



(86) Date de dépôt PCT/PCT Filing Date: 2013/04/10
(87) Date publication PCT/PCT Publication Date: 2013/10/17
(85) Entrée phase nationale/National Entry: 2014/10/14
(86) N° demande PCT/PCT Application No.: US 2013/036005
(87) N° publication PCT/PCT Publication No.: 2013/155203
(30) Priorités/Priorities: 2012/04/13 (US61/624,145);
2013/03/14 (US61/785,517)

(51) Cl.Int./Int.Cl. *G06T 7/20* (2006.01),
G06T 19/00 (2011.01)
(71) Demandeur/Applicant:
LIGHTCRAFT TECHNOLOGY LLC, US
(72) Inventeurs/Inventors:
MACK, NEWTON ELIOT, US;
MASS, PHILIP R., US
(74) Agent: SMART & BIGGAR

(54) Titre : SUIVI DE PRECISION HYBRIDE
(54) Title: HYBRID PRECISION TRACKING

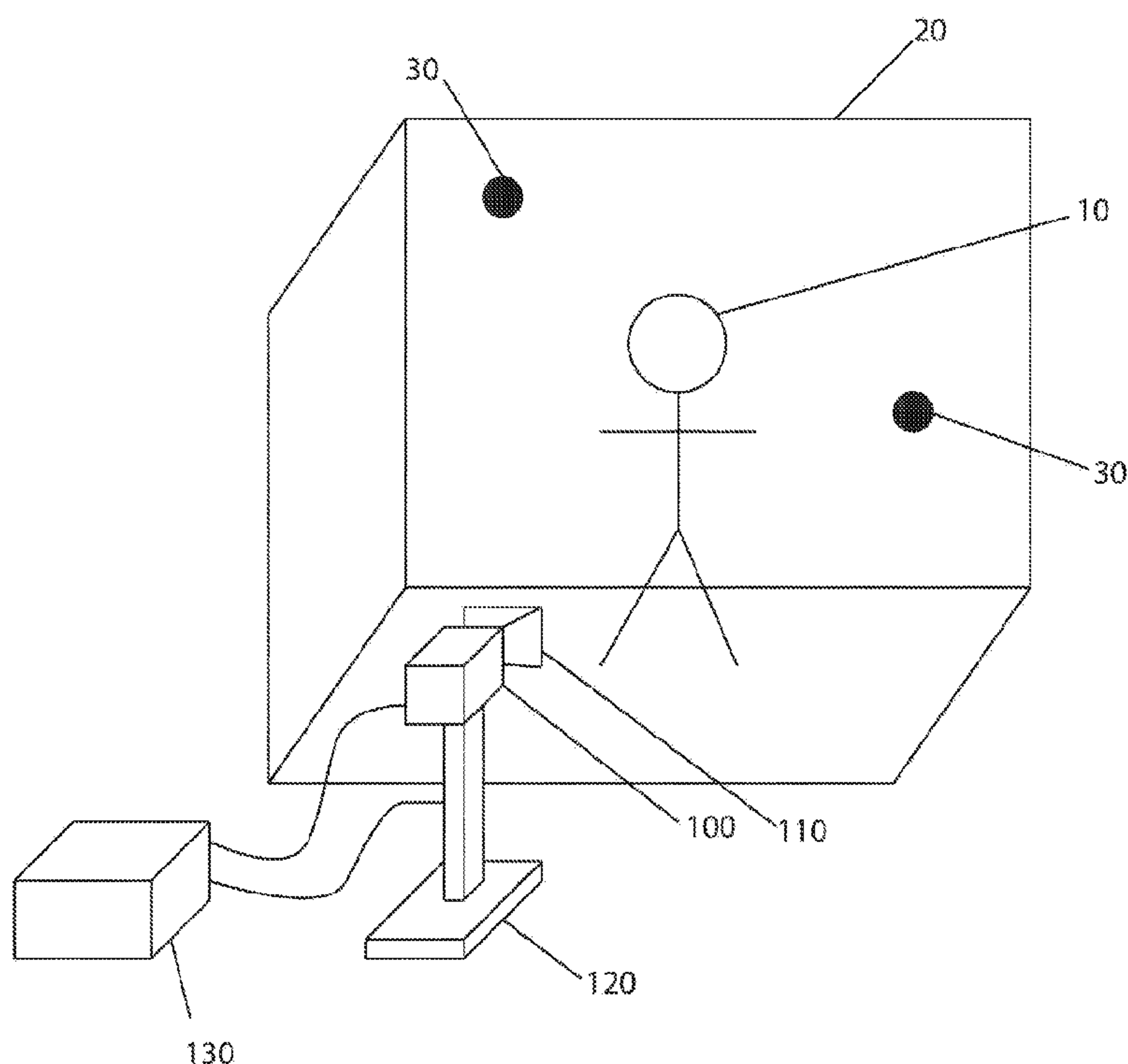


FIG. 1

(57) **Abrégé/Abstract:**

Disclosed herein are through-the-lens tracking systems and methods which can enable sub-pixel accurate camera tracking suitable for real-time set extensions. That is, the through-the-lens tracking can make an existing lower precision camera tracking and compositing system into a real-time VFX system capable of sub-pixel accurate real-time camera tracking. With this enhanced level of tracking accuracy the virtual cameras can be used to register and render real-time set extensions for both interior and exterior locations.

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

(19) World Intellectual Property
Organization
International Bureau

WIPO | PCT

(10) International Publication Number
WO 2013/155203 A3(43) International Publication Date
17 October 2013 (17.10.2013)(51) International Patent Classification:
G06K 9/00 (2006.01)(21) International Application Number:
PCT/US2013/036005(22) International Filing Date:
10 April 2013 (10.04.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/624,145 13 April 2012 (13.04.2012) US
61/785,517 14 March 2013 (14.03.2013) US(71) Applicant: **LIGHTCRAFT TECHNOLOGY LLC**
[US/US]; A Delaware Corporation, 612 Venice Boulevard,
Venice, CA 90291 (US).(72) Inventors: **MACK, Newton, Eliot**; % Lightcraft Techno-
logy Llc, 612 Venice Boulevard, Venice, CA 90291 (US).
MASS, Philip, R.; % Lightcraft Technology LLC, 612
Venice Boulevard, Venice, CA 90291 (US).(74) Agent: **LARSON, Douglas, N.**; Brooks Acordia IP Law,
PC, 1445 E. Los Angeles Avenue, Suite 108, Simi Valley,
CA 93065 (US).(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,
BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP,
KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD,
ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI,
NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU,
RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ,
TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA,
ZM, ZW.(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ,
UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,
TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,
EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,
MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
ML, MR, NE, SN, TD, TG).

Published:

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

[Continued on next page]

(54) Title: HYBRID PRECISION TRACKING

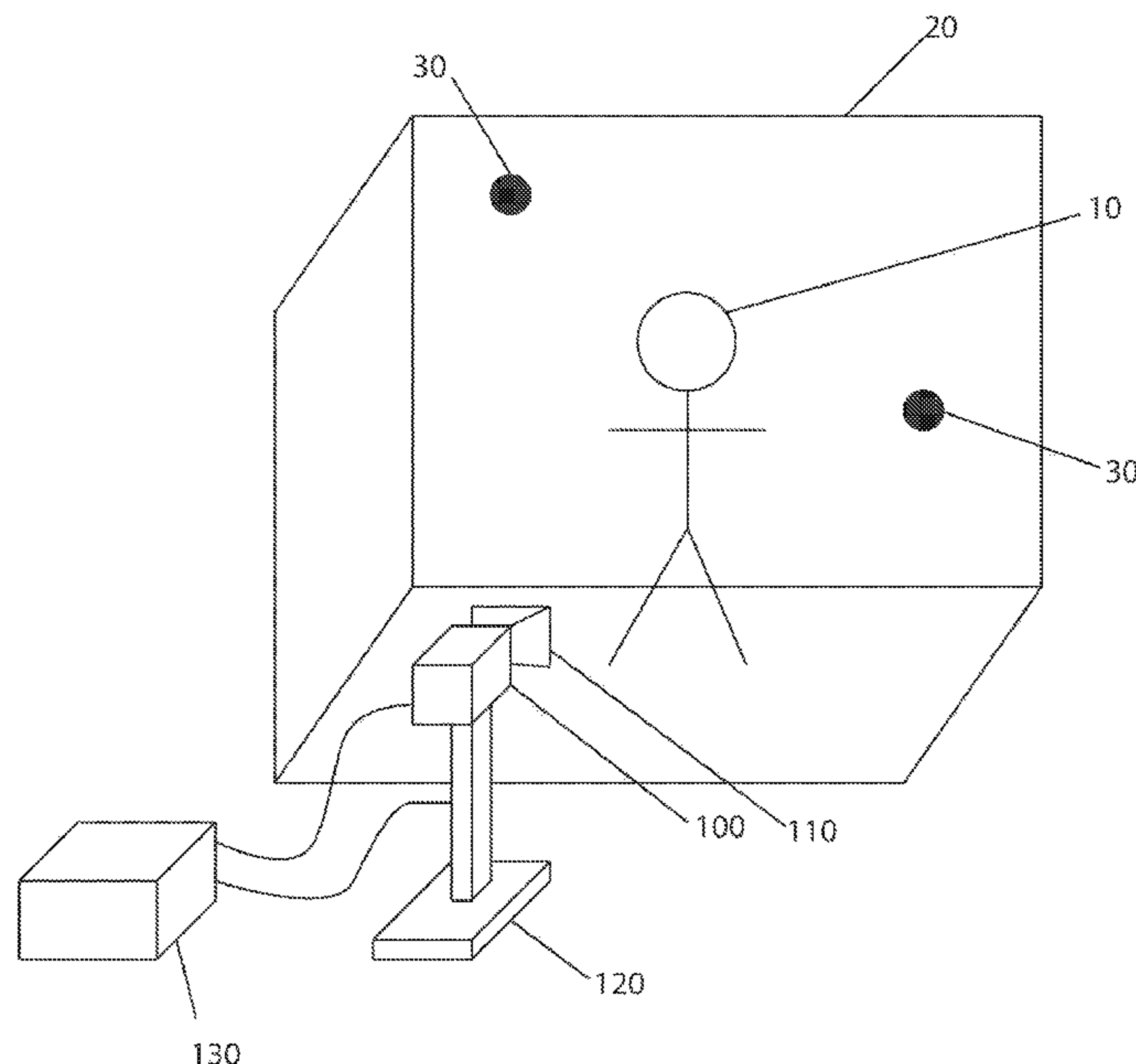


FIG. 1

(57) Abstract: Disclosed herein are through-the-lens tracking systems and methods which can enable sub-pixel accurate camera tracking suitable for real-time set extensions. That is, the through-the-lens tracking can make an existing lower precision camera tracking and compositing system into a real-time VFX system capable of sub-pixel accurate real-time camera tracking. With this enhanced level of tracking accuracy the virtual cameras can be used to register and render real-time set extensions for both interior and exterior locations.

WO 2013/155203 A3



(88) Date of publication of the international search report:
27 December 2013

HYBRID PRECISION TRACKING

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of co-pending provisional application Serial
5 No. 61/624,145, filed April 13, 2012, and whose entire contents are hereby incorporated
by reference, and of co-pending provisional application Serial No. 61/785,517, filed
March 14, 2013, and whose entire contents are hereby incorporated by reference.

TECHNICAL FIELD

10 This application relates to systems and methods for combining real scene
elements from a video, film, digital type camera or the like with virtual scene elements
from a virtual camera into a finished composite image, and more particularly, to systems
and methods for creating matching virtual scene elements with live action scene elements
to a high degree of precision.

BACKGROUND

15 Combining real world imagery with additional imagery from another source
requires careful control over which sections of each image are to be used in the final
composite image. One common application is to combine images generated by a
20 computer with images acquired from a traditional motion picture, video or digital camera.
In order to seamlessly combine the images, the camera parameters from the point of view
of the virtual scene camera must be closely matched to the parameters of the live action
camera. In addition, for use in on set and live applications, the match between virtual and
live action sceneries must be completed quickly, preferably in a matter of milliseconds in
25 order to be useful at standard motion picture and television production frame rates of 24

to 30 frames per second.

Tracking the live action camera and lens may be achieved in a variety of ways. (See U.S. patent application Serial No. 12/832,480, filed July 8, 2010, entitled “Methods and Systems for Calibrating an Adjustable Lens,” which was published on February 3, 5 2011 as Publication No. 20110026014, and which issued as U.S. Patent 8,310,663 (‘663) and whose entire contents are hereby incorporated by reference.) The measurement of the live action camera's position and orientation may be achieved with optical encoders, by solving the pose of fiducial markers located overhead or in the scene, with inertial accelerometers, by matching high contrast points from one image to the next, and many 10 other methods common in the field.

Real-time measuring methods that measure the angular tilt of the camera (pan, tilt and roll) at the camera itself share a common problem. Small errors in angular measurement become large positional errors when projected out several feet, which is where the majority of live action and virtual scenery seams are located. This can be seen 15 clearly in a composited image as a shifting or sliding between the virtual and live action elements when the camera is moved.

A traditional method of solving this high precision seam is to use post-production software tools, such as Syntheyes, 3D Equalizer, Boujou, PF Track, or Nuke that locate and track multiple high contrast regions from one frame to the next. By analyzing the 20 relative motion of several points on the screen, the live action camera motion and lens parameters can be derived and the motion of the virtual background matched cleanly to the live action foreground.

This process has several limitations. Firstly, if an actor walks in front of the high

contrast area being tracked, the camera solution mathematics can easily be disrupted. Secondly, the process of measuring the motion by tracking the individual pixels of the high contrast part of the image is both fragile and time-consuming. It typically cannot be computed in real time, and if a frame of the live action image has a lighting change where the pattern is unrecognizable, the artist must re-specify the high contrast area at the frame of failure to continue the process. In fact, the process of solving an accurate camera track from an arbitrary moving background can exceed the time required to complete the rest of the shot due to the handwork required.

In addition, if the live action camera is zoomed in, the high contrast area that was being tracked can simply disappear from the image, resulting in the camera failing to track the object. It is notoriously difficult for the algorithms used to derive camera intrinsic parameters from shots where the camera is simultaneously changing its focal length (“zooming in”) and translating.

Accordingly, the pure pixel tracking based methods do not work well for the demands of real-time visual effects processing, which must be very rapid to compute as well as robust to the frame-by-frame changes in the live action video image and changes in the adjustment of the cinema lens being used to record the image.

A prior art system is the PREVIZION system, which is available from Lightcraft Technology, LLC. of Venice, California. The PREVIZION system includes a camera tracking system, a lens calibration system, a real-time compositing system, and a built-in 3D renderer. An example of a publication disclosing the prior art PREVIZION system is the PREVIZION product brochure, entitled “Previzion Specifications 2011,” published on April 8, 2011, and whose contents are incorporated by reference.

SUMMARY

Various embodiments of a hybrid precision tracking system are disclosed herein.

In one embodiment, an artist surveys the 3D center position of artificial high contrast

5 marker points that are placed on the stage where virtual scenery will join with live action.

The artist then enters the position of these points into a computer system. The artist selects points from the live action image that correlate with the 3D positions of the centers of the high contrast marker points that were measured in the live action scene.

These points can be selected using a 2D display of the live action image, for example by
10 locating a pointer in the desired location over the visible marker, and pressing a selection button.

The live action camera and lens connected to the computer system are being tracked by a camera and lens tracking system, so that the live action camera and lens parameters are closely known. The computer can then predict where on the 2D screen the
15 3D surveyed points should be, and display a virtual marker there.

The above process generates a list of points with a known 3D position, a predicted 2D screen position of these points, and an actual 2D screen position. The process of minimizing the error between the predicted 2D screen position of the 3D surveyed points and the actual 2D screen position of the 3D surveyed points can create a high precision
20 tracking solution.

Since the 2D-to-3D correlation can be created by clicking on a live action object, the 2D representation is viewed from the position of the current live action camera. For the 3D surveyed point to line up accurately, it can be projected from a virtual camera with the same position, orientation and lens parameters as the live action camera. The

virtual camera can be calibrated to the live action camera using a camera and lens calibration system as described in the above-mentioned '633 patent.

For high accuracy, the system quickly detects the optical center of the scene marker. In one aspect of the disclosure, the marker is a circular dot. The detection of the dot can be achieved by a series of operations that are optimized on the GPU. In one embodiment, this set of operations can include: edge or slope detection; removal of parallel lines to select for circular markers; searching along the normal of the current slope for the opposite side of the circle, and thereby determining a candidate for the center point of that circle; voting for the most likely center point, using the accumulation of all of the centerlines; and readout of the highest weighted point to the CPU with a fast data transfer, such as a geometry shader.

After this is done, the delta between the predicted point XY screen location and the actual XY screen location can be minimized, by performing a least squares or similar error minimizing solution on the pan, tilt and roll rotary axes of the camera to provide a delta angle to the incoming angular data.

As searching the entire screen for markers is costly computationally, the system can predict the XY projected marker location using its list of marker XYZ positions, and only search in a pattern around the position of the marker, taking into account all camera and lens parameters, such as lens field of view and distortion.

Since the optical solution of the centers of the circles may vary from frame to frame, causing a visible "jitter" in the virtual to live action seam, this delta angle in turn can be filtered to remove the "jitter" by a Kalman or other low pass filter to create a smooth motion while retaining accuracy.

When an object is removed from view, the remaining delta angle is also filtered to avoid jitter. Even if all of the markers are excluded, the delta angle between the incoming camera tracking data and the minimized solution is preserved, so that when the markers appear on the other side of the actor they can be re-acquired with minimal visual artifacts.

- 5 The concentration on filtering the delta angle of the solution means that the visual solve is robust to individual marker occlusion.

In traditional optical tracking systems, the XY position of the markers drives the focal length of the solve, and slight marker instability can cause errors in focal length calculation. With a hybrid approach of the present disclosure, the calculation of the lens
10 focal length can be done earlier and stored, and the marker motion can be rapidly stabilized in the XY space without losing lens focal length accuracy.

The separation of the two components (the external sensor based camera and lens tracking, and pixel based delta angle calculation) makes a real-time use of this system possible. In addition, the hybrid approach can easily handle zooming and high distortion
15 lenses.

According to another aspect of the disclosure, the 3D locations of the markers can also be entered using 3D survey data.

According to a further aspect of the disclosure, the XYZ position of the markers can be solved by translating the camera and shooting two or more images of the dots
20 from separate points of view, then using a photogrammetry solution to solve for the XYZ positions of the markers. This enables use of the system without an external survey tool.

According to a still further aspect of the disclosure, the system can use naturally occurring features, such as the SIFT or SURF algorithms.

According to a yet still further aspect of the disclosure, the system can scan the entire screen.

According to another aspect of the disclosure, the computations required to perform subpixel matching can be performed in real time.

5 According to a further aspect of the disclosure, the hybrid precision tracking can automatically handle variations in lens focal length or distortion.

According to a still further aspect of the disclosure, the marker data can be entered in standard 3D survey coordinate form and rapidly modified by the artist during production.

10 According to another aspect a hybrid precision tracking method can be achieved with data that is already existing in a real-time compositing and camera tracking system.

According to a further aspect a Hough transform or other related transform can be used to identify the circle centers.

15 According to a still further aspect the system can use active markers to generate small dots of light that can be seen identically from most orientations.

According to a still further aspect the system can use an artificial bar code or fiducial type marker that can be rapidly identified by standard machine vision techniques.

20 According to a yet further aspect the current camera motion data from the camera tracking system can be used to generate a “deblurred” image of the marker. (This can be important in order to handle tracking with any degree of camera motion.)

Disclosed herein is a through-the-lens tracking system/method which when in a PREVISION system (or method) enables sub-pixel accurate camera tracking suitable for real-time set extensions. That is, the through-the-lens tracking incorporated into a

PREVIZION system makes the system a real-time VFX system capable of sub-pixel accurate real-time camera tracking. With this enhanced level of tracking accuracy the virtual cameras can be used to register and render real-time set extensions for both interior and exterior locations.

5 A through-the-lens tracking system of the present disclosure can be integrated with existing PREVIZION camera tracking tools to provide additional sub-pixel accurate tracking information for any shot especially where optical tracking or encoded cranes cannot provide a suitable level of accuracy to be used in final composites. The PREVIZION system automatically follows pre-measured tracking markers that are
10 visible through the main camera lens (whether mounted directly on the set or a green screen) and merges that data with the standard optical/inertial or encoded crane tracking information.

 In the past, precise registration of virtual elements and real world settings was achieved by doing expensive and time-consuming post-optical tracking. This time
15 consuming process is transformed herein by removing the need to use post-production optical trackers like Syntheyes or 3D Equalizer, saving hours of processing time per shot. With the new combination of tracking techniques, PREVIZION users can achieve sub-pixel accurate tracking in real time and can do complicated set extensions with the same ease and cost effective techniques that are currently being achieved with PREVIZION.
20 That is, the human eye can detect a wobble of a pixel; thus, the present system has a registration with the precision to generally produce tracking accuracy within a fraction of a pixel, e.g., “sub-pixel” accuracy.

 Additional tools verify tracking accuracy and register the through-the-lens

information with a locally surveyed coordinate framework. These tools can take into account complex calculations involving image recognition and analysis to track multiple targets visible in the main camera viewfinder. All of this processing can happen within a tight time frame to maintain the real-time capability of the PREVIZION system. More specifically, video is typically shot at frame rates of anywhere from 24 to 60 frames per second. Since the system also has to do the other keying, tracking, rendering and compositing tasks on the same GPU, the circular detection and solution can be done within just a few milliseconds. In fact, the algorithm of the system can be quick enough to predict and detect a circle in less than two milliseconds.

According to an embodiment herein a hybrid through-the-lens tracking system is disclosed. The system can include a first system and a second system operatively connected thereto. The first system can include a calibrated look-up table and can be configured to use the calibrated look-up table to modify search locations for scene markers when using adjustable lenses (such as for motion picture and television visual effects production). The first system can be configured to transmit predicted 2D locations of the scene markers to the second system, wherein the second system is configured to use a high speed GPU-based scene marker detection scheme that uses a fast geometry based search to calculate the predicted 2D locations of the centers of the scene markers. The high speed detection can be under 10 milliseconds per marker, or between 2 and 10 milliseconds per marker. The scene markers can be circular markers, fiducial markers, bar code markers or natural feature markers.

Also disclosed herein is a hybrid through-the-lens tracking method that includes subjecting surveyed live action scene marker locations, live action images and combined

camera and lens data to a hybrid precision tracking process to produce filtered camera pose data. The hybrid precision tracking process can include: using the incoming live camera and lens data to predict the 2D position of the locations of on screen markers; searching within an area around the predicted 2D position for markers; calculating the
5 actual XY position of the markers; and modifying the camera pose data to align the predicted positions with the actual positions. The process can further include filtering the resulting angular changes to remove noise from the process.

A method of hybrid through-the-lens tracking is disclosed. This method can include: reading the current incoming camera and lens information; using the lens
10 position information to look up the current lens parameters; using the stored lens parameters and camera information to predict the location of 2D markers in the scene; detecting the positions of the actual 2D markers in the scene; and calculating the necessary angular corrections from the incoming camera position data to best correct the camera's angular position to match the actual 2D markers. And the information can
15 include the current measured camera position and orientation, and the lens parameters include focal length, distortion and entry pupil location.

According to another aspect a hybrid through-the lens tracking method is disclosed. The method can include: using a calibrated look-up table of a first system to modify search locations for scene markers when using adjustable lenses (such as for
20 motion picture and television visual effects production); transmitting, using the first system, predicted 2D locations of the scene markers to a second system; and using the first system and the predicted 2D locations to minimize the reprojection error between the estimated positions and the actual detected positions of the scene markers.

A hybrid precision tracking method is disclosed which includes calculating a filtered pose using predicted camera location data, 3D location data of surveyed tracking markers and a live action image. This method can further include sending the filtered pose to a 3D renderer to generate a correctly aligned rendered 3D image. The calculating
5 step can include: measuring the camera's position and orientation; measuring the adjustment position of the camera lens; looking up the optical parameters of the lens, based on the measured adjustment lens positions; predicting the on-screen XY position of surveyed physical locations in the scene, using the camera and lens information; searching in an area surrounding the predicted XY position for a detectable marker that
10 represents the actual position of the physical location; measuring the difference in camera angular movement required to minimize the error between the predicted and actual physical XY location. And the calculating step can further include filtering the angular difference value over time to remove any angular noise introduced by the measuring steps.

15 A method of finding the center of a circle is also disclosed. The method can include: detecting the edge or slope of an image luminance or color, and storing this information in a derivative image; removing parallel lines from the resulting derivative image by searching along the normal of the current derivative slope to select for circular markers; searching along the direction of the current slope for the opposite side of the
20 circle, and thereby determining a candidate for the center point of the circle; voting for the most likely center point, using the accumulation of all of the center points; and reading out the highest weighted point to a CPU with a fast data transfer.

BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing and other features and advantages of the present disclosure will be more fully understood from the following detailed description of illustrative
5 embodiments, taken in conjunction with the accompanying drawings.

FIG. 1 is a perspective view of an embodiment in accordance with the present disclosure.

FIG. 2 is a front view of a live action image with a set of points around it representing a rough matte outline in accordance with an embodiment of the present
10 disclosure.

FIG. 3 is a perspective view of a 3D compositing scene in accordance with the present disclosure.

FIGS. 4A, 4B, 4C and 4D are image views of several steps in an edge detection process in accordance with the present disclosure.

15 FIG. 5 is an image view of several steps in a parallel line removal process in accordance with the present disclosure.

FIG. 6 is an image view of several steps in a circle center detection process in accordance with the present disclosure.

FIG. 7 is a data flow diagram in accordance with the present disclosure.

20 FIG. 8 is a data flow diagram showing an exemplary existing virtual studio system modified to include a hybrid precision tracking process of the present disclosure.

DETAILED DESCRIPTION

The following is a detailed description of the presently known best mode(s) of

carrying out the inventions. This description is not to be taken in a limiting sense, but is made merely for the purpose of illustrating the general principles of the inventions.

A rapid, efficient and reliable system for generating a rapid and precise camera and lens match that significantly speeds the integration of live action and virtual composite images is disclosed. Applications ranging from video games to feature films can implement the system in a fraction of the time typically spent tracking multiple areas of high contrast in the image by hand. The system thereby can greatly reduce the cost and complexity of matching virtual and live action backgrounds, and enables a much wider usage of the virtual production method.

Since the present process is primarily for joining live action with computer-generated elements, its applications for video games may be limited. The process works with a real-time video feed from a camera, which is also available on most “still” cameras. The process can work with a “video tap” mounted on a film camera, in systems where the image is converted to a standard video format that can be processed. An objective of the present disclosure is to provide a method and apparatus for creating rapid high precision measurement of camera rotation.

Referring to FIG. 1, an embodiment of the present disclosure is depicted. A scene camera 100 with a lens 110 is positioned to capture a live action image 90 (FIG. 2) of a subject 10 in front of a background 20. The subject(s) 10, for example, can be actors, props and physical sets. The background 20 can have markers 30 that are placed in known positions in the scene. Per an aspect of the disclosure, the centers of the markers 30 can be measured accurately in 3D space using a survey tool, and used to increase the accuracy of the camera solve.

The scene camera 100 can be mounted on a camera tracking system 120. And the camera tracking system 120 can be an encoded pedestal, dolly, jib, crane or any other form of camera position, orientation and field-of-view measuring system. Focus distance can also be measured, as the parameters of a lens can change while focusing. There may
5 be more than one scene camera to enable different views of the subject's performance to be captured.

The scene camera 100 and the camera tracking system 120 are connected to a video processing system 130, as depicted in FIG. 7. The video processing system 130 takes the incoming live action video image 90, recognizes the live action markers 30,
10 generates the corresponding background virtual marker XY locations of markers 40 (FIG. 2) and performs a delta angle minimization process using the two sets of XY data. The video processing system 130 can include a computer with a live video input, a camera tracking data input, and a video card capable of processing 2D and 3D computer graphics calculations.

15 The hardware components of the video processing system 130 can typically include a computer, a video input and output board that converts high definition serial digital video into the computer's memory, and a graphics card or GPU where the circle detection algorithms are run.

An embodiment of the disclosure is illustrated in FIG. 2. A live action subject 10
20 is shown in the center of a live action image 90. A virtual marker 40 is displayed along with the live action image 90 in a user interface 400 (FIG. 7). The virtual marker 40 is drawn at the predicted XY position where the corresponding surveyed marker 3D space (from surveying marker 30) would be projected onto the screen.

A 2D tracker 60 is displayed around the live action marker 30. The user selects which marker 30 to track by clicking on it. The live action marker 30 can be a circle, for example. This enables the computer to rapidly detect the center of the circle to subpixel resolution using one of a variety of machine vision transforms that are well known to those skilled in the art. Once the XY positions of the 2D tracker 60 and the corresponding virtual marker 40 are known, the error between the two (which is mostly caused by slight error in camera angle) can be minimized by creating a camera delta angle added to the incoming real-time camera tracking data that minimizes the distance between the live action and the virtual XY positions of the markers.

The 2D tracker 60 follows the position of the motion of marker 30. This can be done by using the incoming data from the newest frame of camera tracking data, adding the delta angle already found for the previous frame, and then calculating the new delta angle to minimize 2D error between the live action and virtual markers. In this way, the error between the two is continually minimized, keeping the solution error small and automatically handling high-speed motion.

The XY position of the virtual marker 40 can be calculated from the measured 3D position of the live action marker 30, and transformed to 2D coordinates using projection techniques that are well known to those skilled in the art, using the live action camera and lens information measured by camera tracking system 120. This produces a basic projected geometry XY position. To more accurately locate the XY position of the virtual marker 40, the system can take into account lens distortion, which can be measured as a lens parameter (see the above-mentioned lens calibration '633 patent) and used to adjust the basic XY projection to the distorted XY projection. This calculation, for example, can

be computed by the cvProjectPoints2 function in the OpenCV computer vision library.

In some cases, the incoming live action image must be “deblurred” by performing a reverse blur on the image, using the incoming camera motion data from the sensor. Since the camera motion is known, the amount of motion blur in the live action image
5 can be predicted and removed to increase tracking accuracy. In a preferred embodiment, the deblurring algorithm is a reverse Gaussian algorithm described below, and well known to practitioners in machine vision.

The display in the user interface is 2D and but for correct alignment, all of the various components exist as 3D objects in a virtual scene. Referring to FIG. 3, the
10 transformation of 3D to 2D coordinates can be done using standard projection geometry calculations that are well known to those skilled in the art. A virtual camera 70 has a frustum 80 that describes the field of view of the virtual camera. The parameters of the virtual camera 70 and frustum 80 match the parameters of the live action camera 100 and the lens 110.

15 A live action image 90 containing a subject 10 is located at a distance from the virtual camera 70. The virtual marker 40 is shown at the point where the system predicts it to be, while the actual marker 30 is located to the right. The 2D tracker 60 follows the live action marker, and the delta angle 82 is calculated between the predicted XY and the actual XY.

20 FIGS. 4A-4D represent several stages in the image processing to locate the centers of circular markers 30 within the scene. In FIG. 4A, circular marker 30 is shown with a pixel array 200 consisting of individual pixels. A digital image 90 is made up of millions of these individual picture elements (or pixels) 210. The edge of the circular

marker 30 can be seen to cause some of the pixels in array 200 to be dark and some to be light. For reference, a square marker 230 is shown in FIG. 4B to illustrate how the circle detection algorithm will reject square markers 230.

To achieve the speed required by real time operation, each of these operations can
 5 take place on a dedicated graphics processing unit, or GPU. While the GPU provides speed, it has a restricted set of programming instructions as compared to a general purpose central processing unit, or CPU.

The next step is to find the slope or derivative 220 of the local image luminance at a center pixel 228, as shown in FIG. 4A. This can be calculated for each pixel in the
 10 image 90, based around a center pixel 228 in grid 200. This can be calculated using a dot product operating over a surrounding area of pixels, which is a method well known in the literature. The process of finding the derivative can be calculated, for example, by defining top vector 222, mid vector 224 and bottom vector 226 to correspond to the pixels surrounding the current pixel being processed on the GPU. The slope 220 for this
 15 center pixel 228 can then be calculated with:

$$dx = \text{dot}(\text{vec}(-1.0, 0.0, 1.0), \text{top}) + \text{dot}(\text{vec}(-2.0, 0.0, 2.0), \text{mid}) + \text{dot}(\text{vec}(-1.0, 0.0, 1.0), \text{bot})$$

$$dy = \text{dot}(\text{vec}(1.0, 2.0, 1.0), \text{top}) + \text{dot}(\text{vec}(-1.0, -2.0, -1.0), \text{bot})$$

where $\text{dot}(x,y)$ is the dot product of the two vectors specified.

20 The slope can then be thresholded to find only points with high derivatives (such as the visually sharp edge of marker 30). This generates an image consisting only of pixels with the local derivative at the edges, such as the edges of the circle 240 in FIG. 4C and the square 250 in FIG. 4D. Square 250 is the high derivative image that is

extracted from square marker 230 in FIG. 4B.

A further step can then be made to reject parallel lines, such as those generated in square 250. An embodiment is shown in FIG. 5. Derivative image 250 contains a pixel 228 that contains x and y derivative information if it is over a threshold, and 0 if not. For derivatives that are over that threshold (edges), the x and y derivative values comprise the slope of the line that is normal to the local edge. The algorithm then calculates the direction 320 that is perpendicular to the local derivative 220.

The parallel line rejection algorithm then for each pixel 228 searches along the positive and negative directions of the perpendicular 320 to find an offset pixel 328, and compares the derivative found there to the original derivative in pixel 228. If the derivatives are identical, the value of pixel 328 is set to 0. This removes straight edges with constant slopes from consideration in the circle center finding algorithm, and preserves curves such as circle 240. In a preferred embodiment, the search distance can be an adjustable number of pixels, generally less than five. The resulting image can consist only of filtered circles 260 (FIG. 6).

In another embodiment, the system can perform a color difference extraction on the live action image, using techniques well known to practitioners in the art. In this way, the only areas that will be searched for circles are those that are shades of blue or green, enabling the use of darker color green circles on a lighter colored green screen as circular markers. This allows use of markers without interfering with the extraction of the actors from the green or blue screen.

A next step to determine the circle centers is made and an embodiment thereof is shown in FIG. 6. A pixel array 200 consists of the filtered derivative values of filtered

circles 260. For each pixel 228 with derivative magnitude values above a user specified threshold, the program searches along the direction of the derivative 220 over a minimum and maximum radius specified by the user. At the maximum derivative value found (the opposite edge of the circle, if it exists), the coordinates of the center of the found circle
5 310 (calculated by averaging the XY coordinates of the first edge with the second edge) are written into the current pixel value 228. This creates an array of found circle coordinates 462 (FIG. 7).

The “votes” for the circle coordinates are now accumulated. This can be accomplished with a geometry shader, which is a type of graphics processing program
10 well known to those skilled in the arts. The original purpose of the geometry shader is to create a new vertex when programmer specified conditions are met. The geometry shader is written to iterate over the array of found circle coordinates 462. When a pixel with previously written center coordinates is found, the shader emits a vertex to a buffer on the GPU that has been mapped to CPU main memory. In this way, the data transfer between
15 CPU and GPU is limited to a list of XY coordinates of the found circle centers, which is much lower bandwidth (and thus much faster) than reading back a pixel array to the CPU and doing a search for the various coordinates, slopes, etc. in the CPU. Although this would be much easier from a programming standpoint, it would not provide the real time performance that is required for this application.

20 Once the list of centers is in the CPU, the list can then be tallied up and averaged to find the groupings of found centers. The XY position of the marker center can be predicted, for example, from the camera and lens tracking system's position and orientation, so that each search only takes place centered around one dot. Thus, the

average of the found centers for that section is the found center of the circle.

This technique for finding the center of a circle can also be used for high speed alignment and tracking for many types of robotic positioning that need fast accuracy using small graphics processors.

5 Now, the actual XY positions of the trackers 60 centered around markers 30 are known. Since each tracker 60 has a corresponding predicted XY location 40, the delta angle 82 can be calculated as is shown in FIG. 3. In FIG. 3, the only error is a horizontal error that is corrected by calculating the delta angle 82. The delta angle 82 can be calculated, for example, for the camera pan, tilt, and roll angles by minimizing the sum of
10 the squares of the errors between the predicted XY positions 40 and the actual measured positions 60 of the images of the markers 30.

Since the delta angle 82 is computed for each new frame, image noise can affect the stability of the angle. Delta angle 82 must then be filtered to generate a visually smooth result. This filtering can be achieved, for example, by using a Kalman filter, a
15 method that is well known in the field.

FIG. 7 shows a data flow diagram in a further embodiment. Referring thereto, scene camera 100 generates a live action image 90, which is sent to a target XY predictor 490. Camera tracking system 120 measures the motion of camera 100 and sends camera data 122 to data combiner 420. Lens 110 sends lens position data 112 to lens calibration
20 table 410, which has been previously generated with a lens calibration system such as disclosed in the '663 patent. Using lens position data 112, the lens calibration table 410 can produce lens data 412, which is sent to the data combiner 420. Data combiner 420 sends the combined camera and lens data 310 to target XY predictor 490. Target XY

predictor 490 then predicts the XY positions of the live action markers 30 in the live action image 90, and sends an image subset 492 centered around the predicted position of live action marker 30 to the derivative detector 440.

The derivative detector 440 performs a derivative calculation and thresholding
5 operation on image subset 492, which results in derivative image 442. Parallel line remover 450 then removes from the derivative image 442 derivatives that are found to be part of parallel lines, leaving filtered image 452 that consists only of non-parallel edges such as circles. Circle center finder 460 then takes filtered image 452 and searches for the center of the circle predicted by each derivative circle. This results in an image 462 that
10 contains the accumulated circle center predictions. Center voting module 470 accumulates the predicted centers in circle center image 462, and reads out the average predicted center value 472. Since there is typically only one circle found within the predicted search area, this prediction is the center average of live action marker 30.

In user interface 400, the user enters the data for the surveyed live action marker
15 3D locations 32. This is sent to a point projection module 430. Using the combined data 310 from the data combiner 420, the point projection module 430 generates a set of predicted 2D virtual marker positions 40, each of which corresponds to an actual circle center average 472. These two data streams are sent to an XY error minimizer 480, which calculates the adjusted pose 482 that provides the minimum error between the predicted
20 and actual circle centers. To remove noise in the solution, the adjusted pose 482 is sent to noise filter 500, which performs a noise reduction calculation on adjusted pose 482 to produce filtered pose 502, which can be the final output.

FIG. 8 is a data flow diagram showing an existing virtual studio system, such as

the PREVIZION system, modified to include a hybrid precision tracking process (shown in the middle of the drawing by box 800) of the present disclosure, and as discussed above. The PREVIZION system includes a camera tracking system 120, a lens calibration table 410 that interprets lens position data 112 from lens 110, a data combiner 420 to bring the camera data 122 and lens data 412 data together, a 2D compositing engine 620 and a 3D renderer 610. The system reads camera data 122 and lens position data 112 along with live action image 90, creates a rendered 3D virtual image 612, and combines this with the live action image 90 in the 2D compositing engine 620 to create a finished composite image 630.

Previously, the 3D camera pose data went directly from the data combiner 420 to the 3D renderer. In contrast a system herein (and as depicted in FIG. 8, for example) can use the predicted camera location from the data combiner 420 along with the 3D locations of the surveyed tracking markers, and calculate a new filtered pose 502. Filtered pose 502 can then be sent to a 3D renderer 610 to generate the correctly aligned rendered 3D image 612, which is then composited with live action image 90 in a 2D compositing engine 620 to generate a finished composite image 630 with the virtual set extension correctly aligned to the live action set.

PREVIZION is a real-time visual effects system that provides high precision camera tracking, sophisticated rendering and VFX quality keying, which makes it ideal for the real time compositing of virtual backgrounds and CGI characters for on-set production. The PREVIZION system provides filmmakers with the freedom to create visual effects shots with dynamic camera moves and actually see the completed effects while shooting green screen or virtual elements. This ability enables producers to use

high quality visual effects with a minimal increase in production costs and significant savings in post production.

Referring again to FIG. 7, two major areas of the present disclosure are shown. The first area depicted by dotted line block 600 can include elements 490, 420 and 410.

5 It relates to the use of a calibrated look-up table to adaptively modify the search locations for scene markers when using adjustable lenses, such as those used in motion picture production.

The second area depicted by dotted line block 700 can include elements 440, 450, 460 and 470. The second area 700 relates to the use of a high speed, GPU based circle
10 center detection scheme using a fast circle geometry based search. This keeps the operation on the GPU except for the final readback of circle centers in the voting process, and provides a speed increase over other traditional circle center finding algorithms such as the Hough transform.

Another novel area/aspect of this disclosure is the use of arbitrary camera tracking
15 data 122 coming in from camera tracking system 120. An embodiment of the present system can work with arbitrary external tracking sources, such as encoders, GPS and so forth.

A flow chart of a software process of the present disclosure can be essentially that of the video processing system 130 in FIG. 7.

20 As mentioned above the current camera motion data from the camera tracking system can be used to generate a “deblurred” image of the marker. This can be important in order to handle tracking with any degree of camera motion.

The main steps in this “deblurring” process can include:

- For each pixel, using the present and previous camera position and orientation, calculate the 2D motion vector.
- Take several pixel samples along the 2D motion vector, starting with the present XY position.
- 5 - Add a cumulative total of the color samples to one buffer, sumColor.
- Add a cumulative total of the differences between the original pixel color and the sampled colors.
- Divide the color sample total by the number of samples to get an average.
- Subtract the averaged color from the current pixel to get a buffer with the removed color
- 10 values.
- Generate the final unblurred image by linearly interpolating between the original color image and the removed average, using the color difference buffer as the weighting factor.

This can work out to the following code:

```

vec2 sCoord = coord;
15    vec3 sumColor = vec3(0.0, 0.0, 0.0)
       float diff = 0.0;
       for(int iSample = 0; iSample < samples; iSample++)
       {
        sCoord += coordInc;
20    vec3 sColor = texture(image, sCoord).rgb;
        sumColor += sColor;
        diff += abs(dot(vec3(1.0, 1.0, 1.0), (sColor - color)));
       }
       sumColor /= float(numSamples);
25    float sumMult = 0.5;
       float mult = (1.0/(1.0 - sumMult));
       vec3 removed = mult * (color - (sumMult * sumColor));

```

```
diff *= scale;  
diff = min(diff, 1.0);  
vec3 unblurred = mix(color, removed, diff);
```

5 Examples of ways/methods in which the present disclosures can be used, commercialized and/or built can include: as a feature as part of Lightcraft Technology LLC's existing virtual studio (PREVIZION) system; as an add-on software or hardware component for other existing virtual studio systems such as ORAD, Viz-RT, and Brainstorm; as an embedded program running inside a digital motion picture camera; and
10 as a software addition to a post production tool. The additional hardware component can be a small box with a HDSDI video input for the incoming signal, serial or network inputs for the incoming tracking and lens data, and a serial or network output for the refined tracking data. The user can optionally turn on a menu overlay to adjust the system parameters. The add-on software can be transmitted over the Internet, on a disk, USB
15 drive, etc. The software can be in the form of a (non-transitory) (tangible) computer-readable medium (e.g., a disk, hard drive, or other data storage device) containing program instructions for causing a computer to perform a particular process, as disclosed herein.

 That is, the present disclosure can be used as a software addition to a post-
20 production tool. For this tool, the software code herein can be easily ported over. That is, since the code already works in real time, those skilled in the art would know from this disclosure how to make it work in non-real time.

 The above-mentioned add-in software can be compiled directly into the PREVIZION product, and so can be transmitted to the user by sending him a new build

with the feature enabled. In a separate hardware form, it can be sent to a user as a separate box.

Adapting the present disclosure in a real-time virtual studio system, aside from that of the PREVIZION system, can be done, for example, by making a hardware box
5 that contains most of the code and hardware used for the PREVIZION system minus the 3D rendering code. (The video hardware I/O, serial connections to the external camera and lens tracking devices, and a GPU on which to run the high speed marker detection functions are still used.)

In a preferred embodiment, the video hardware I/O card can be a MerlinIP
10 HDSDI card from Skymicro Inc. of Simi Valley, CA. The serial connections can be through a standard serial interface card known to computer technology practitioners. The GPU can be a nVidia Quadro type GPU. The main computer can be a standard graphics workstation type machine, the specifics of which are well known to practitioners of computer graphics.

15 Although the inventions disclosed herein have been described in terms of preferred embodiments, numerous modifications and/or additions to the above-described embodiments would be readily apparent to one skilled in the art. The embodiments can be defined, for example, as methods carried out by any one, any subset of or all of the components as a system of one or more components in a certain structural and/or
20 functional relationship; as methods of making, installing and assembling; as methods of using; as methods of commercializing; as kits of the different components; as an entire assembled workable system; and/or as subassemblies or sub-methods. It is intended that the scope of the present inventions extend to all such modifications and/or additions.

WHAT IS CLAIMED IS:

1. A hybrid through-the-lens tracking system, comprising:
 - a first system that includes a calibrated look-up table and is configured to use the calibrated look-up table to modify search locations for scene markers when using adjustable lenses for motion picture and television visual effects production;
 - the first system is configured to transmit predicted 2D locations of the scene markers to a second system; and
 - the second system is configured to use a high speed GPU-based scene marker detection scheme that uses a fast geometry based search to calculate the predicted 2D locations of the centers of the scene markers.
2. The system of claim 1 wherein the high speed detection is under 10 milliseconds per marker.
3. The system of claim 1 wherein the high speed detection is between 2 and 10 milliseconds per marker.
4. The system of claim 1 wherein the first system further includes a data combiner operatively connected to the look-up table and a target XY predictor operatively connected to the data combiner.
5. The system of claim 1 wherein the system is embodied in a hardware component.

6. The system of claim 1 wherein the second system includes a motion blur removal stage, a derivative detector, a parallel line remover, a circle center finder, and a center voting module.
7. The system of claim 1 wherein the scene markers are circular markers.
8. The system of claim 1 wherein the scene markers are fiducial markers, bar code markers or natural feature markers.
9. A hybrid through-the-lens tracking method, comprising:
subjecting surveyed live action scene marker locations, live action images and combined camera and lens data to a hybrid precision tracking process to produce filtered camera pose data.
10. The method of claim 9 wherein the hybrid precision tracking process includes:
using the incoming live camera and lens data to predict the 2D position of the locations of on screen markers; searching within an area around the predicted 2D position for markers; calculating the actual XY position of the markers; modifying the camera pose data to align the predicted positions with the actual positions; and filtering the resulting angular changes to remove noise from the process.
11. The method of claim 10 wherein the markers are circular markers.
12. The method of claim 10 wherein the 2D location prediction step is before the

center position solving step.

13. The method of claim 9 further comprising subjecting the filtered pose to 3D rendering to produce a rendered 3D image.

14. The method of claim 13 further comprising 2D compositing the live action image and the rendered 3D image to produce a composite image.

15. A hybrid through-the-lens tracking method, comprising:
reading current incoming camera and lens information;
using the lens position information to look up the current lens parameters;
using the stored lens parameters and camera information to predict the location of 2D markers in the scene;
detecting the positions of the actual 2D markers in the scene; and
calculating the necessary angular corrections from the incoming camera position data to best correct the camera's angular position to match the actual 2D markers.

16. The method of claim 15 wherein the information includes the current measured camera position and orientation, and the lens parameters include focal length, distortion, and entry pupil location.

17. The method of claim 16 wherein the current measured camera position and orientation and the lens parameters are used to predict the XY position of the measured 3D

locations of the scene markers when photographed by the camera.

18. The method of claim 15 wherein the detecting the positions of the markers includes: deblurring to get a sharp circle; derivative detecting to find sharp edges of the markers; removing parallel lines to remove things that are not circles; center detecting to find the possible centers of circles; and center voting to detect the actual centers.

19. A hybrid precision tracking method, comprising:
calculating a filtered pose using predicted camera location data, 3D location data of surveyed tracking markers and a live action image.

20. The method of claim 19 further comprising sending the filtered pose to a 3D renderer to generate a correctly aligned rendered 3D image.

21. The method of claim 20 further comprising compositing the correctly aligned rendered 3D image with a live action image to generate a composite image.

22. The method of claim 20 wherein the calculating step includes:
measuring the camera's position and orientation;
measuring the adjustment position of the camera lens;
looking up the optical parameters of the lens, based on the measured adjustment lens positions;
predicting the on-screen XY position of surveyed physical locations in the scene, using

the camera and lens information;

searching in an area surrounding the predicted XY position for a detectable marker that represents the actual position of the physical location;

measuring the difference in camera angular movement required to minimize the error between the predicted and actual physical XY location; and

filtering the angular difference value over time to remove any angular noise introduced by the measuring steps.

23. The method of claim 22 wherein the detectable marker is a circle.

24. The method of claim 22 wherein the detectable marker is a fiducial marker or a natural feature marker.

25. A hybrid through-the lens tracking method, comprising:

using a calibrated look-up table of a first system to modify search locations for scene markers when using adjustable lenses for motion picture and television visual effects production;

transmitting, using the first system, predicted 2D locations of the scene markers to a second system; and

using the first system and the predicted 2D locations to minimize the reprojection error between the estimated positions and the actual detected positions of the scene markers.

26. A method of finding the center of a circle, comprising:

detecting the edge or slope of an image luminance or color, and storing this information

in a derivative image;

removing parallel lines from the resulting derivative image by searching along the normal of the current derivative slope to select for circular markers;

searching along the direction of the current slope for the opposite side of the circle, and thereby determining a candidate for the center point of the circle;

voting for the most likely center point, using the accumulation of all of the center points; and

reading out the highest weighted point to a CPU with a fast data transfer.

27. The method of claim 26 wherein the fast data transfer is a geometry shader followed by a readback operation.

28. The method of claim 26 wherein the detecting, removing, searching and voting steps are conducted by a set of fragment and geometry shaders operating on a graphics processing unit.

29. The method of claim 26 wherein the voting includes traversing through the set of found center positions in the image, and writing out a vertex through a geometry shader for each found circle center.

30. A hybrid through-the-lens tracking method, comprising:
subjecting surveyed live action scene marker locations, live action images and combined camera and lens data to a hybrid tracking process to produce filtered camera pose data, the

hybrid tracking process using adjustable lenses for motion picture and television visual effects production wherein the lenses are zooming and focusing during shooting;

wherein the hybrid tracking process includes: using the incoming live camera and lens data to predict the 2D position of the locations of on-screen markers; searching within an area around the predicted 2D position for markers; calculating the actual XY position of the markers; modifying the camera pose data to align the predicted positions with the actual positions; and filtering the resulting angular changes to remove tracking motion noise from the process; and

wherein the hybrid tracking processes uses an external camera tracking system that uses tracking devices to calculate the initial pose of the camera that is then refined through the hybrid precision tracking process.

31. The method of claim 30 wherein the filtering includes removing sudden angular changes in the background caused by a tracking marker moving onto or off of the visible image.

32. The method of claim 31 wherein the filtering uses a Kalman or other mathematical filter to remove sudden variations in the tracking camera angles that are inconsistent with the previous motion of the camera.

33. The method of claim 30 wherein the tracking devices include overhead markers or motion encoders.

34. The method of claim 30 wherein the markers are circular markers.

35. The method of claim 30 wherein the 2D location prediction step is before the center position solving step.

36. A hybrid through-the-lens tracking method, comprising:
reading current incoming camera pose and adjustable zooming and focusing lens information;
using the lens position information to look up the current lens parameters;
using the stored lens parameters and camera pose information to predict the location of 2D markers in the scene;
detecting the positions of the actual 2D markers in the scene;
calculating angular corrections from the incoming camera position data to correct the camera's angular position to match the actual 2D markers; and
correcting the camera's angular position using the angular corrections and thereafter filtering the corrections to minimize motion noise and improve tracking accuracy.

37. The method of claim 36 wherein the information includes lens distortion information, and further comprising using the lens distortion information to render the virtual background with correct distortion and thereby match the live action image.

38. The method of claim 36 wherein the information includes the current measured camera position and orientation, and the lens parameters include focal length, distortion, and entry pupil location; and further comprising using the current measured camera position and orientation and the lens parameters to predict the XY position of the measured 3D locations of

the scene markers when photographed by the camera.

39. A hybrid through-the-lens tracking system, comprising:

a first system that includes a calibrated look-up table and is configured to use the calibrated look-up table to modify search locations for scene markers when using adjustable lenses for motion picture and television visual effects production wherein the lenses are zooming and focusing during shooting;

the first system is configured to transmit predicted 2D locations of the scene markers to a second system;

the second system is configured to use a high speed GPU-based scene marker detection scheme that uses a fast geometry-based search based on the predicted 2D locations to calculate the actual 2D locations of the centers of the scene markers; and

the first system includes a data combiner operatively connected to the look-up table and a target XY predictor operatively connected to the data combiner; and

the second system uses the actual 2D positions of the scene markers and the original position of the camera to calculate angular corrections to the camera, to increase the accuracy of the known angular position of the scene camera.

40. The system of claim 39 wherein the target XY predictor is configured to calculate the estimated XY position of the surveyed markers in the image visible to the scene camera.

41. The system of claim 39 wherein the calculating the actual 2D marker locations in the live action image is at motion picture and television frame rates.

42. A hybrid through-the-lens tracking system, comprising:

a first system that includes a calibrated look-up table and is configured to use the calibrated look-up table to modify search locations for scene markers when using adjustable lenses for motion picture and television visual effects production and wherein the lenses are zooming and focusing during shooting;

the first system is configured to transmit predicted 2D locations of the scene markers to a second system;

the second system is configured to use a high speed GPU-based scene marker detection scheme that uses a fast geometry based search to calculate the actual 2D locations of the centers of the scene markers; and

the fast geometry based search using camera position data and the centers of markers discovered by the geometry search to correct the camera's angular position to match actual 2D markers.

43. The system of claim 42 wherein the second system includes a derivative detector, a parallel line remover, a circle center finder, and a center voting module.

44. The system of claim 43 wherein the derivative detector detects areas of high contrast, the parallel line remover filters out high contrast areas that fall along a line, and the circle center finder calculates a series of estimated center locations, which the center voting module then adds up to determine the actual XY center of the circle.

45. The system of claim 42 wherein the second system uses motion data from the external camera tracking system to deblur the incoming live action marker image.

46. A hybrid through-the-lens tracking method, comprising:
calculating pose data in a hybrid tracking process that uses predicted XY scene marker location data from a data combiner with 3D location data of surveyed tracking markers;
the lens of the scene camera being a zooming and focusing lens;
generating correctly aligned rendered 3D image in a 3D renderer using the pose data; and
compositing the rendered 3D image with a live action image in a 2D compositing engine to generate a composite image in real time.

47. The method of claim 46 wherein the generating includes correcting the camera's angular position using camera position data and the actual XY center locations of scene markers in the live action scene camera image.

48. The method of claim 46 further comprising using data from an external camera tracking system to deblur the incoming live action image of the markers.

49. The method of claim 46 wherein the markers are located in the visible field of view of the scene camera.

50. A hybrid precision tracking method, comprising:
calculating a filtered pose using predicted camera location data, 3D location data of

surveyed tracking markers and a live action image;

wherein the calculating step includes:

measuring the camera's position and orientation;

measuring the adjustment position of the camera lens;

looking up the optical parameters of the lens, based on the measured adjustment lens positions;

predicting the on-screen XY position of surveyed physical locations in the scene, using the camera and lens information;

searching in an area surrounding the predicted XY position for a detectable marker that represents the actual position of the physical location;

measuring the difference in camera angular movement required to minimize the error between the predicted and actual physical XY location; and

filtering the angular difference value over time to remove any angular noise introduced by the measuring steps;

sending the filtered pose to a real time 3D renderer to generate an aligned rendered 3D image; and

using the aligned rendered 3D image to composite a finished image by combining the live action image with the aligned rendered 3D image.

51. The method of claim 50 wherein the detectable marker is a circle.

52. The method of claim 50 wherein the detectable marker is a fiducial marker or a natural feature marker.

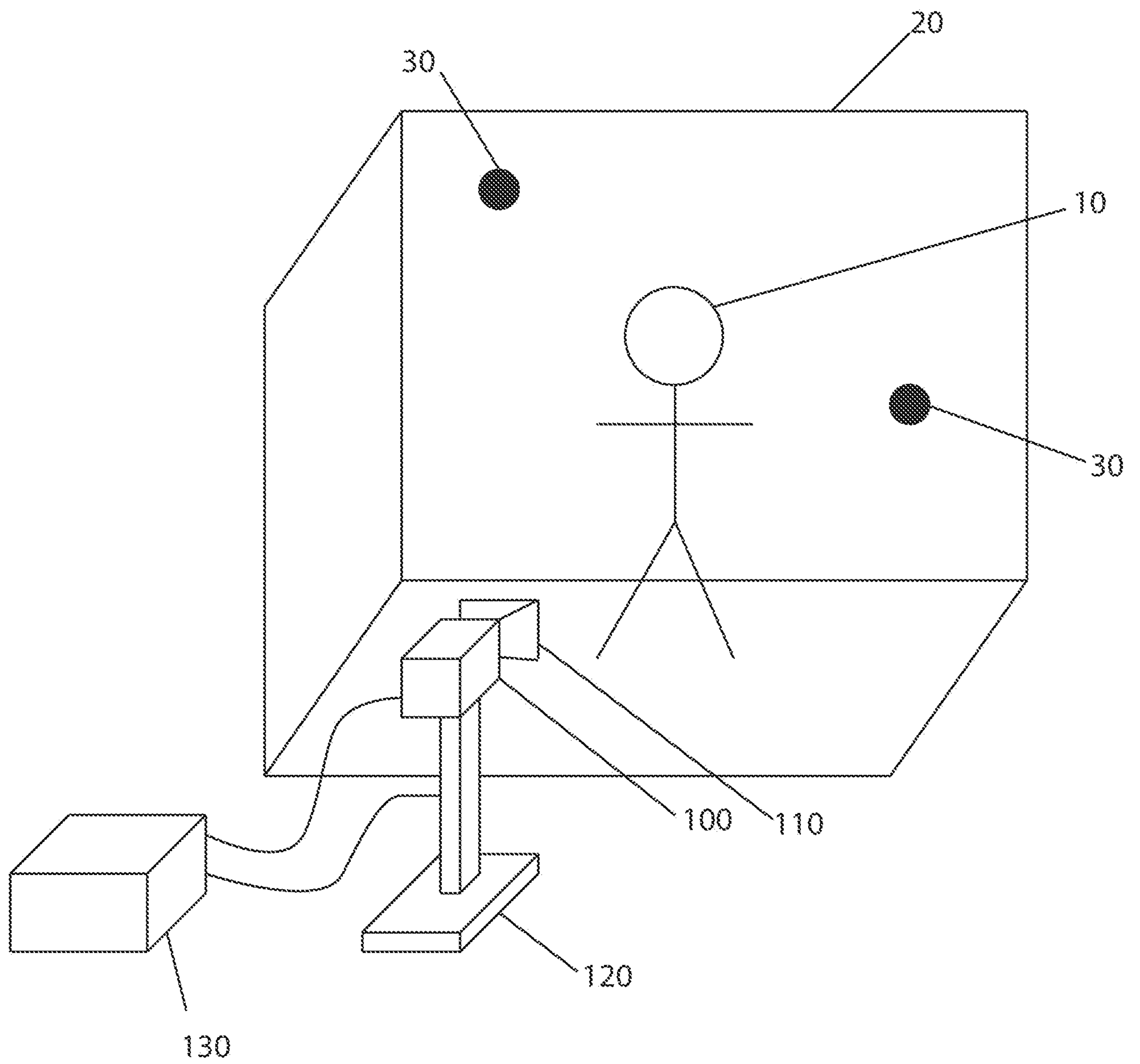
53. A hybrid through-the lens tracking method, comprising:

using a calibrated look-up table of a first system to modify search locations for scene markers when using adjustable lenses for motion picture and television visual effects production wherein the lenses are zooming and focusing to generate predicted 2D locations of the scene markers;

using the first system, the predicted 2D locations, and the actual found 2D locations of the scene markers to minimize the reprojection error between the estimated positions and the actual detected positions of the scene markers; and

transmitting, using the first system, the actual found 2D locations to a second system, which uses the calculated angular corrections to render an accurately located virtual image using updated camera pose and lens distortion data that is aligned at a sub-pixel resolution to the live action image at the location of the scene markers.

1/8

**FIG. 1**

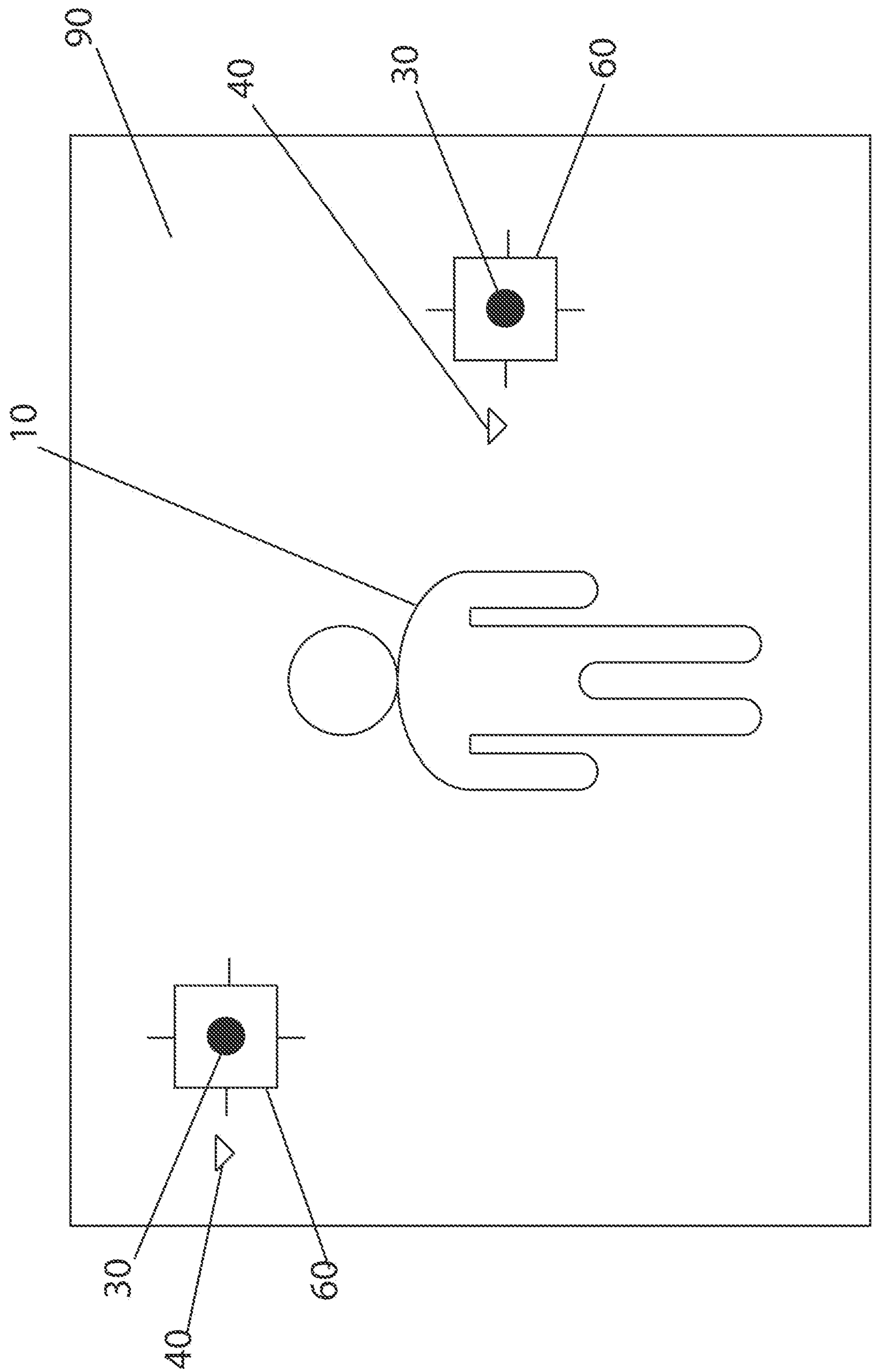
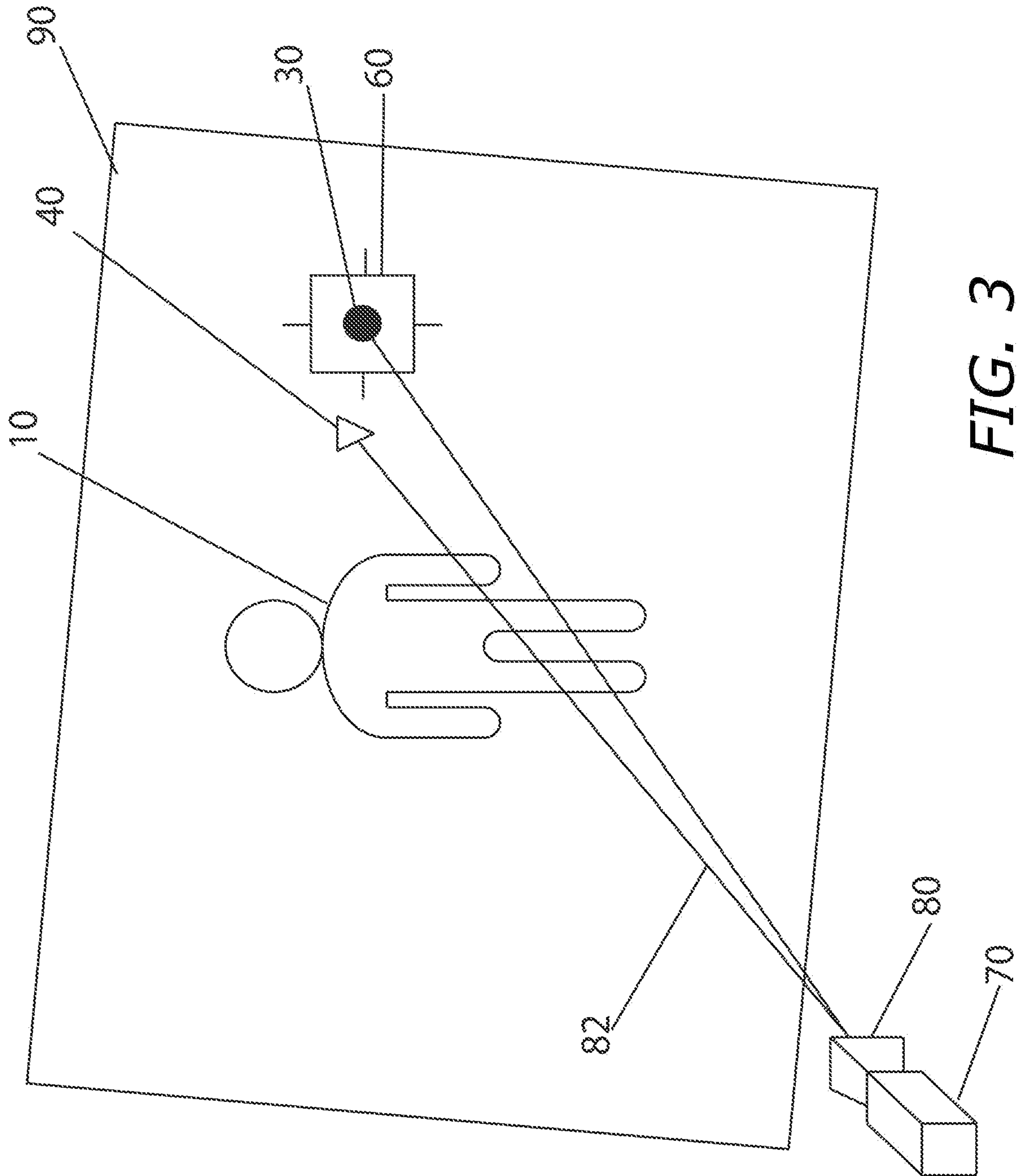
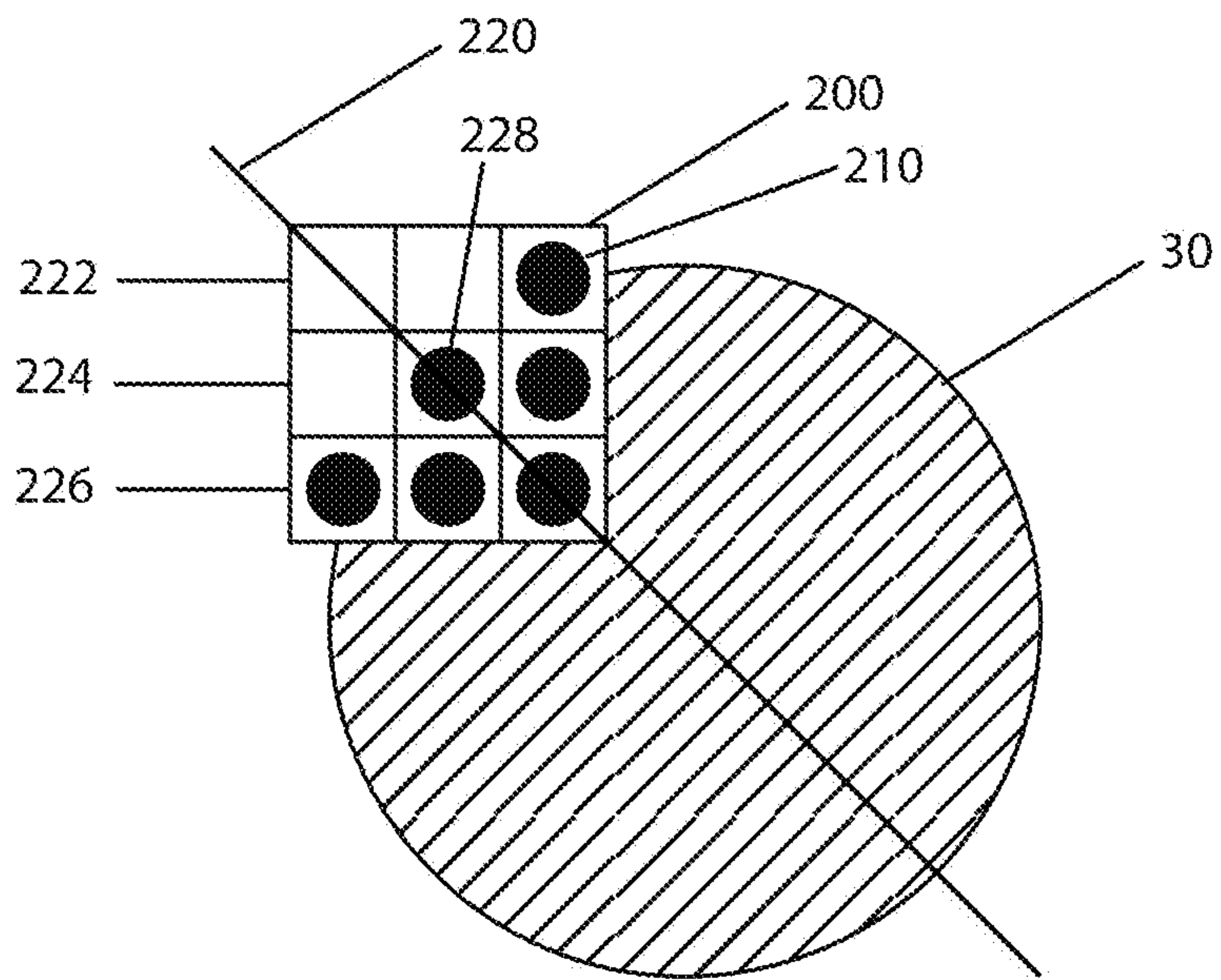
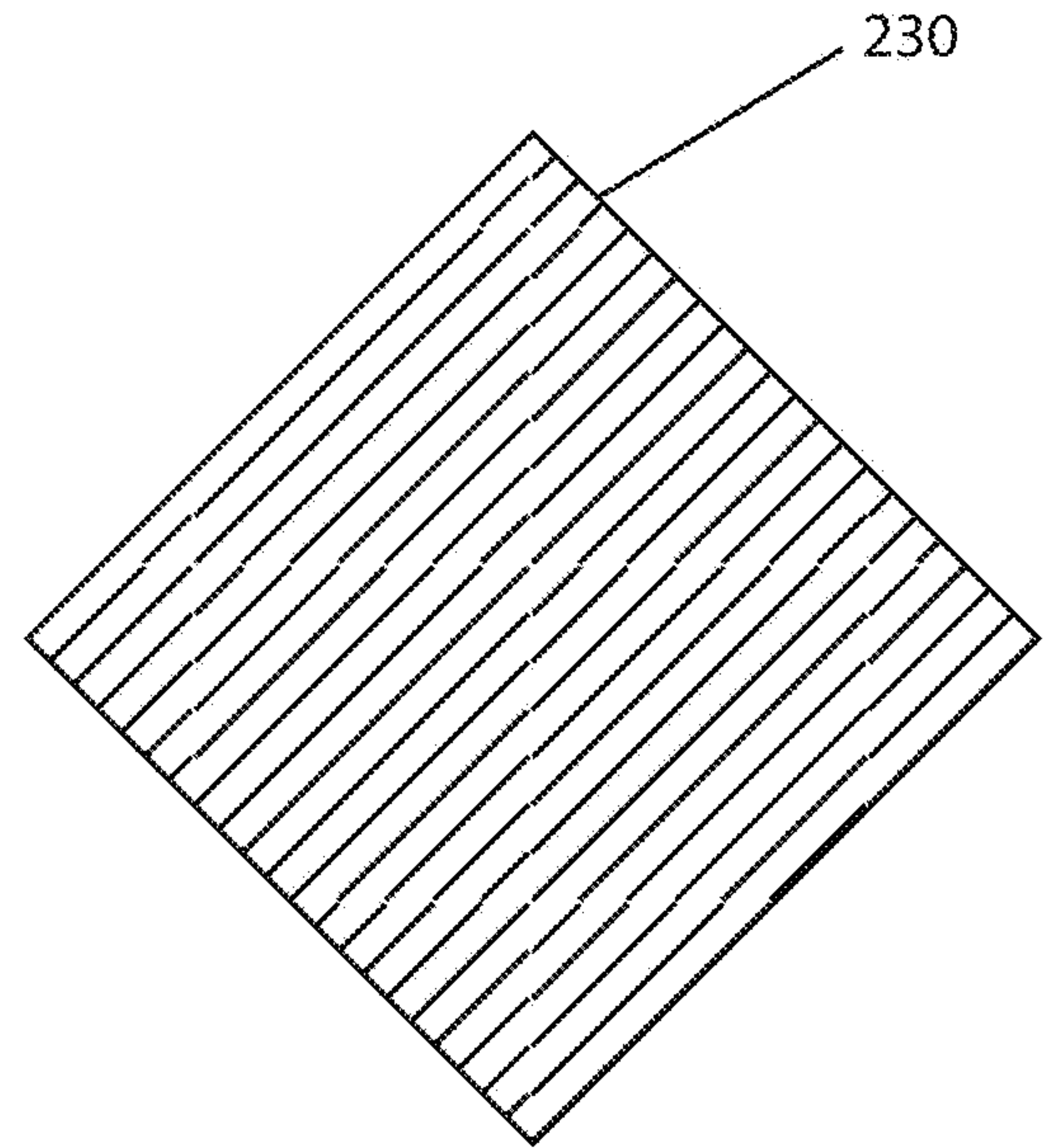
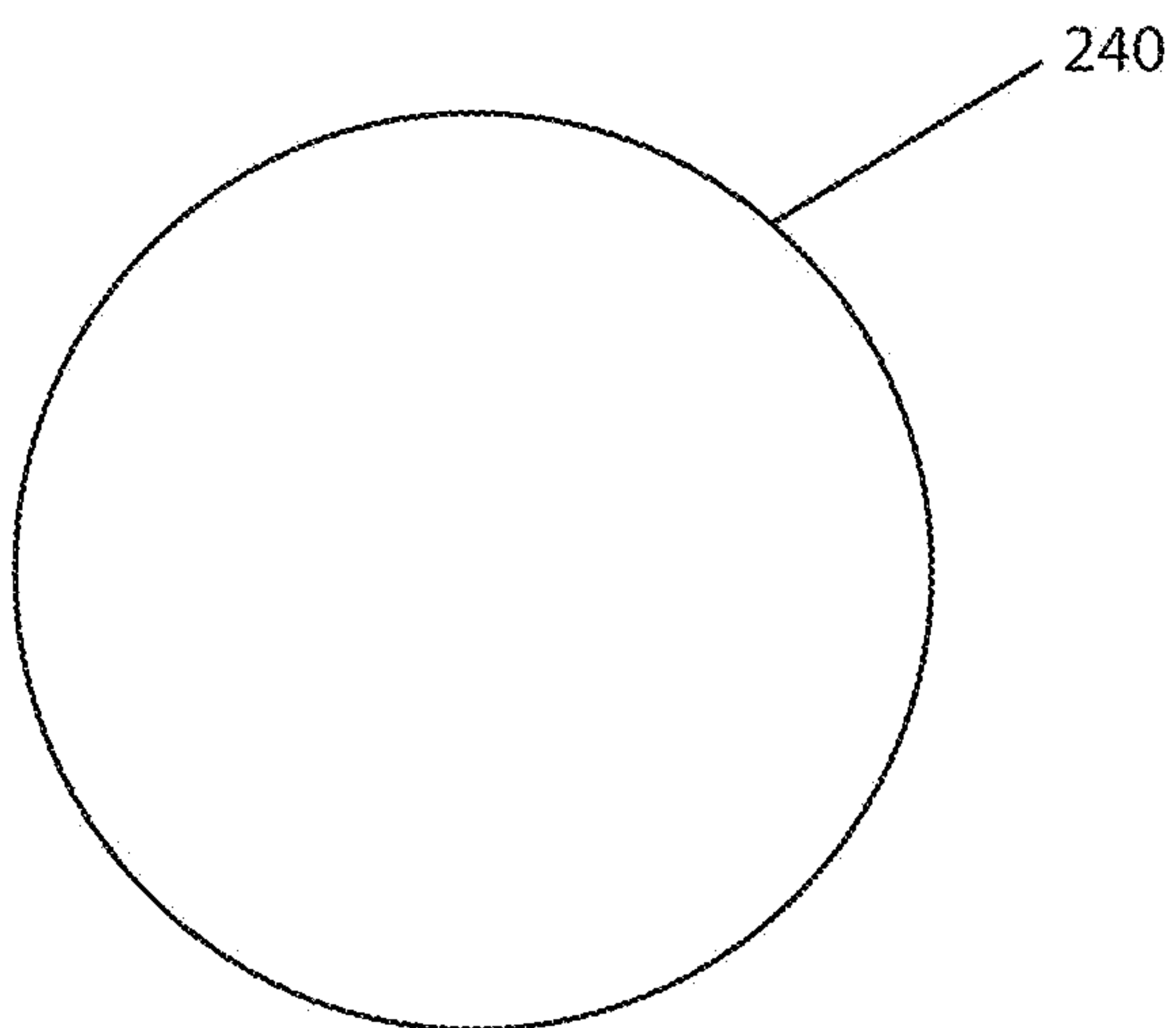
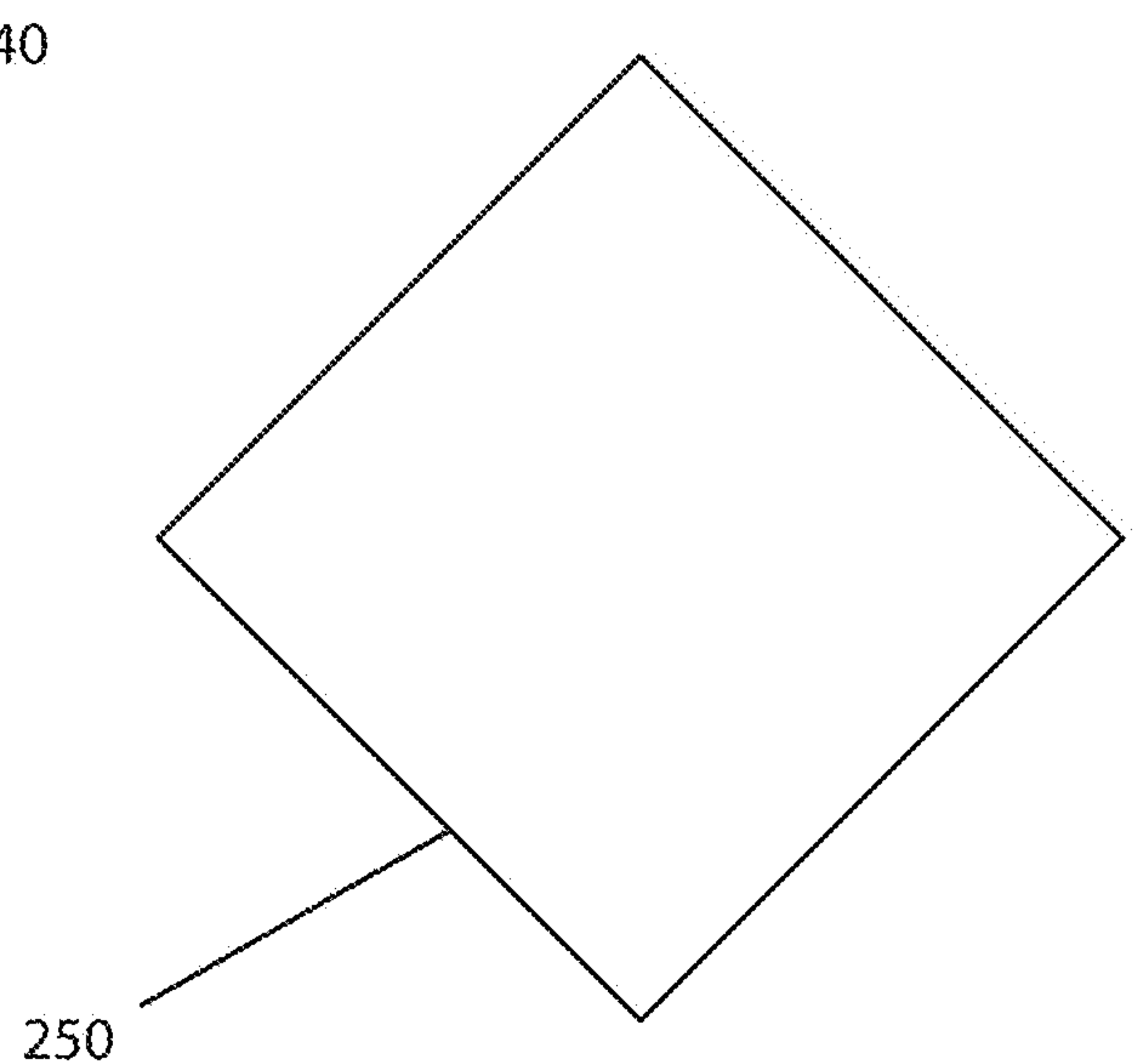


FIG. 2

3/8



4/8

**FIG. 4A****FIG. 4B****FIG. 4C****FIG. 4D**

5/8

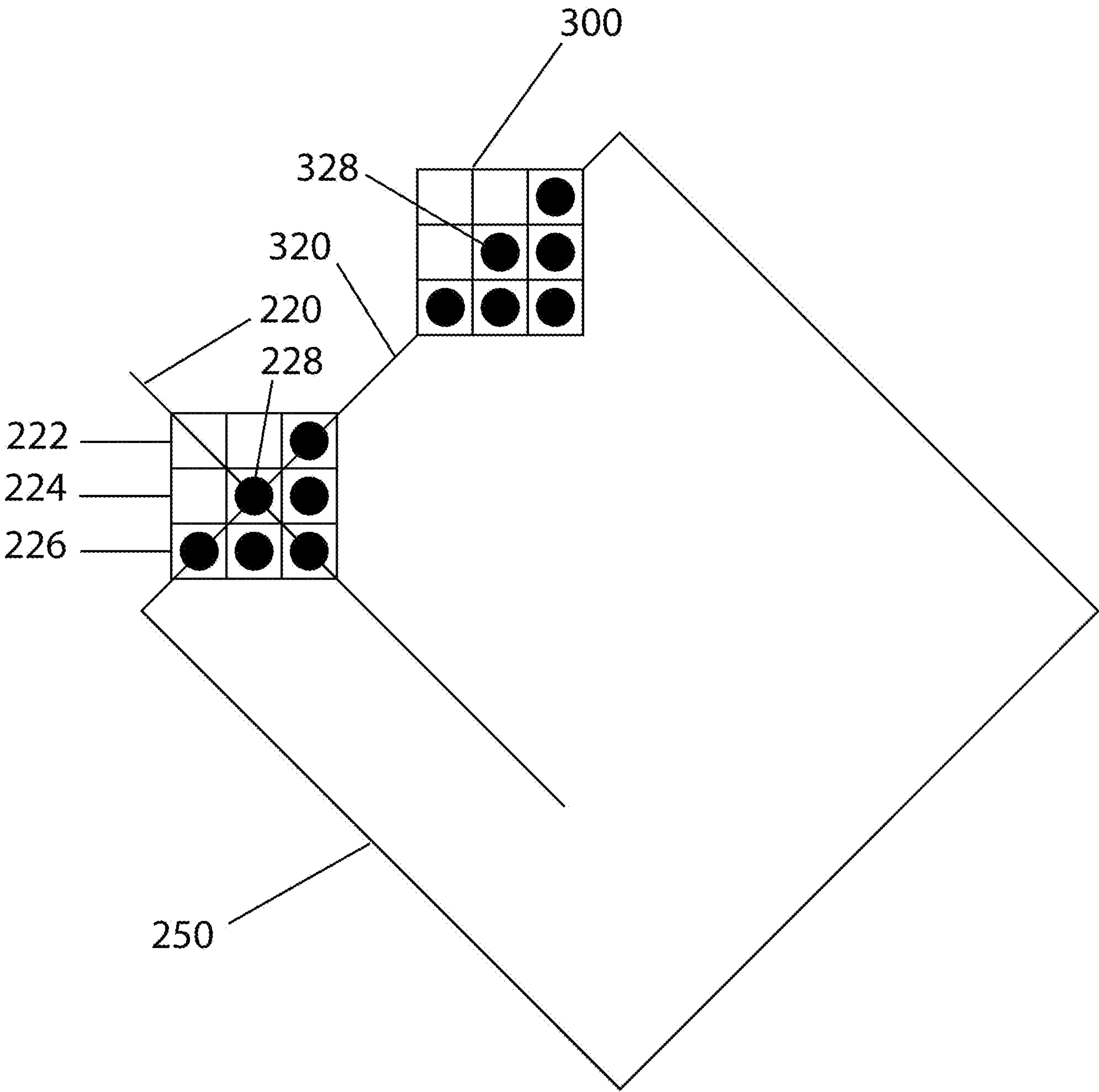
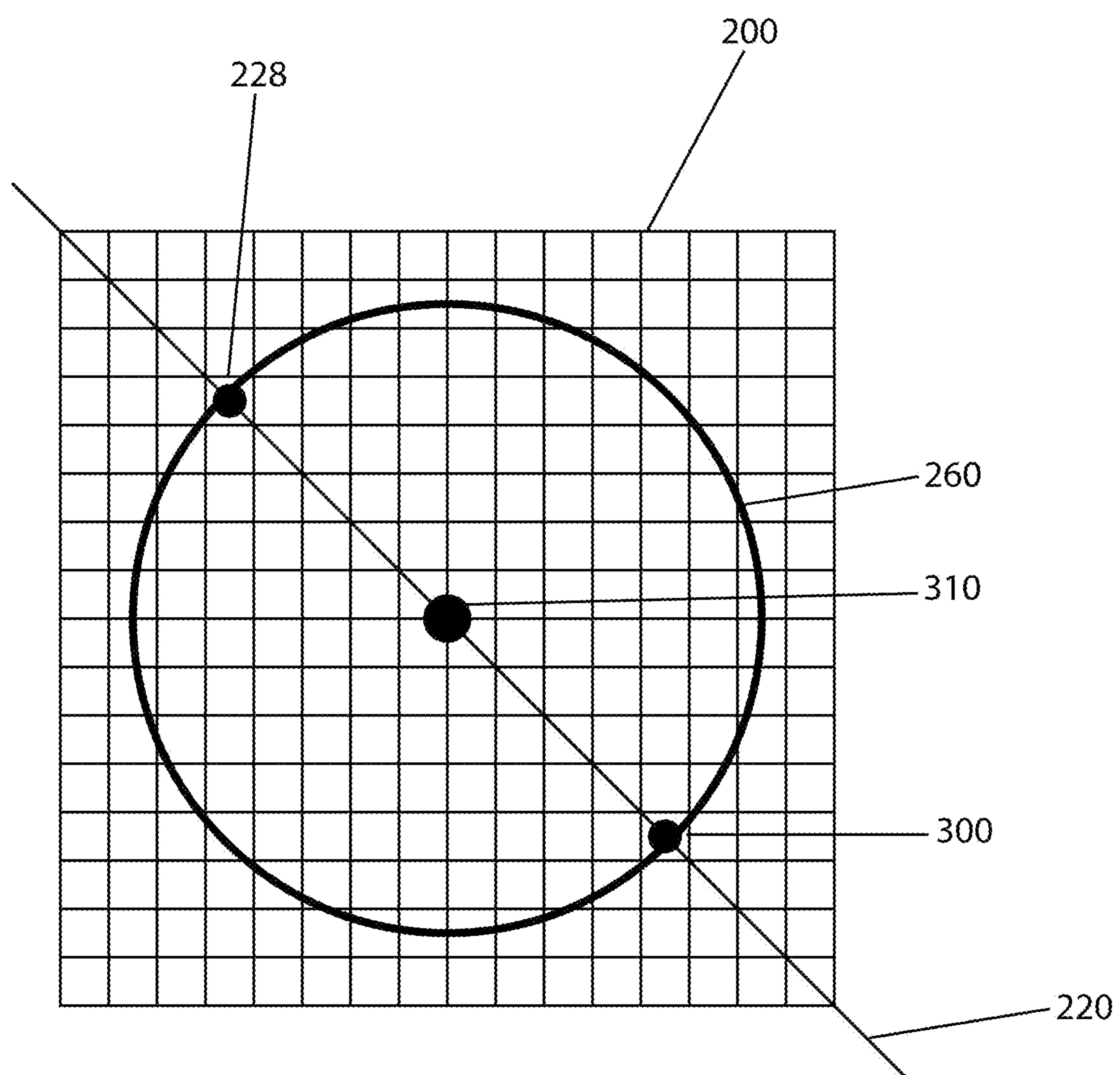


FIG. 5

6/8

*FIG. 6*

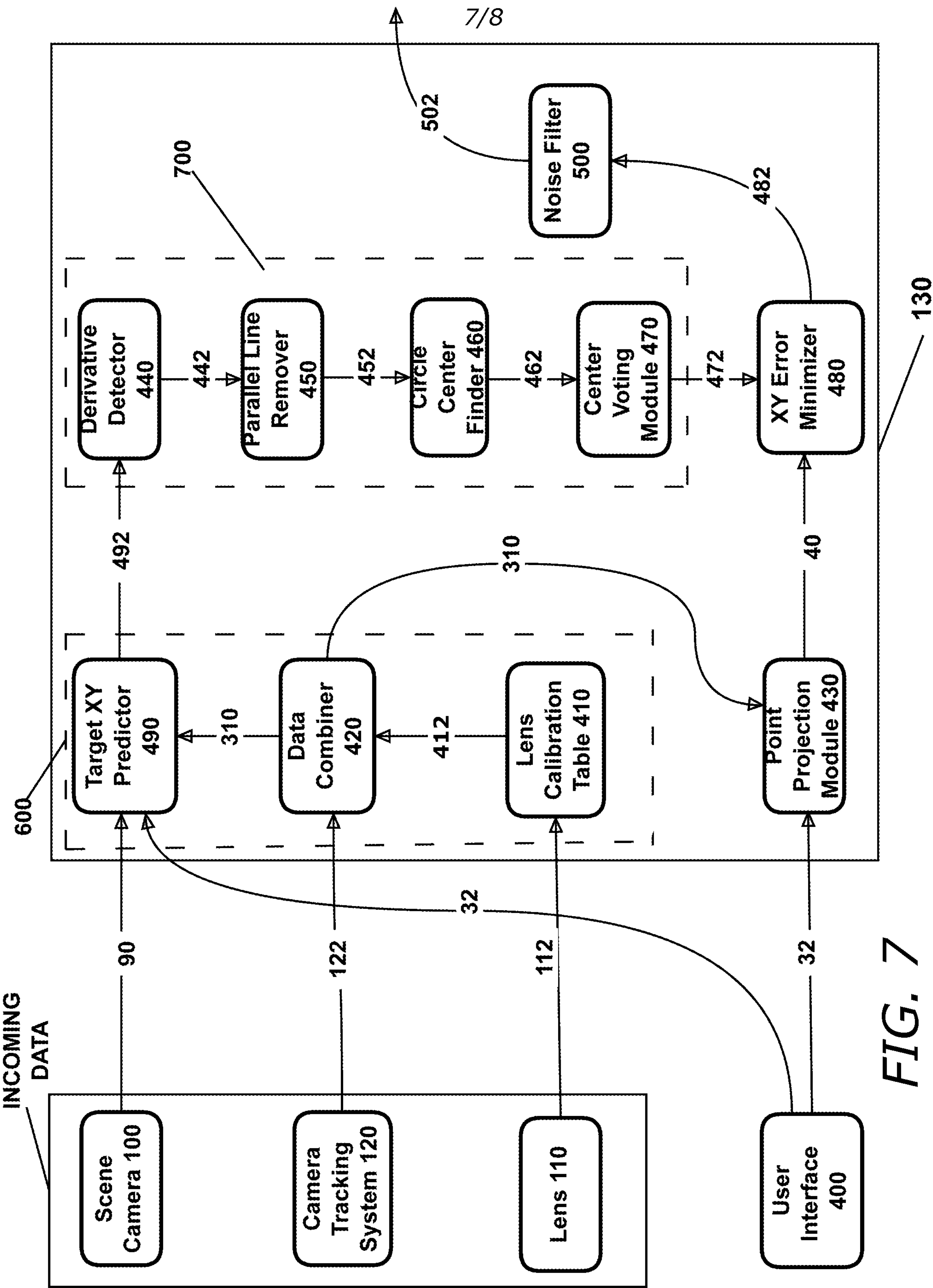


FIG. 7

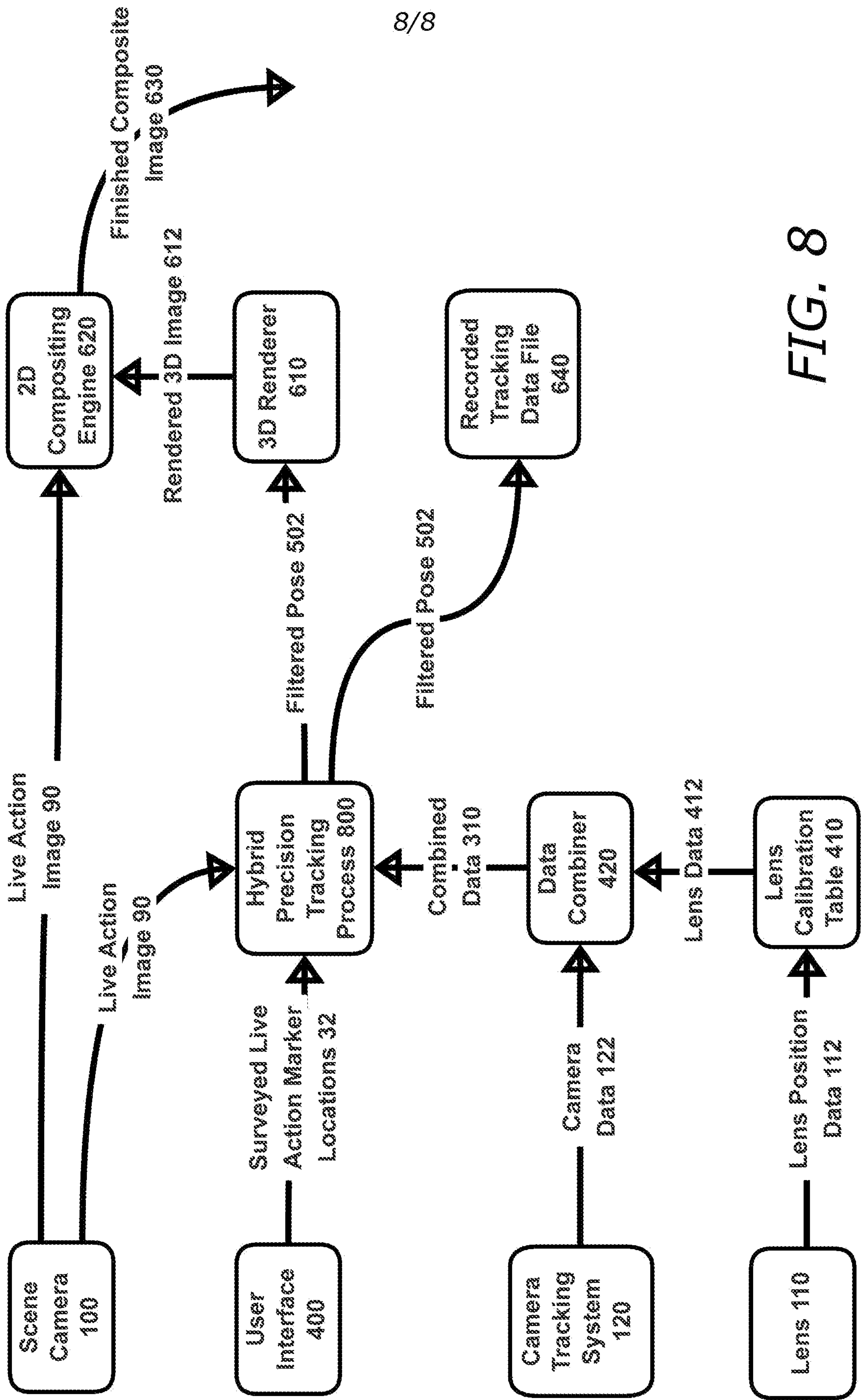


FIG. 8

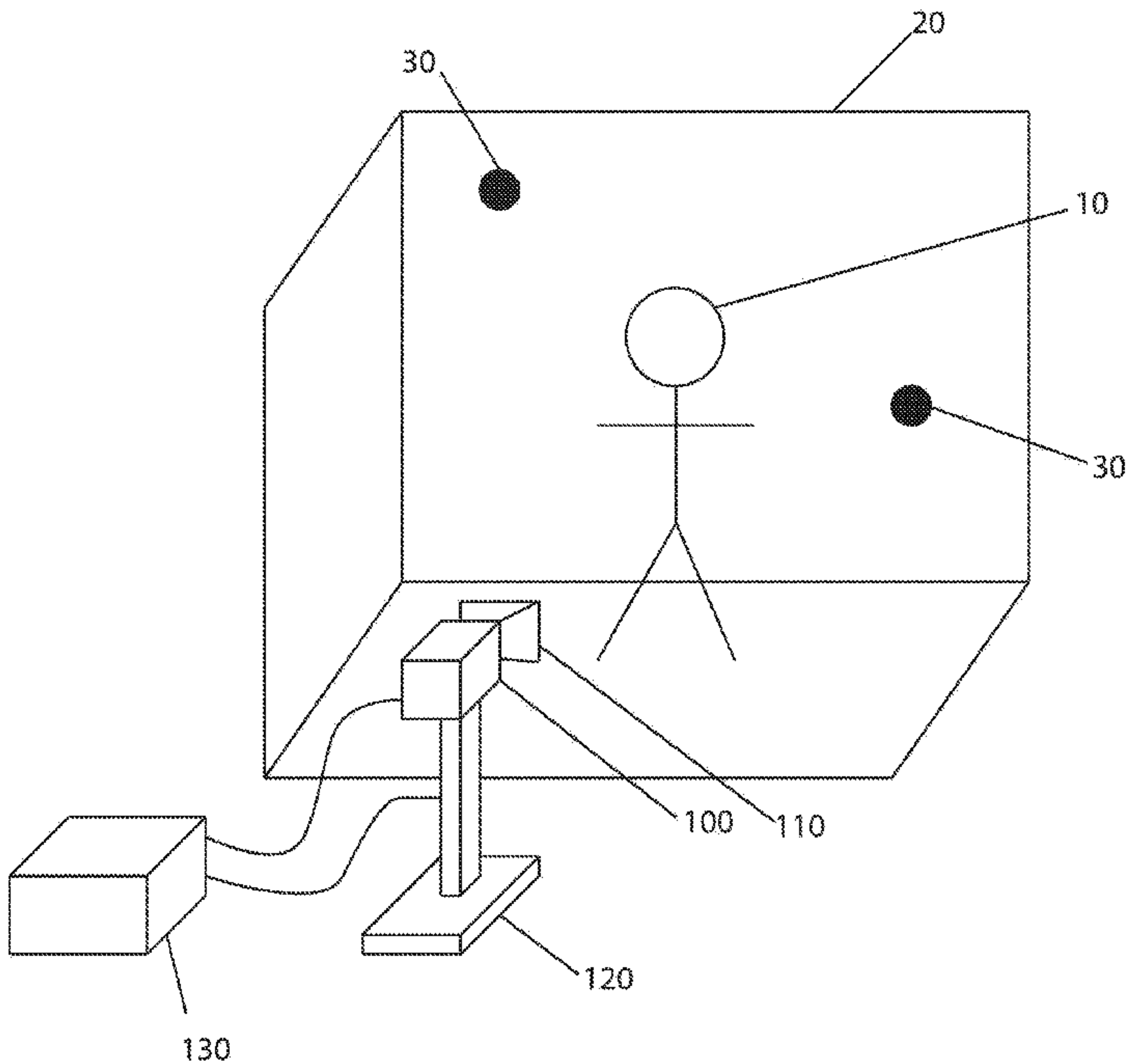


FIG. 1