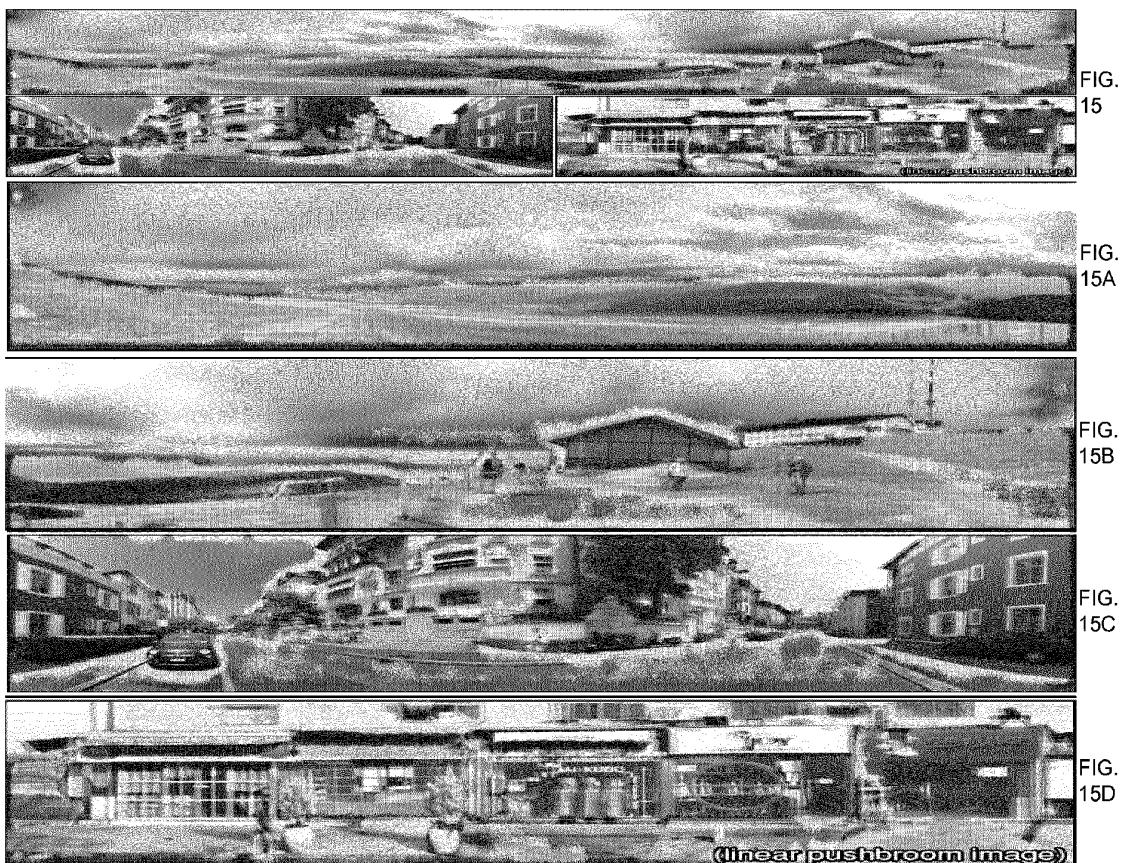


FIG. 14D



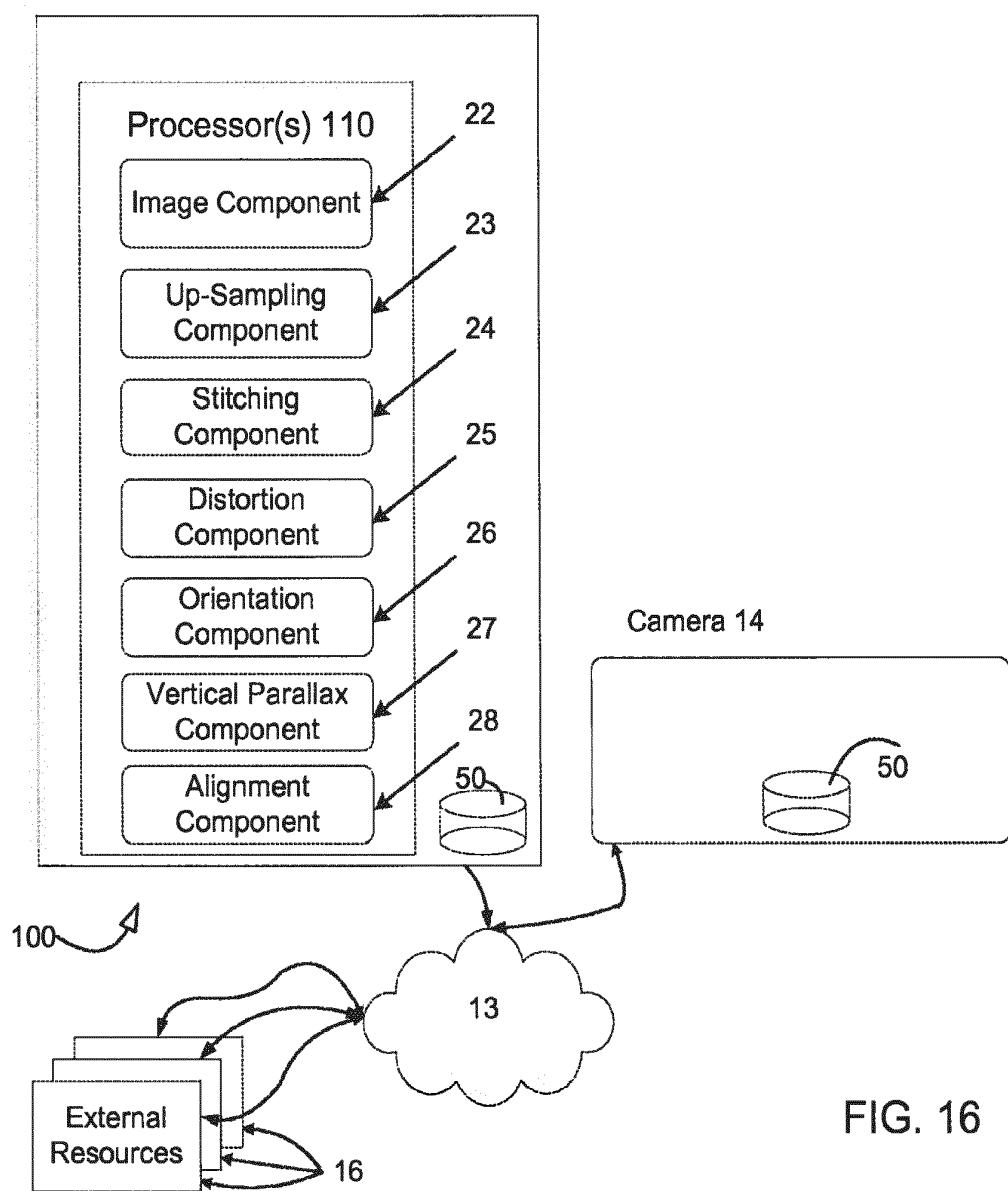


FIG. 16

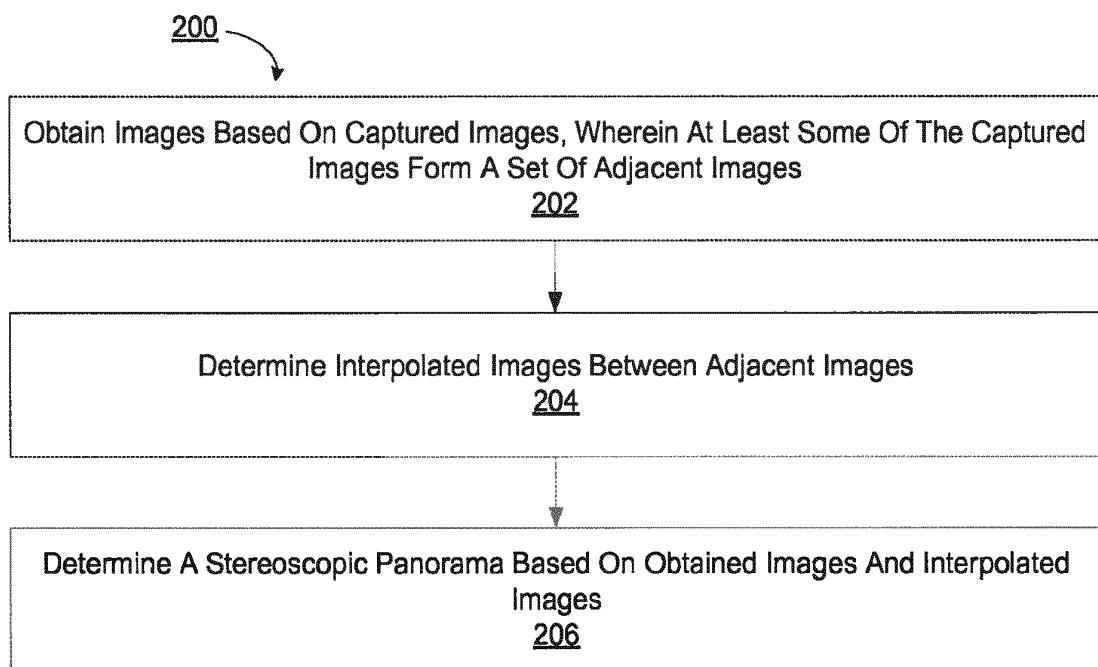


FIG. 17

5



EUROPEAN SEARCH REPORT

Application Number

EP 14 16 4965

10

15

20

25

30

35

40

45

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
Y	US 6 665 003 B1 (PELEG SHMUEL [IL] ET AL) 16 December 2003 (2003-12-16) * claim 1 *	1-16	INV. G06T3/40
Y	----- US 2003/035592 A1 (CORNOD KATHERINE H [US] ET AL) 20 February 2003 (2003-02-20) * paragraph [0003] - paragraph [0005] *	1-16	
A	----- PELEG S ET AL: "OMNISTEREO: PANORAMIC STEREO IMAGING", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, IEEE COMPUTER SOCIETY, USA, vol. 23, no. 3, 1 March 2001 (2001-03-01), pages 279-290, XP001005773, * page 280, right-hand column, paragraph 1 - page 281, left-hand column, paragraph 2 *	1-16	
A	----- US 6 795 109 B2 (PELEG SHMUEL [IL] ET AL) 21 September 2004 (2004-09-21) * the whole document *	1-16	
A	----- SHUM H-Y ET AL: "Systems and Experiment Paper : Construction of Panoramic Image Mosaics with Global and Local Alignement", INTERNATIONAL JOURNAL OF COMPUTER VISION, KLUWER ACADEMIC PUBLISHERS, NORWELL, US, vol. 36, no. 2, 1 January 2000 (2000-01-01), pages 101-130, XP002578824, * page 115, right-hand column, paragraph 4 - page 116, right-hand column, paragraph 4 *	1-16	TECHNICAL FIELDS SEARCHED (IPC) G06T
	----- -/-		
1	The present search report has been drawn up for all claims		
50	Place of search Munich	Date of completion of the search 11 September 2014	Examiner Pierfederici, A
55	CATEGORY OF CITED DOCUMENTS X: particularly relevant if taken alone Y: particularly relevant if combined with another document of the same category A: technological background O: non-written disclosure P: intermediate document		T: theory or principle underlying the invention E: earlier patent document, but published on, or after the filing date D: document cited in the application L: document cited for other reasons &: member of the same patent family, corresponding document

5



EUROPEAN SEARCH REPORT

Application Number
EP 14 16 4965

10

15

20

25

30

35

40

45

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
A	WO 98/34195 A1 (YISSLUM RES DEV CO [IL]; PELEG SHMUEL [IL]; ROUSSO BENNY [IL]) 6 August 1998 (1998-08-06) * page 7, paragraph 1 - paragraph 2 * * page 15, paragraph 3 - paragraph 5; claim 2 *	1-16	
A	----- SAWHNEY H S ET AL: "TRUE MULTI-IMAGE ALIGNMENT AND ITS APPLICATION TO MOSAICING AND LENS DISTORTION CORRECTION", PROCEEDINGS OF THE 1997 IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION. SAN JUAN, PUERTO RICO, JUNE 17 - 19, 1997, 17 June 1997 (1997-06-17), pages 450-456, XP000776535, * page 238, right-hand column, paragraph 4 *	1-16	
A	----- LI Y ET AL: "STEREO RECONSTRUCTION FROM MULTIPERSPECTIVE PANORAMAS", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, IEEE COMPUTER SOCIETY, USA, vol. 26, no. 1, 1 January 2004 (2004-01-01), pages 45-62, XP001185855, * the whole document *	1-16	TECHNICAL FIELDS SEARCHED (IPC)
The present search report has been drawn up for all claims			
1	Place of search Munich	Date of completion of the search 11 September 2014	Examiner Pierfederici, A
50	CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		
55	T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document		

**ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.**

EP 14 16 4965

5

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

11-09-2014

Patent document cited in search report		Publication date		Patent family member(s)		Publication date
US 6665003	B1	16-12-2003	AT EP IL JP JP JP US WO	420528 T 1048167 A1 136128 A 5358015 B2 2003524927 A 2010166596 A 2012257329 A 6665003 B1 0039995 A2		15-01-2009 02-11-2000 31-01-2013 04-12-2013 19-08-2003 29-07-2010 27-12-2012 16-12-2003 06-07-2000
US 2003035592	A1	20-02-2003	US US US	6665450 B1 2003035592 A1 2004091170 A1		16-12-2003 20-02-2003 13-05-2004
US 6795109	B2	21-09-2004	NONE			
WO 9834195	A1	06-08-1998	AT AU DE DE EP IL JP JP US US WO	280978 T 5573698 A 69827232 D1 69827232 T2 0954828 A1 131056 A 4007619 B2 2001510585 A 6532036 B1 2003076406 A1 9834195 A1		15-11-2004 25-08-1998 02-12-2004 20-10-2005 10-11-1999 06-07-2003 14-11-2007 31-07-2001 11-03-2003 24-04-2003 06-08-1998

^{EPO} For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 61812561 A [0001]

摘要

基於已有圖像來建立立體全景獲取圖像的系統和方法。此獲取圖像可被處理和/或預處理，例如補償獲取期間不理想相機定向造成的透視變型，減少垂直視差，對齊相鄰圖像，和/或減少相鄰圖像之間的旋轉和/或位置漂移。此獲取圖像可以用於內插飛行中的虛擬圖像來減少全景圖像中的可視偽像。獲取和/或內插的圖像（或圖像片段）可以被縫合一起來形成立體全景。