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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 August 2018 (23.08.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/150336 A1

(54) Title: HIGH-SPEED OPTICAL TRACKING WITH COMPRESSION AND/OR CMOS WINDOWING
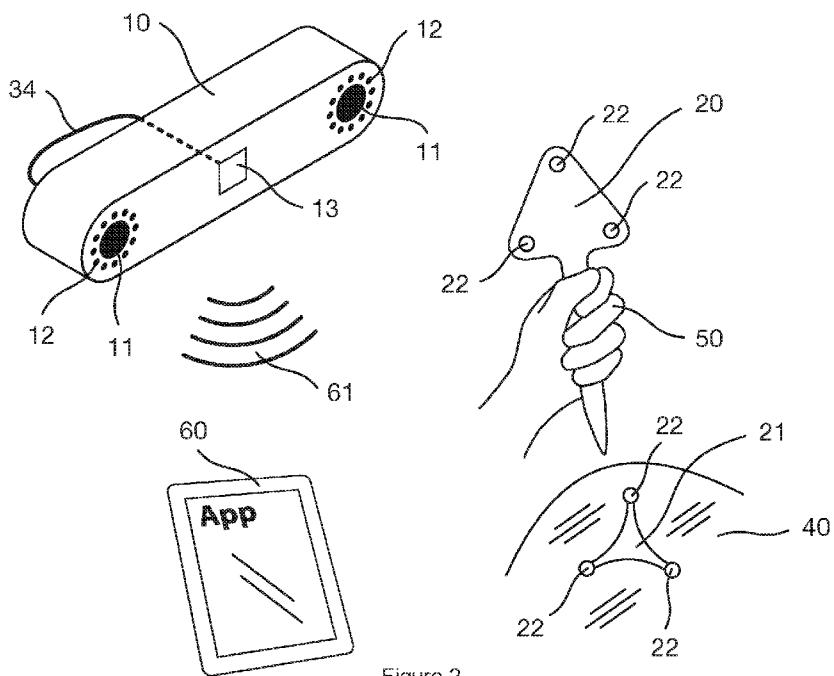
Figure 2

(57) Abstract: A high-speed optical tracking with compression and CMOS windowing for tracking applications that require high-speed and/or low latency when delivering pose (orientation and position) data. The high-speed optical tracking with compression and CMOS windowing generally includes an optical tracking system (10) that sense fiducial (22) and by extension markers (20). The markers can either be integrated on tools (20) handled by an operator (e.g. surgeon) (50) or fixed on the subject to track (e.g. patient) (40). The low-level processing (101, 107, 140) is realized within the tracking system to reduce the overall data to be transferred and considerably reduce the processing efforts. Pose data (127) are finally computed and transferred to a PC/monitor (30,31) or to tablet (60) to be used by an end-user application (130).

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

Declarations under Rule 4.17:
— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

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

HIGH-SPEED OPTICAL TRACKING WITH COMPRESSION AND/OR CMOS WINDOWING

CORRESPONDING APPLICATION

5   The present application claims priority to the earlier US provisional application USSN 62/458,697 filed on February 14, 2017, the content of this earlier application being incorporated by reference in its entirety in the present application.

TECHNICAL FIELD AND BACKGROUND OF THE INVENTION

10   The present invention relates generally to hi-speed optical tracking and more specifically it relates to a high-speed optical tracking with compression and CMOS windowing for tracking applications that require high-speed and/or low latency when delivering pose (orientation and position) data.

        Typically, such optical tracking systems are used in the field of computer assisted surgery

15   or dentistry, robotic assisted surgery or dentistry, rehabilitation and other similar medical applications. On a more general level, they are used in the fields of industrial part checking, or other applications where the tracking is required.

        Traditional optical pose tracking systems comprise two or more cameras. They use triangulation to determine the 3D position of light generating elements in space. These light

20   generating elements may either be active: they transmit light (e.g. LEDs) or passive: they reflect light (e.g. reflective disks or spheres), or a combination of active and passive. In the passive case, the light used to be generated on rings of LEDs located around the cameras of the optical tracking system. Emitted light is further reflected on tracked objects and sensed by the cameras.

        Such tracking systems are able to compute (that is determine) the 3D position of every

25   single light generating element. If at least 3 light generating elements with known relative positions are measured, it is further possible to compute (that is determine) a 3D orientation. Such an object is called a marker and markers are fixed on target objects/subjects to be tracked by the system.

30   In a medical application context, a user (e.g. a surgeon) touches a region of interest on the patient's body using a distal tip of an object (e.g. a probe or a surgical instrument). A tracking

system views the marker(s) affixed to the object and is able to retrieve its/their pose(s). Then, on the basis of the known relationship between the location of the marker(s) and the location of the object tip, the marker-tracking system determines the coordinates of the object's tip as well as it's orientation. Such pieces of information can be used for example by an application to register the patient with a preoperative planning and/or to steer the interventional radiologist when inserting a biopsy needle.

Moreover, medical and especially surgical applications require the markers to be sterile. Common passive markers are usually disposable, autoclavable or a mix of the two.

Examples of tracking systems and methods are given in publications WO 2017/208186 and US 2017/0086941.

BRIEF SUMMARY OF THE INVENTION

The invention generally relates to a hi-speed optical tracking which includes an optical tracking system (10) that sense fiducials (22) (by extension markers 20). The fiducials (respectively markers) can either be integrated on tools (20) handled by an operator (e.g. surgeon) (50) or fixed on the object/subject to track (e.g. patient) (40). The low-level processing (101, 107, 140) is preferably realized within the tracking system. Pose data (127) are computed and transferred to a PC/monitor (30,31) or to a tablet (60) to be used by an end-user application (130).

There has thus been outlined, rather broadly, some of the features of the invention in order that the detailed description thereof may be better understood, and in order that the present contribution to the art may be better appreciated. There are many additional features of the invention that will be described hereinafter.

In this respect, before explaining at least one embodiment of the invention in detail, it is to be understood that the invention is not limited in its application to the details of construction or to the arrangements of the components set forth in the following description or illustrated in the drawings. The invention is capable of other embodiments and of being practiced and carried out in various ways according to different embodiments. Also, it is to

be understood that the phraseology and terminology employed herein are for the purpose of the description and should not be regarded as limiting.

In a first aspect, an object and aim of the present invention is to improve the known devices and methods.

5

Another object is to provide a high-speed Optical Tracking with compression and/or CMOS windowing for tracking applications that require high-speed and/or low latency when delivering pose (orientation and position) data.

10   Another object is to provide a High-speed Optical Tracking with Compression and/or CMOS Windowing that improves the tracking speed and reduce the latency of pose data delivery.

Another object is to provide a High-speed Optical Tracking with Compression and/or

15   CMOS Windowing that enables high-speed / low latency within embedded processing as less data are to be processed with respect to a conventional solution (that would not have these features).

Another object is to provide a High-speed Optical Tracking with Compression and/or

20   CMOS Windowing that enables to transfer and process raw data within a compressed form.

Another object is to provide a High-speed Optical Tracking with Compression and/or CMOS Windowing that uses CMOS windowing and image-based object tracking to speed-up the readout phase of the imaging sensor.

25

Another object is to provide a High-speed Optical Tracking with Compression and/or CMOS Windowing that uses an additional micro-lens array in the front of the image sensor to enable light field applications (e.g. simultaneously acquiring 2D images and 3D depth maps).

30   Another object is to provide a High-speed Signal Processing with Compression and/or CMOS Windowing to speed-up the data acquisition with respect to a conventional solution.

Other objects and advantages of the present invention will become obvious to the reader and it is intended that these objects and advantages are within the scope of the present invention. To the accomplishment of the above and related objects, this invention may be embodied in the form illustrated in the accompanying drawings, attention being called to the fact, however, that the drawings are illustrative only, and that changes may be made in the specific construction illustrated and described within the scope of this application which is not limited to the exemplary embodiments described herein.

In an embodiment, the invention concerns a method for high speed tracking of optical fiducials and markers, wherein

-) at least one fiducial is attached on a marker which is placed on an object and/or on a person, said fiducial is detected by at least one optical sensor, said optical sensor providing sensor data;

-) the position and/or orientation of said marker(s) detected by said optical sensor is determined by a data processing step of said sensor data;

-) the determined position and/or orientation is transmitted through communication means as processed data to an application for further treatment,

wherein in the data processing step said sensor data is transferred to a compression module for compression of said sensor data and wherein said position and/or orientation of said marker is determined from said compressed data.

In an embodiment, the sensor data readout is partial, meaning that only a part of the sensor data is read and used for subsequent treatment, whereby said readout contains sufficient information for the determination of the fiducials to be tracked.

In an embodiment, the partial sensor data of said fiducials is not compressed for the determination of said position and/or orientation of said marker.

In embodiments of the invention, the data may be compressed or not and/or partial or not. Preferably, the acquired data is treated (for example compressed and/or partial, i.e. windowed) to reduce the overall size of data and reduce treatment time.

In an embodiment, in the data processing step of the method, the compressed or not-compressed data is segmented depending on the intensity of the pixels.

In an embodiment, after segmentation of the data, blobs are detected in the segmented image and their position determined.

In an embodiment, a geometry matching is made between a cloud of 3D reconstructed points and the geometry of the object.

In an embodiment, the further treatment of data comprises a navigation application for tracking said optical fiducials and/or markers.

In an embodiment, the data are used to control a robotic structure.

In an embodiment, the robotic structure is used in a surgical application.

In an embodiment, the robotic structure is held by an operator, such as a surgeon. For example, the robotic structure may be an intelligent tool such as the NAVIO™ surgical system manufactured by the Company Smith Nephew and disclosed for example at the following link http://www.smith-nephew.com/professional/microsites/navio/.

Other robot and device may of course use the present invention as well, this example being non-limiting.

The present invention also concerns a device for the high speed tracking of optical fiducials and markers according to one of the preceding claims, comprising
-) at least a fiducial attached to a marker placed on an object and/or a person,
-) an optical tracking system with an optical sensor detecting said fiducial(s) or said marker(s) via sensor data,

-) electronic means, comprising at least a memory and sensor data treatment means for determining the position and/or orientation of said marker(s) detected by said sensor by treatment of sensor data,

-) communication means to transfer the determined position and/or orientation of said marker(s)

5   to an application for further treatment,

wherein the data treatment means comprise at least

a compression module for compression of said sensor data and wherein said position and/or orientation of said marker is determined from said compressed data by said treatment means.

10   In an embodiment, said treatment means receive partial sensor data from a window of said sensor.

       In an embodiment, the device comprises an electronic device receiving said determined position and/or orientation, said electronic device being a computer device.

15

       In an embodiment, the computer device is a tablet or another equivalent device.

       In an embodiment, the object is marker.

20   In an embodiment, the transfer of data is made via wireless means. Of course, the transmission may also be made via wire or a combination of wire and wireless.

       In an embodiment, the optical sensor comprises a least one camera.

25   In an embodiment, the device is used to control a robotic structure.

       In an embodiment, the robotic structure is used in a surgical application.

       In an embodiment, the robotic structure is an intelligent tool and may be held by an

30   operator, such as a surgeon. The robotic structure may be the Navio^{TM} surgical system disclosed above or another equivalent structure/device.

In an embodiment, at least one optical sensor used in the method and the device may be a plenoptic camera.

In an embodiment, the plenoptic camera may comprise a micro-lens array placed in front of the sensor array, each micro-lens and the underlying pixels on the sensor array forming a micro-camera.

In an embodiment, the partial sensor data read in the method may be defined to retrieve only specific micro-camera pixels.

In an embodiment, the information from at least one micro-camera may be used to define a rough position of the fiducials in the images.

All technical means of the embodiments of the invention will be described in more detail in the following description of the invention and its features.

BRIEF DESCRIPTION OF THE DRAWINGS

Various other means, objects, features and attendant advantages of the present invention will become fully appreciated as the same becomes better understood when considered in conjunction with the accompanying drawings, in which like reference characters designate the same or similar parts throughout the several views, and wherein:

FIGURE 1: Figure 1 is an upper perspective view of a computer assisted surgery (CAS) setup for a typical registration step using optical tracking. The optical tracking system (10) is composed of two cameras (11), surrounded by infrared lights rings (11). The device is connected to a computer trough a cable (32) (e. g. USB, Ethernet, etc.). The optical tracking system (10) detects fiducials (22) (e. g. reflective material or infrared emitting diodes) that are in the field of view of the cameras. Fiducials that are rigidly affixed at a known position define a marker (21). A pointer (20) is a marker with a tool-tip that is used to point anatomical landmarks. The marker (21) is rigidly fixed (e. g. screwed, taped, etc.) on the patient (40). The pointer is hold by the

surgeon (50) which records anatomical landmarks on the patient (40). Corresponding landmarks are at the same time visible on the screen (31). The screen is connected to the PC via a cable (33). The correspondence between virtual and pointed landmarks is called registration and allows to know the position of the marker in the virtual reference. Registration is used in a CAS application to put in correspondence the planning and the real patient.

FIGURE 2: Figure 2 is an upper perspective view of a compact computer assisted surgery setup for a typical registration step using optical tracking. The process is similar to Figure 1 except that it is more embedded and compact. In this example, the optical tracking system (10) has embedded computing capacities (13). Instead on being connected to an external computer, the connection between the cameras (11) and the embedded computing unit is done internally (34), for example within the system (10). Note that the cable on the image is only for illustration. This link would practically and preferably be integrated in an electronic circuitry. After internal pose computation, data are transferred wirelessly (61) (e. g. through a WiFi, Bluetooth™ or LiFi connection) to a tablet PC (60). The end user application (APP) runs on the tablet.

FIGURE 3: Figure 3 is a flowchart of a conventional optical tracking system. Fiducials (22) are sensed by camera sensors (e. g. CMOS - 11). For each camera, pixels of the entire image or a sub-part of it are readout (105) to a memory (110). Images are then retrieved by the data processing unit (120) either via a wired link to an external PC (see Figure 1), or through an internal bus in case of an embedded processor (see Figure 2). The data processing unit first segments the image depending on the intensity of the pixels (121), detects blobs (122) in the segmented images, find their centroids, the right and left correspondence (123), and further determines their 3D position by mean of triangulation (124). The correspondence between the cloud of 3D points and the geometry of the markers is optionally performed. This step is called geometry matching (125). If a marker (21) comprises three or more fiducials (22), it is possible to compute its pose (position and orientation - 126). Pose data (127) are finally transferred to the navigation application (130) (e. g. The registration Software application described in Figure 1).

FIGURE 4: Figure 4 is a flowchart illustrating the overall operation of an embodiment of the present invention. Fiducials (22) are sensed (i.e. detected) by camera sensors (11). Camera

9

sensors are connected to a low-level tracking engine (101) that determine a window or several windows in the image provided by the sensor (11) where are located the fiducials (22), respectively the markers to track (20,21). Partial images are readout from the sensors (106) and further transferred to an image compression module (107) that will reduce the memory footprint (size) of the sub-images defined by the windows. This compression will also enable to later process the images without requiring to decompress them (140). Compressed images are further stored in the memory (110). The compressed data processing unit (140) retrieves the partial compressed images. It first segments the image depending on the intensity of the pixels (141) directly using the compressed images (without requiring any preliminary decompression), detects blobs (142) out of segmented image, defines their centroids, finds the right and left correspondence (143), and determines their 3D position by mean of triangulation (144). The correspondence between the cloud of 3D points and the geometry of the markers is optionally performed. This step is called geometry matching (145). Pose data (127) are finally transferred to the navigation application (130). As mentioned above, in one embodiment, partial images are readout and the corresponding data is compressed. In another embodiment, one may envisage to compress the data of the entire image (and not only of partial images) or alternatively to use directly the data of partial images without compression of data. Depending on the performance and capabilities of the system, this may be enough to ensure the desired tracking speed. However, if possible, the compression of partial data is a preferred method.

FIGURE 5: Figure 5 presents the timing diagram of a conventional sensor (up) versus a sensor using windowing (down). Note that for a sensor using windowing, the sensor integration time is the same but the readout is considerably reduced. In both situations, the readout time is directly proportional to the number of effective pixels to retrieve.

FIGURE 6: Figure 6 is an upper perspective view of a robotic assisted surgery setup using optical tracking. The process is similar to Figure 2 except that the human operator (50) is replaced by a robotic arm (51). The end-effector of the robot (52) holds a marker (20) and a surgical tool (53). The geometric relation between the marker and the surgical tool is known. Note that the transformation between the patient (via the marker) and the surgical tool is used to steer the robot during the operation process. It enables to control the robot's movements even if

the patient's anatomy is moving. Low latency and fast update rates of the optical pose estimation are required to safely and precisely control the robot. It also has an impact on the operating speed that could be increased.

5   **FIGURE 7:** Figure 7 is an upper perspective view of an intelligent tool (54) in a computer assisted surgery setup using optical tracking. An intelligent tool is defined as being a handheld robot. The intelligent tool is held by the surgeon and moves given the relation between the marker and the patient position. As such, it is similar to setup of figure 6, except that the base of the robot is not fixed and is held by the operator.

10
        FIGURE 8: illustrates the readout procedure of a CMOSIS CMV2000.

        FIGURE 9: illustrates a simulator in action. The white disks are the tracked blobs, the red areas (or clear lines) are the row-windows and the black (or dark) area is the invisible portion of the

15   frame.

        FIGURE 10: illustrates the HTML output of a simulator. The top graph shows the curve of the readout ratio during the capture. The user can query events to see screenshots using his mouse. The bottom graph is a histogram to assess the distribution of the readout ratio.

20

DETAILED DESCRIPTION OF THE INVENTION
A.   *Overview*

        Turning now descriptively to the drawings, in which similar reference characters
25   denote similar elements throughout the several views, the figures illustrate an optical tracking system (10) that sense fiducials (22). A marker (20) is formed by three or more fiducials fixed together at known locations. The markers can either be integrated within/on tools (20) held by an operator (e.g. a surgeon, or a robot) (50) or fixed on the object/subject to track (e.g. a patient) (40). The low-level processing (101, 107, 140) is preferably realized
30   within the tracking system. Pose data (127) are computed internally and sent to a tablet (60) or transferred and computed on an external PC (30). Poses are finally used by an end-user

application (130). The overall system is designed to improve the overall performances (e.g. improve speed, reduce latency) by reducing the quantity of data to transfer during the processing and thereby win time, for example.

5   *B.   Optical Tracking System*

The optical tracking system comprises one or more cameras (11). The device emits infrared light from rings (12), for example using IR LEDs surrounding each infrared camera. Cameras are sensing fiducials (22) placed on objects/subject to track like markers (20,21). There are mainly two types of fiducials, both being possible in embodiments of the present
10   invention, either as alternative solutions, or in combination:

1. Passive fiducials which are coated with a retro-reflective material containing small spheres that mirror the light (for example from the IR LEDs of the ring 12) back to the lenses, causing the fiducials to appear as bright spots in the infrared images.

2. Active fiducials which are light sources (e.g. infrared light emitting diodes - IR-
15   LEDSs). When using exclusively such fiducials, the rings (12) are then not necessary.

Raw data are further processed (101, 107, 140) to provide 3D position of the fiducials. When three or more detected fiducials are part of markers, it is possible to compute their poses.
20

Most commonly used tracking systems are stereo-cameras (system comprising two cameras) which are fixed on a baseline. The transformation from one camera to the other is usually factory calibrated or alternatively calibrated in the field. Once a fiducial is simultaneously detected by both cameras, it is possible to compute its 3D position by means
25   of triangulation. Example of commonly used passive optical tracking systems are for example Atracsys infiniTrack™ , fusionTrack™, spryTrack™, Northern Digital Polaris™, and Axios CamBar™.

Optical tracking systems usually operate in the near infrared spectrum but can
30   alternatively be designed for the visible spectrum or any other spectrum.

Optical tracking systems can alternatively detect tags or known pattern. Systems like ARToolkit™ (http://artoolkit.org) or MicronTracker™ from the company ClaroNav are tracking objects with specific tags. Tags/patterns are not to be considered as fiducials. Tags/patterns have feature points (see features points on Wikipedia). These feature points

5   then form fiducials, whereas tags/patterns are to be considered as markers in the context of the present invention. Other equivalent marking means (fiducials or markers) are of course possible in the scope of the present invention.

Optical tracking systems with three or more cameras are alternatively possible in

10   other embodiments of the invention.

Other types of setup are also possible like fixing single tracking cameras in the room and having an extra calibration step to define their respective transformation. This methodology is usually used for motion capture applications. It is alternatively possible to fix

15   several multiple camera tracking systems (with a known baseline) in the room.

Systems capturing active fiducials (e. g. Infrared LEDs) do not need to have the ring of LEDs (12) described on Figure 1 and Figure 2 but they may be present (but not used) in order to provide standard products

20

Optical tracking systems may have embedded processing capabilities (e. g. Figure 2) or delegated processing (e. g. Figure 1). By delegated processing, the pose computation from the raw signals is partly realized within the tracking system and partly on an external CPU like a PC. One configuration or the other may be chosen according to circumstances.

25

C.   **Tracked Tools**

A tracked tool comprises a marker rigidly fixed to (respectively integrated in) an instrument (20). The marker - also called rigid body in the literature - (20,21) comprises several fiducials (22) rigidly fixed altogether. The relative position of the fiducials forming

30   the marker is called the geometry and should be well known in order to estimate the pose of the marker from the 3D points of cloud. During the tracking process (140), these fiducials are

13

detected (142), identified (143) and their 3D position is calculated by the tracking system by means of triangulation (144). If at least three fiducials are detected, the Arun registration algorithm ("Least-squares fitting of two 3-D point sets ", Arun, 1987) can be used to calculate the pose (position + translation) of the marker in the tracking system referential.

5   Note that a tracked tool may comprise less than three fiducials. In this case, only reduced spatial information (less than 6 degrees of freedom) can be retrieved by the tracking system. Tracked tools are designed to be autoclaved or partly/entirely disposable in surgical applications.

10   The tracked tools may be a digitizer for registration purposes, or any surgical tools including drill, biopsy needle, or intra-operative imaging like ultrasound probe, endoscope, CT-scanner, MRI, robots, intelligent tools, etc.

The tracked tool may be operated by a human operator (e.g. the surgeon) or a robot
15   (see Figure 6) an intelligent tool (see Figure 7), in an automatic, semi-automatic or manual manner. An example of robot is the Mako System™ from the company Stryker. An example of an intelligent tool is the NAVIO™ from Smith and Nephew.

D.   PC System

20   The PC system alternatively called navigation PC (30) retrieves processed data (32), computes pose of the markers (20, 21) if this process has not already been done within the tracking system (13). The PC is connected (33) to a (touch-)screen (31) to display the end-user application (130).

25   Typical PC systems (including the screen) used in the Operating Room (OR) are placed on a cart and are fixed on a pole with a gimbal to aim the tracking zone.

The PC can run any operating system (windows, Linux, etc.).

30   The input device for the PC is any appropriate system such as a touch-screen, a mouse, a keyboard, a Kinect, etc.

14

E.   *Tracked Subject*

The tracked subject (40) is typically a patient under surgery. The marker (21) is fixed (as rigidly as possible) on the anatomy targeted by the surgery. For example, during a hip replacement procedure, the marker can be fixed on a Steinmann's nail.

Different surgical procedures use navigation like implant placement, image-guided surgery, neuro, ENT, spine, maxillofacial, liver surgeries, dentistry and may benefit from the present invention. Markers (21) are fixed (screwed, taped, etc.) on the anatomy to track. After registration step is realized, the marker should stay affixed to the anatomy, otherwise a re-registration is necessary.

F.   *Operator*

The operator (50) holds the tracked instruments. In the application of figure 1, the surgeon uses a digitizer (20) to point anatomical landmarks on the patient (40). In a surgical application, the operator is the surgeon, assistant and/or nurse. The operator may alternatively be a robot or a combination of both in case of an intelligent tool.

G.   *Tablet*

The tablet (60) is similar to a PC having a touch-screen except that data transfer (61) is usually performed wirelessly.

Tablet may be a portable PC (e.g. Lenovo Yoga ®), a phablet (Android®, iOS®, Windows ®) a tablet (iPad®) or any mobile computer including a touch screen.

The tablet is usually medical grade and/or wrapped in a sterile bag to avoid contamination within the sterile field of the OR.

H.   *Camera Sensor*

A camera sensor (11) comprises an image sensor and electronics to control it. For example, CMOS (Complementary Metal Oxide Semiconductor) image sensors (100) convert

15

light into electric charge and process it into electronic signals. With this technology, each pixel has its own charge-to-voltage conversion, and the sensor often also includes amplifiers, noise correction, and digitization circuits, so that chip outputs are digital bits. A CMOS imager converts charge to voltage at the pixel, and most functions are integrated into the

5   chip. CMOS technology has the ability to read out only a portion of the image (also called windowing mechanism) (106). This technique is used to increase the overall frame rates and reduce the latency as only small regions of interest (pixels) are readout and may typically be used in the present invention.

10   Sensors are linked to dedicated digital electronics to both control the chip (e.g. set the integration time, set the pixel depth retrieval, etc.), as well as estimate the position of the fiducials (22) within the images (101) to set the corresponding region of interest (windows) within the CMOS sensor. A smart windowing heuristic is able to reduce the readout time of the CMOS sensor – as less pixels are readout - by a factor typically ranging from 1. 5 to 3

15   enabling to considerably speed up the entire acquisition process. An example of heuristic is presented later in the present description and my be applied in embodiments of the present invention.

Any image-sensor technologies having windowing capabilities may alternatively be

20   used in embodiments of the present invention.

Partial images readout from the image sensor are further compressed (107). The image-compression algorithm is designed such that the image segmentation process (141) can be done without requiring (any / a full) decompression of the images. Resulting

25   compressed images are further transferred (108) to a shared memory (110).

The camera sensor may comprise a CMOS image sensor, connected to a micro-controller or FPGA control circuit. The electronic to control the CMOS may alternatively be embedded on the chip (ASIC). The camera sensor may be fully designed for this task

30   meaning that it will include the effective array of pixels, the windowing capabilities and the low-level tracking engine (101).

The camera lens may comprise a classic main lens or be composed with an additional micro-lens array allowing the image sensor to record 2D image and 3D depth information of a scene with a single capture (eg. https://raytrix.de/technology/, https://www.lytro.com). The

5   placement of an additional micro-lens array in front of the image sensor modifies the captured light on the sensor. Each single micro-lens creates a single image on the sensor, the size of that image being equal of the size of the pixels behind each single micro-lens (also known as a micro-camera). That captured sub-image represents the scene with a specific angle of view. The footprint of that image on the sensor is limited to the size of the micro-

10   lens. It means that in that case, with a partial read-out capturing pixels from a unique micro-lens, the complete 2D overview view of the scene is accessible. That type of lens can be mounted on a tracking system, on one or multiple cameras to help the low-level tracking engine to fine-tune the windowing algorithm. The windowing algorithm could use that partial read-out to visualize the complete field of view even with a lower resolution. Noisy field of

15   view can be easily identified with a predefined limited read-out time.

Such a light-field (aka plenoptic) camera allows among other things the refocusing of the depth of field on post-capture. The windowing algorithm can reconfigure on the fly the focus of the scene for each captured partial data from a unique single capture. Different

20   criteria such as optical jitter of a blob or its size, its relative distance or the angle of view of the seen 3D marker can be considered.

That lens may be mounted on IR limited or visible spectrum ranged image sensor. Only 2D image partial or complete data can be processed without analysis of 3D depth

25   information. 3D depth information can be partially reconstructed, even with a limited resolution.

That light-field camera may alternatively be used as an additional data source (2D and/or 3D image) synchronized or not with other cameras.

30

17

Any other light-field technology (e.g. multi-camera arrays) may alternatively be used.

I.   *Memory*

5   The memory (110) stores the compressed partial images coming from the camera sensors modules (100). If the processing is embedded, the main processor will access (via an internal bus) these partial images for further processing. If the processing is done on an external PC, partial image data are transferred to the PC by a communication channel (32). This communication channel may be wired (e. g. USB, Ethernet, etc.) or wireless (e.g. Li-Fi, 10   Wi-Fi, Bluetooth, etc.).

A fast memory is required to store the real-time flow of data. A RAM (e.g. DDR2, DDR3) is typically used for example.

15   *J.   End-user Application*

Typical end-user application is dedicated for image-guided surgery, computer-assisted surgery, or robotic-assisted surgery.

Alternate applications for motion capture or rehabilitation may benefit from the 20   proposed invention.

Alternate industrial application in the fields of quality control and reverse engineering may also be envisaged.

25   Other general purpose metrological application may benefit from the invention.

K.   *Compressed Partial Images Processing*

The goal of the compressed partial images processing (140) is to compute the pose estimation of the marker (146) from the raw compressed partial images with the minimum 30   CPU resources and as fast as possible. The overall process is well known in the literature and can for example be found in the paper: "Affordable Infrared-Optical Pose-Tracking for

Virtual and Augmented Reality", Tomas Pintaric, 2007, which is incorporated by reference in the present application. The originality of the proposed approach is to perform the processing without requiring to preliminary decompress the images data. It has the advantages to require less bandwidth to transfer and to process the data which will increase

5   the update speed and reduce the overall latency.

The processing can basically be defined as followed: the compressed partial images are first segmented (141) to find the blobs. A blob is normally the image of a fiducial (but can alternatively be some noise). An example of such a compression algorithm is described

10   later in the present description. The centroid of each candidate blob (142) is then computed using standard image processing techniques. The blobs are further identified (143) in order to determine which fiducial/marker they are from. As the geometry of the camera sensors (intrinsic and extrinsic parameters) (100) is known, it is possible to find the 3D positions of each fiducials by mean of triangulation (144). The system further tries to find the geometries

15   of the markers within these reconstructed 3D points (145). Finally, the pose (orientation and position) of the markers and/or the 3D position of the individual fiducials is computed (146) and further transferred (127) to the end-user application (130).

This processing is done in an embedded processor like an FPGA (Field

20   Programmable Gate Array), a SoC (System on Chip), a micro-controller, a CPU, a DSP or directly within a dedicated ASIC chip.

Captured partial image data moved into memory can leverage multiple optimization offered by the platform. A typical optimization could be to trigger an interrupt signal on each

25   partial sets of data moved into memory. A set of data could be an aggregate of pixels to reconstruct one single blob.

The processing unit could be a SoC FPGA with separated hardware FPGA circuit (configured with predefined operations) and a hardware processor such as an ARM based

30   HPS, executing software algorithms such as image blob analysis. The FPGA unit will trigger a signal as soon as that partial data is ready to be process and pushed in memory. The

hardware/software processor could then start its software computation before waiting other sets of data including other blob positions. The overall process can also be parallelized.

5   *L.   Connections of Main Elements and Sub-Elements of Invention*

In a typical optical tracking setup, each camera (11) of the optical tracking system (10) is sensing fiducials that are in the field of view. Camera sensors - typically CMOS - are connected to a low-level tracking engine (101) that both detects the location of new fiducials in the images as well as tracks the previously detected fiducials. This action is realized by

10   estimating the position of several windows in the image where the fiducials are expected to be located. The chapter "Example of heuristic for marker tracking in the camera sensor images" detailed hereunder in the description presents a possible implementation of such an algorithm. Partial images are readout from the sensors (106) and further transferred to an image compression module (107) that will reduce the memory footprint (size) of the image.

15   This compression will also enable to later process the image without requiring a partial/full decompression (140). The chapter: "Example of algorithm to retrieve fiducials centroid in compressed partial images" hereunder presents a possible implementation of such an algorithm. Compressed partial images are further transferred to the main memory (110). The processing module is informed that new data are available either by an internal signaling

20   system in case of an embedded processing (13) or by transmission means (USB, Ethernet, Wi-Fi, etc.) in case the processing is realized on an external PC (30). The compressed data processing unit (140) first segments the image depending on the intensity of the pixels (141) directly using the compressed images as input - without decompressing the data -, detects blobs (142) out of segmented image, defines their centroids, finds the right and left

25   correspondence (143), and determines their 3D position by mean of triangulation (144). The correspondence between the cloud of 3D points and the geometry of the markers is further performed if the fiducials are affixed to markers (20,21). This step is called geometry matching (145). Pose data are computed (146) and are further transferred (127) to the end-user application (130). The end-user application will format the measurements according to

30   the interface. For example, it will inform the surgeon on the trajectory of a biopsy needle, on the position of a digitizer (20) with respect to the patient (40), etc.

M.   *Alternative Embodiments of Invention*

If the invention comprises only the windowing mechanism, it enables the readout time to be reduced. It has the consequences to diminish the overall latency and/or to increase the update rate. As presented later, a typical configuration without windowing will run at 32 Hz due to the communication time to transfer the images. From several experimental setups, adding the windowing mechanism will increase the frame rate by a factor comprises between 30% to 100% (41. 6 Hz to 64 Hz) depending on the fiducials configuration.

If the invention comprises only the compression mechanism, it enables to transfer the images at a higher frame rate (typically 335 Hz with 4-ms latency with an Atracsys fusionTrack^{TM} system) compared to the 32 Hz without compression. High frame rate enables new types of application like integrating the optical tracking in the control loop of a robotic system.

Combining the windowing and compression mechanism (also called the overdrive mode) enables to get the performances of a high-end system (typically 1000 Hz with 3ms latency for an Atracsys fusionTrack^{TM}) with mid-range components and with only few hardware modifications of a conventional tracking system. This allow an overall cost reduction of the hardware.

In an embodiment of the invention, the overdrive mode could be used to add tracking features on conventional monocular or stereo video camera. By conventional, one mean grayscale or color cameras acquiring full images and working in the visible and/or near IR spectrum. These cameras could have illumination rings around their lenses or not. In this embodiment, tracking images are interleaved between standard camera images. In a single camera configuration, an algorithm like "Model-based object pose in 25 lines of code », Dementhon, 1995 could be used to estimate the pose of an object based on 4 or more fiducials which are located on the object/subject to track. By using the overdrive mode tracking will require fewer processing in the camera electronics and a low transmission bandwidth. As such, the overall frame rate of this hybrid tracking/conventional camera will

be almost the frame rate of a conventional one. In the other hand, if the dual tracking and normal acquisition are integrated in a conventional camera, the global update rate would be two times less (as the entire tracking image needs to be transferred without compression).

5   An example of application of the previous embodiment is head tracking with a single (or alternatively a stereo) camera. The user wears glasses that have multiple fiducials attached. A camera with the overdrive mode enables to track head movements (respectively eyes positions) at fast update rate / low latency. A hybrid camera could be used to track and acquire images simultaneously. Hybrid acquisition could be used to operate simultaneously 10   tracking of the glasses and a video acquisition (e.g. video conference). Estimation of eyes position could be used to drive an autostereoscopic 3D screen.

A similar hybrid approach could be used to track the head of a user with one or more fiducials located on the head (e.g. on glasses, hat, etc.) and simultaneously use image 15   processing on conventional images to detect other features like eye, nose, mouth. These new detected features enable to refine the initial pose estimation of the tracking. Alternatively tracking of fiducials enable to bootstrap the feature detection in the images and/or make it faster.

20   In an embodiment of the invention, instead of having rings to illuminate reflective fiducials, active LEDs could be wirelessly synchronized (optical synchronization, Wi-Fi, Bluetooth, Li-Fi, etc.) to the camera(s) acquisition so that LEDs are active only when the camera is sensing photons. This mechanism enables to have a hybrid camera (tracking/image acquisition) as previously described by just using an existing camera hardware. Such cameras 25   could be webcams, industrial cameras, security cameras, etc.

In an embodiment of the invention, the hybrid camera is used for security applications. Tracking is performed to check if objects (equipped with one or more fiducials) have moved or to estimate specific parameters within the field of view (e.g. the position of 30   the floor, door, restricted area, etc.).

In an embodiment of the invention, instead of tracking the active/passive fiducials previously described, the system is extracting and tracking SIFT (Scale-Invariant Feature Transform) features at high frame rate on partial image using the windowing mechanism. The hybrid mode would enable to acquire complete images interleaved with partial images.

5   Both rates could be independently configured (synchronously or asynchronously). Any other image feature could alternatively be used (e.g. contour/edge descriptor, color/texture descriptor, etc.).

In an embodiment of the invention, instead of tracking fiducials, the system may be

10   capturing high-speed general purpose 1D signals (e. g. spectroscopy). Using a 2D sensor (as presented in this invention) instead of 1D one may be interesting either to be able to parallelize the processing (e.g. having multiple 1D signals on the 2D sensor) and/or to be able to correct the perfect alignment of the signal on the image sensor which is not possible with a 1D sensor. The overdrive mechanism enables to increase the overall update rate (data

15   rate) and to reduce the latency.

In an embodiment of the invention, the windowing heuristic could input information provided by external sensors like Inertial Measurement Units (IMUs). IMUs have the ability to work at higher frame rate or with an interleaved data acquisition. As such, it is possible to

20   extrapolate the position of a fiducial in the camera reference frame in order to narrow the windowing search. Such an approach will reduce the search zones (windows), make the detection heuristic more robust and enable a higher frame rate / reduced latency.

In an embodiment of the invention, the described optical tracking system can be used

25   in a medical robotic application. Fiducials (respectively markers) would be placed on the robot (e.g. on its end-effector) and on the patient. Robot control would include the data from the high-speed, low latency optical tracking system, for example, to compensate in real-time for the movements of the patient during the procedure.

30   In an embodiment of the invention, the described optical tracking system can be used in an intelligent tool robotic application. This configuration is very similar to the previous

one except that the surgeon will hold the robot. In addition to the previous benefits, this setup enables to have the surgeon move faster (and more naturally) as the latency is reduced and update speed increased.

5   What has been described and illustrated herein is a preferred embodiment of the invention along with some of its variations. The terms, descriptions and figures used herein are set forth by way of illustration only and are not meant as limitations. Those skilled in the art will recognize that many variations are possible within the spirit and scope of the invention in which all terms are meant in their broadest, reasonable sense unless otherwise

10   indicated. Any headings utilized within the description are for convenience only and have no legal or limiting effect.

         Exemplary embodiments have been described to provide an overall understanding of the principles of the structure, function, manufacture, and use of the systems and methods

15   disclosed herein. One or more examples of these embodiments are illustrated in the accompanying drawings. Those skilled in the art will understand that the systems and methods specifically described herein and illustrated in the accompanying drawings are non-limiting exemplary embodiments and that the scope of the present invention is defined not solely by the claims. The features illustrated or described in connection with one exemplary

20   embodiment may be combined with the features of other embodiments. Such modifications and variations are intended to be included within the scope of the present invention. A number of problems with conventional methods and systems are noted herein and the methods and systems disclosed herein may address one or more of these problems. By describing these problems, no admission as to their knowledge in the art is intended. A

25   person having ordinary skill in the art will appreciate tha, the scope of the present invention is not so limited. Moreover, while this invention has been described in conjunction with a number of embodiments, it is evident that many alternatives, modifications and variations would be or are apparent to those of ordinary skill in the applicable arts. Accordingly, it is intended to embrace all such alternatives, modifications, equivalents and variations that are

30   within the scope of this invention.

INDEX OF ELEMENTS

10: **Optical Tracking System**
11: Camera
12: Ring of Ir Leds
13: Internal Cpu Ressources

20: **Tracked Tools**
21: Marker
22: Fiducials

30: **Pc System**
31: (touch-)screen
32: Wired Connection To Pc
33: Connection Between Screen And Pc
34: Internal Connection To Embedded Cpu

40: **Tracked Subject**

50: **Operator**
51: Robot
52: End Effector of the Robot
53: Surgical Tool
54: Intelligent (Robotic) Tool

60: **Tablet**
61: Wireless Connection To Optical Tracking System

70: **N/A**

80: **N/A**

90: **N/A**

100: **Camera Sensor**
105: Pixel Readout To Memory
106: Partial Images Readout To Image Compressor
107: Image Compressor
108: Transfer of Compressed Images To Memory

110: **Memory**

120: **Data Processing**
121: Image Segmentation Module
122: Blob Detection Module
123: Blob Identification Module
124: Triangulation Module

125: Geometry Matching Module
126: Pose Computation Module
127: Transfer of Pose Data To Application

130: End-user Application

140: Compressed Partial Images Processing
141: Image Segmentation Module With Partial Compressed Images
142: Blob Detection Module
143: Blob Identification Module
144: Triangulation Module
145: Geometry Matching Module
146: Pose Computation Module
147: Transfer of Pose Data To Application

Annex 1: Example of heuristic for marker tracking in the camera sensor images.

The use of robotics in surgical application is on the rise. These new applications are by design more latency-constrained than traditional applications of computer assisted surgery systems. Indeed, smaller latencies are necessary for robots to react to external events on time. This is especially critical in surgical applications where human life is at stake. The high-end tracking solution of the company Atracsys is called the fusionTrack™. It is able to stream and process stereoscopic images at more than 330 Hz. While its frame rate is fairly high, customers pointed out that its average latency of 4 milliseconds doesn't meet their real-time requirements. To reduce it to 3 milliseconds, we are going to:

1. roughly break down the contributions to the latency of the fusionTrack™ to show that the bottleneck is the readout of the CMOS sensor (section 1),
2. devise a heuristical windowing predictor to track blobs at a higher frame rate (section 2).

A1. Latency break down of the fusionTrack™

We define the latency of a stereoscopic tracking system as the time elapsed between the images' exposure and their delivery, after processing, to the user application. On the fusionTrack™, this time can roughly be divided in the exposure time, the readout time, the compression time, the memory access time, the transmission time and the processing time. Some of these contributors are pipelined: for instance, images are compressed and stored in memory by the hardware as they are read out from the sensor. However, we have to pay the full price of the sensor readout, transmission and processing time.

The CMOSIS CMV2000, the CMOS sensor used in the fusionTrack™, already attempts to amortize the readout cost by providing 16 readout channels to read 2048 x 1088 x 10-bit frames. More precisely, 10-bit pixels are burst out sequentially in batches of 128 adjacent columns from each channel. Figure 8 outlines this readout process.

The readout clock being a 240-Mhz DDR clock, the time to read out a single pixel is therefore:

$$10 \text{ bits/pixel} \times 2.083 \text{ ns/bit} = 20.8 \text{ ns.}$$

So, the time to read a full row is a full row is

$$10 \text{ bits/pixel} \times 128 \text{ pixels/burst} \times 2.083 \text{ ns/bit} = 2.7 \text{ µs.}$$

The total time to read a full frame follows:

$$1088 \text{ rows/frame} \times (20.83 + 2667) \text{ ns/row} = 2.9 \text{ ms.}$$

In conclusion, the readout time accounts for roughly 75% of our 4-ms latency. By Amdahl's law, breaking down the rest of the contributions is unnecessary as we would need to entirely annihilate it to reach our 3-ms latency goal.

27

A2. Height rows to rule them all

The CMOSIS CMV2000 features row-windowing capabilities. More precisely, 8 ranges of contiguous rows can be read out from the sensor instead of the full picture. Our idea is to limit the section of the frame being read out by predicting the row-windows of interests, i.e. those containing tracked blobs, of the next frame based on the current frame. This work is based on the following assumptions:

1. reading one nth of the frame will roughly divide the readout time by n, meaning that
   a. the time to write and for the sensor to adopt the new row-windowing configuration is negligible, and
   b. the next frame row-windowing can be computed during the readout of the current frame.

Each of these assumptions are to be discussed in the following sections.

A2.1. Configuration overhead

The row-windowing configuration is stored in the sensor's registers programmed over SPI. The SPI clock has 48-MHz frequency and up to 34 8-bit registers need to be programmed (2 specifying the overall 16-bit number of lines being read out and 4 per windows). Writing an 8-bit register takes requires 16 bits to be written (1 R/W bit, 7 address bits and 8 data bits). As multiple SPI writes can be burst, the time to update the row-windowing configuration can be estimated as

$$34 \text{ registers} \times 16 \text{ bits/register} \times 20.833 \text{ ns/bit} = 11.3 \text{ μs.}$$

Being more than 270 times smaller than the current readout time, this appears to validate assumption 1a.

A2.2. A row-windowing heuristic for optical tracking

In this section, we devise a heuristic for row-windowed optical tracking. As it is to be implemented in programmable hardware, the heuristic must be simple enough to limit implementation cost.

As a starting point, simplistic approach would be to evenly divide the frame into 8 row- windows and activate them on frame i+1 based on the content of frame i. In other words, for each of the 8 evenly-divided row-windows, we activate it on frame i + 1 if the one on frame i contains a tracked blob. While this approach works, it has numerous downsides, the first of which being its lack of effectiveness if tracked blobs are smalls compared the height of a uniform row-windows. To solve this problem, we might look at the content of a given row-window on frame i and update its boundaries so that it is clamped to its content on frame i + 1. While this will improve the heuristic's performance, it greatly diminishes its reliability as moving blobs will quickly be lost. Therefore, the boundary clamping must be performed so as to allocate a safety zone on top and bottom of the content of a given row-window. Determining the height of this safety zone is

28

dependent on the user application. More precisely, it will depend on the speed of the blobs and the optical jitter of the sensor.

Finally, another issue is that untracked blobs might show up in invisible area of the frame as the predictor doesn't have access to the whole frame. A first solution might be to ignore this problem and make sure that new blobs are never introduced for the duration of the experiment. This scenario results in an overhead latency only in the initialization phase until all the blobs are tracked. However, in the undesirable event of the loss of a blob whether it is because of its speed or because it has gone out of the working volume, this would force the user to go back to the initialization phase. In conclusion, choosing whether to solve the problem of invisible blobs should be left to the user. Therefore, we considered two approaches:

1. at given intervals, a full frame can be captured, or
2. a sliding row-window can scan the frame continuously.

The first approach generates spikes in the readout period. It means that, for the duration of a frame, the system might rollback to its original 4-ms latency. This is unacceptable as, if it coincides with an event of interest for the robot controller, this would result in non-conforming feedback latency. We therefore restricted our research to the second approach.

Finally, note that while this potentially validates assumption 1b, we still need to show that this approach works in practice and come up with a design that parallelize the frame readout and the prediction detailed above.

A2.3. Results

To assess the impact of the row-windowing heuristic described in section A2.2, we have implemented a simulator. It tries to predict the row-windowing configuration of frame $i + 1$ based on the con- tent of frame i. For a qualitative assessment, it graphically displays the row-windows and the tracked blobs in real time, and saves screenshots of events that needs further attention (change in the readout ratio, i.e. the portion of the frame being read out, or invisible blobs). Finally, it outputs an automated HTML report that shows the readout ratio curve, histogram and the blobs that are lost. Figure 9 shows a screenshot of the simulator, namely the simulator in action, where the white disks are the tracked blobs, the red/clear areas or lines are the row-windows and the black/dark areas are the invisible portion of the frame. Figure 10 shows an example of a generated report.

In conclusion, we showed that our heuristic for row-windowed optical tracking is valuable provided the environment is controlled. To overcome the environmental risk, a readout time monitoring mechanism must be added to the acquisition pipeline in addition of the row-windowing predictor. This would allow the end-user to detect if it is basing a decision on outdated data depending on the application-specific time constraints.

Figure 10 illustrates an HTML output of the simulator. The top graph shows the curve of the readout ratio during the capture. The user can query events to see screenshots using his mouse. The bottom graph is a histogram to assess the distribution of the readout ratio.

Annex 2: Example of algorithm to retrieve fiducials centroid in compressed partial images

In delegated processing (as presented in Figure 1), image data are entirely processed on the computer. This is not a limitation with modern processor as the still following the Moor's law. The only drawback is the transmission of the data from the tracking system to the computer. If we consider a Gigabit Ethernet connection -- which is now available on any PC -- and a system composed of two Full-HD grayscale cameras with a pixel depth of 8 bits, we get a maximum theoretical update rate of:

$$1024\text{^}3 / (8 * 2 * 1920 * 1080) = 32 \text{ Hz}$$

A good way to improve the update rate would be to compress the images in the tracking systems and decompress them in the computer. This approach would unfortunately require much more data processing means on the computer as images must be first decompressed before being further processed.

Modern optical tracking requires precision and speed. Precision is achieved by using cameras with high resolution and fast update rate is necessary to reduce the latency between a real movement and its corresponding measurement in the application.

As an example, the speed of a robotic system driven by optical tracking is directly linked to the update rate and latency of the tracking system.

We propose to solve the drawback of the decompression overhead on the computer by providing a compression representation that can directly be used by the processing unit (e.g. PC) without requiring any decompression.

- A first advantage of this approach is that the update rate is limited by the compression efficiency (as the transmission is the weakest link in the processing chain).
- Another advantage of our invention is as the compressed image can be more efficiently stored in the cache of the processor, it will be process faster.

Practically, with our compression we have developed, we get a maximum update rate improved by a factor 10 in a typical configuration (see Figure A1.2). This compression algorithm is only able to compress full 8bits images. It is presented below:

The compression algorithm produces the following output:

- the output data consists of 8 bits, the MSB discriminates between white and black pixels:
  - o 1 means the 7 LSBs are the original pixel intensity divided by 2;
  - o means the 7 LSBs are the number of consecutive black pixels. The following exceptions exist:
    - ▪ 0x80 means 128 consecutive black values;
    - ▪ 0x00 followed by a value encoded on 2 bytes (little endian, i.e. LSBs first) indicates a number of completely black lines. The address of this 0x00

byte must be a multiple of 16;
- any other 0x00 are ignored (this is a no-op character, used by the FPGA which must write multiples of 16 bytes).

The advantage of such a compression is that it is easy to get the (x,y) position of a blob directly in the compressed image (without need for decompression).

This representation enables to efficiently extract blobs from compressed images by using an adapted version of a line-scan fill algorithm. The scanline fill algorithm (https://en.wikipedia.org/wiki/Flood_fill#Scanline_fill) is an optimization of the flood fill algorithm. Instead of pushing each potential future pixel coordinate on the stack, it inspects the neighbor lines (previous and next) to find adjacent segments that may be filled in a future pass; the coordinates (either the start or the end) of the line segment are pushed on the stack. In most cases this scanline algorithm is at least an order of magnitude faster than the per-pixel one.

We are able to run the described configuration at a speed of 335 Hz which is more than 10 times higher than the non-compressed version.

CLAIMS

1.   A method for high speed tracking of optical fiducials and markers, wherein

-) at least one fiducial (22) attached to a marker (20) placed on an object (20) and/or on a person (40) is detected by at least one optical sensor (11), said optical sensor (11) providing sensor data;

-) the position and/or orientation of said marker(s) (21) detected by said optical sensor (11) is determined by a data processing step of said sensor data;

-) the determined position and/or orientation is transmitted through communication means as processed data to an application for further treatment,

wherein in the data processing step

said sensor data is transferred to a compression module for compression of said sensor data and wherein said position and/or orientation of said marker is determined from said compressed data.

2.   The method as defined in claim 1, wherein sensor data readout is partial, whereby said readout contains sufficient information for the determination of the fiducials to be tracked.

3.   The method as defined in one of the preceding claim, wherein said partial sensor data of said fiducials is not compressed for the determination of said position and/or orientation of said marker.

4.   The method as defined in one of the preceding claims, wherein in the data processing step, the compressed data is segmented depending on the intensity of the pixels.

5.   The method as defined in the preceding claim, wherein after segmentation, blobs are detected in the segmented image and their position determined.

6.   The method as defined in the preceding claim, wherein a geometry matching is made between a cloud of 3D reconstructed points and the geometry of the object.

7.   The method as defined in the preceding claim, wherein the further treatment of data comprises a navigation application for tracking said optical markers.

8.   The method as defined in one of the preceding claims, wherein said raw and/or processed data are used to control a robotic structure.

9.   The method as defined in the preceding claim, wherein the robotic structure is used in a surgical application.

10.   The method as defined in claim 8, wherein said robotic structure is held by the operator.

11.   The method as defined in one of claims 1 to 10, wherein said at least one optical sensor is a plenoptic camera.

12.   The method as defined in the preceding claim, wherein said plenoptic camera comprises a micro-lens array placed in front of the sensor array, each micro-lens and the underlying pixels on the sensor array forming a micro-camera.

13.   The method as defined in the preceding claim, wherein the said partial sensor data are defined to retrieve only specific micro-camera pixels.

14.   The method according to one of claims 11 to 13, where information from at least one micro-camera is used the define the rough position of the fiducials in the images.

15.   A device for the high-speed tracking of optical fiducials and markers, comprising
-) at least a fiducial (22) attached to a marker (21) placed on an object (20) and/or a person (40),
-) an optical tracking system (10) with an optical sensor (11) detecting said fiducial(s) (22) or said marker(s) (21) via sensor data,
-) electronic means, comprising at least a memory (110) and sensor data treatment means (140-146) for determining the position and/or orientation of said marker(s) (21) detected by said sensor (11) by treatment of sensor data,
-) communication means to transfer the determined position and/or orientation of said marker(s) to an application (130) for further treatment,

33

wherein the data treatment means comprise at least

a compression module (107) for compression of said sensor data and wherein said position

and/or orientation of said marker is determined from said compressed data by said treatment

means.

16.   The device according to claim 15, wherein said treatment means receive partial sensor

data from a window of said sensor.

17.   The device according to claim 15 or 16, wherein it comprises a electronic device

receiving said determined position and/or orientation, said electronic device being a computer

device.

18.   The device according to claim 17, wherein said computer device is a tablet (60).

19.   The device according to one of claims 15 to 18, wherein said object is marker (20).

20.   The device according to one of claims 15 to 19, wherein the transfer of data is made via

wireless means.

21.   The device according to one of claims 15 to 20, wherein said optical sensor comprises a

least one camera (11).

22.   The device according to one of claims 15 to 21, wherein said device is used to control a

robotic structure.

23.   The device as defined in the preceding claim, wherein the robotic structure is used in a

surgical application.

24.   The device as defined in one of claims 15 to 23, wherein said robotic structure is held by

the operator.

25. The device as defined in one of claims 15 to 24, wherein at least one optical sensor (11) is a plenoptic camera.

26. The device as defined in the preceding claim wherein the plenoptic camera comprises a micro-lens array placed in front of the sensor array, each micro-lens and the underlying pixels on the sensor array forming a micro-camera.

27. The device as defined in the preceding claim, wherein the partial sensor data are defined to retrieve only specific micro-camera pixels.

28. The device according to one of claims 25 to 27, where information from at least one micro-camera is used the define the rough position of the fiducials in the images.

Figure 1

Figure 2

CMOS sensor #1	CMOS sensor #2

100

105 — full image pixels readout — 105

110 — Memory

data transfer (USB, ethernet or internal bus)

120 — Data processing

121 — - image segmentation
122 — - blob detection
123 — - blob identification
124 — - triangulation
125 — - geometry matching
126 — - pose computation

127 — pose data
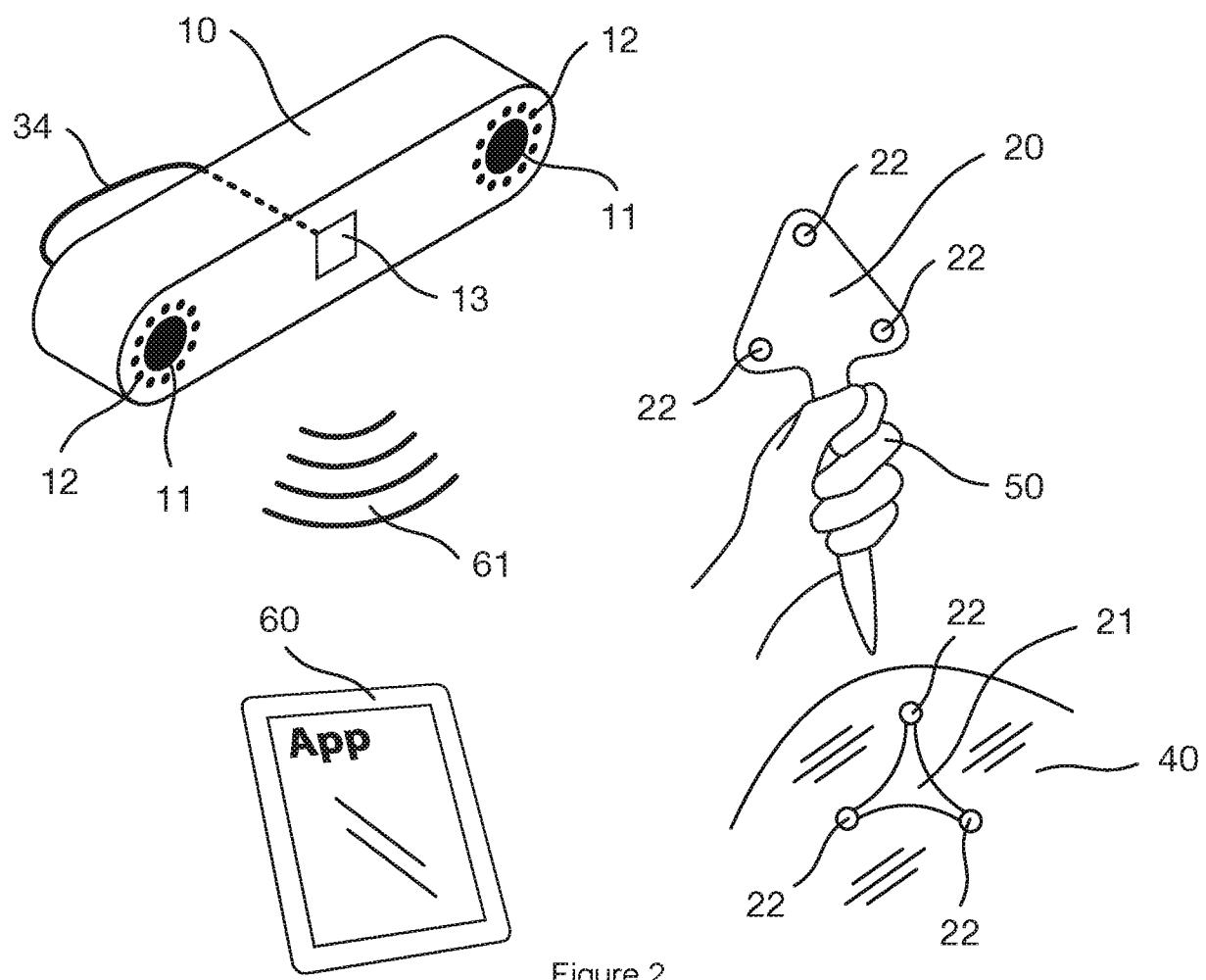
130 — Application
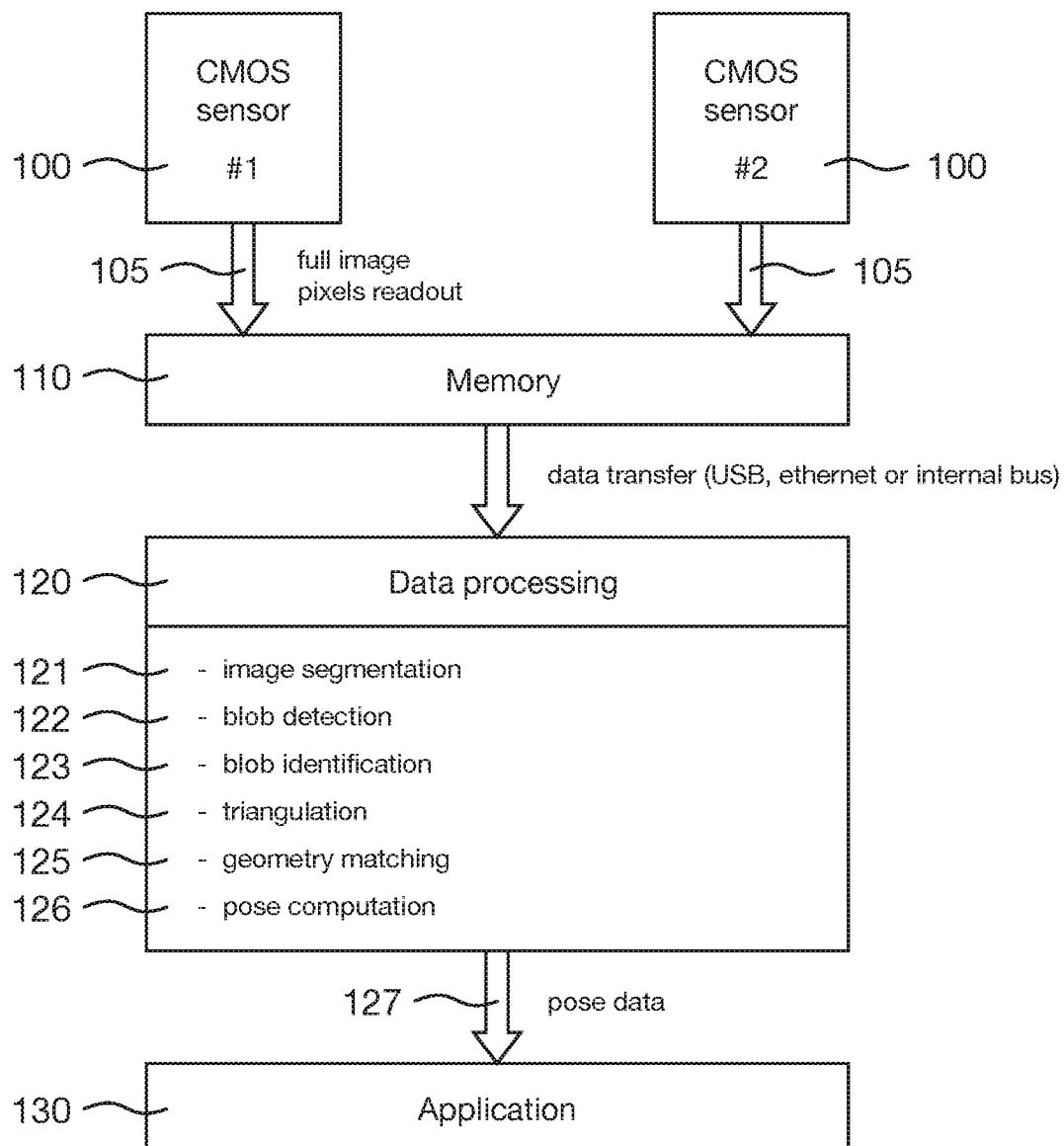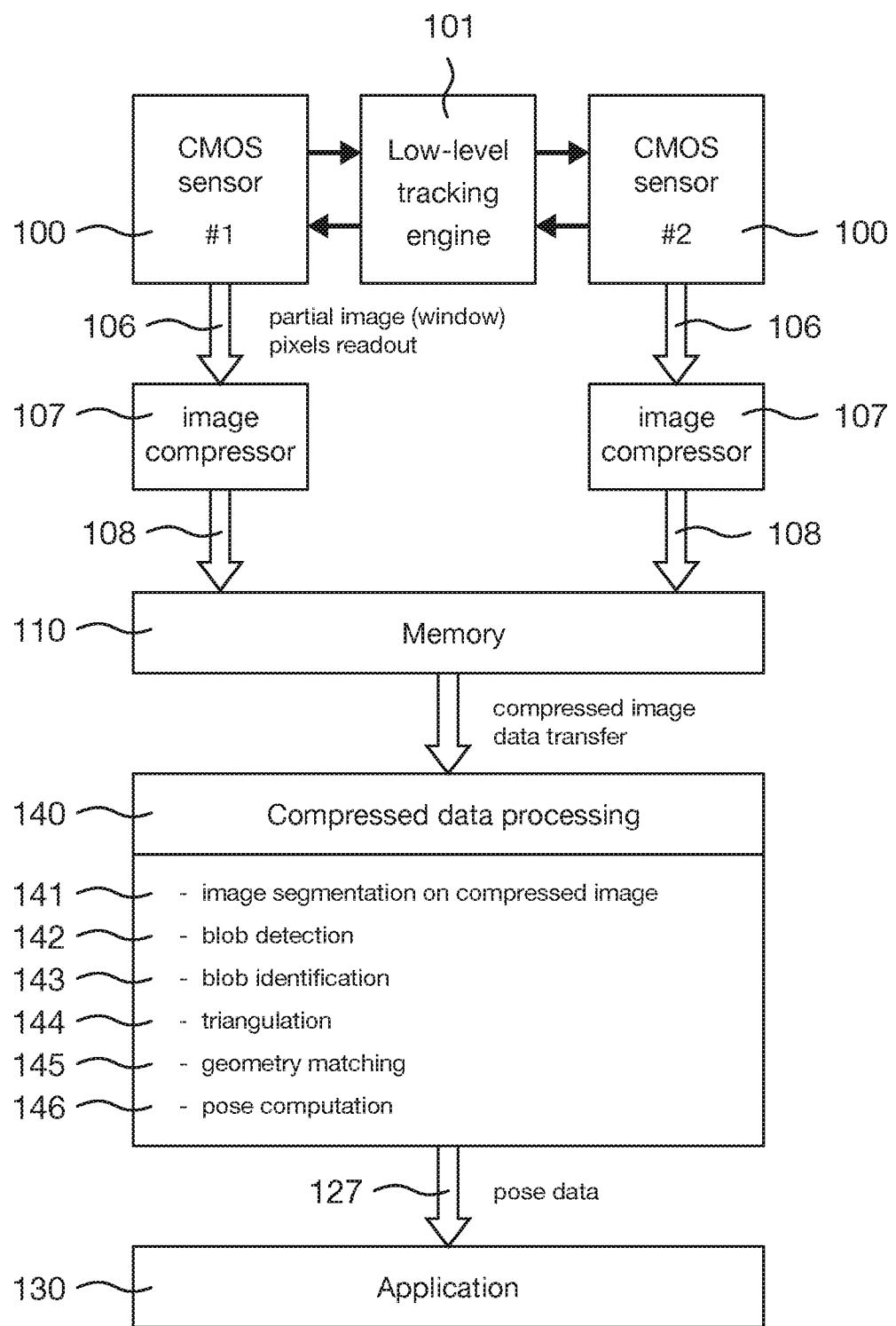
Figure 3

Figure 4

Conventionnal sensor :
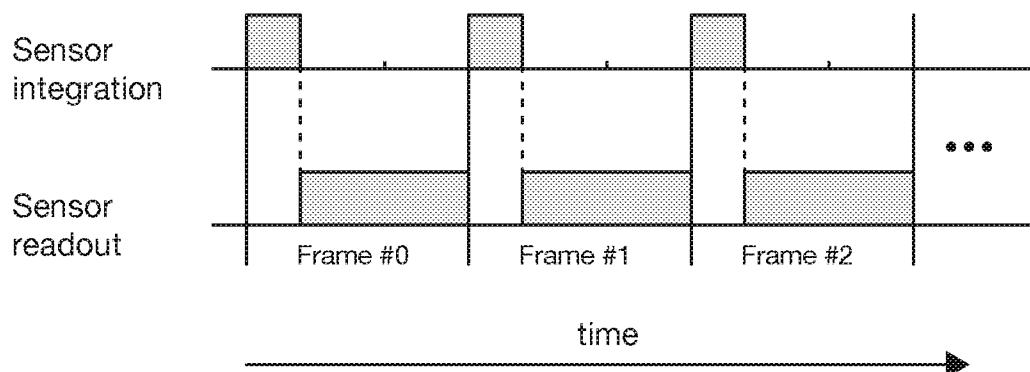
**CMOS sensor with windowing :**
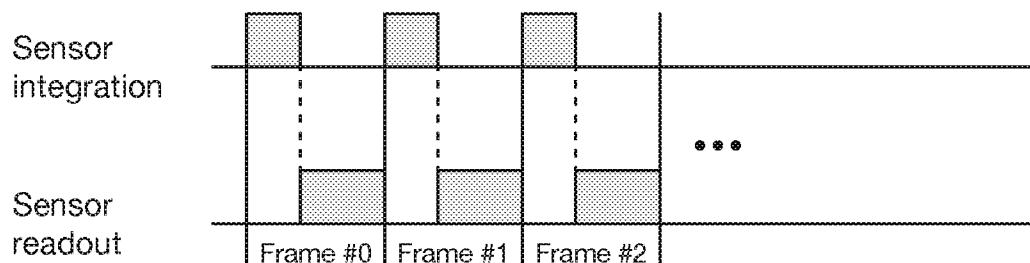

Figure 5
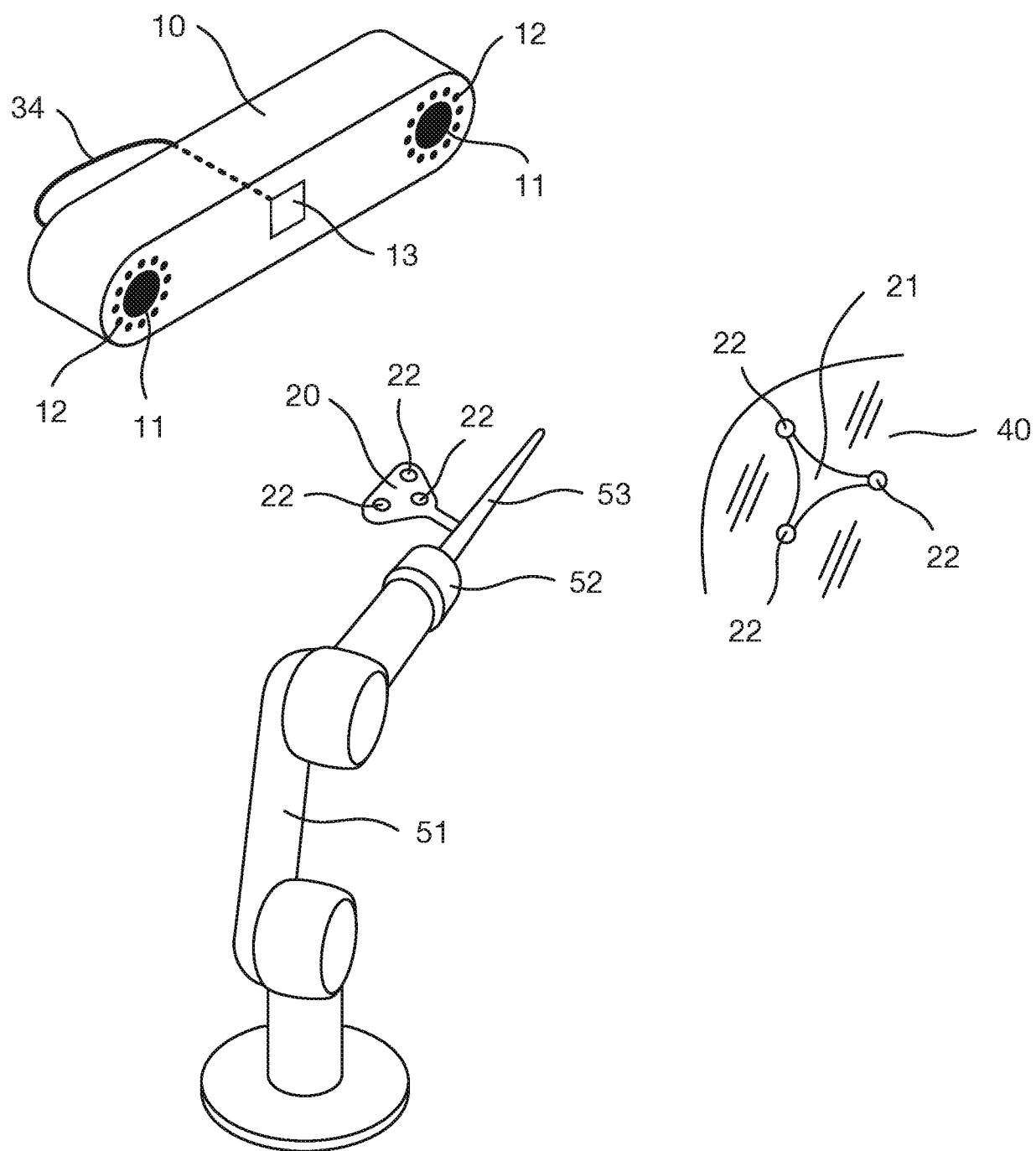
Figure 6
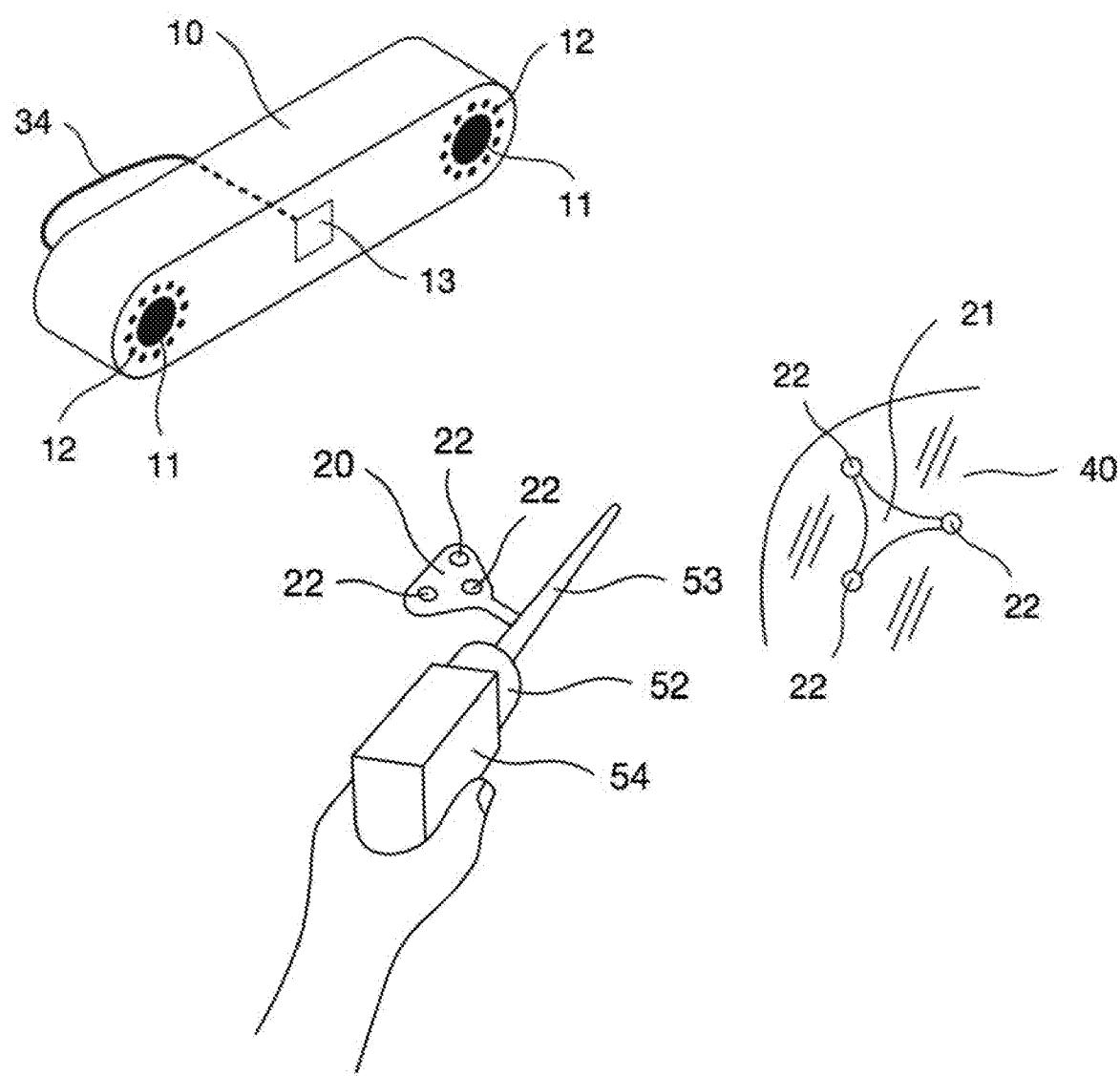
Figure 7
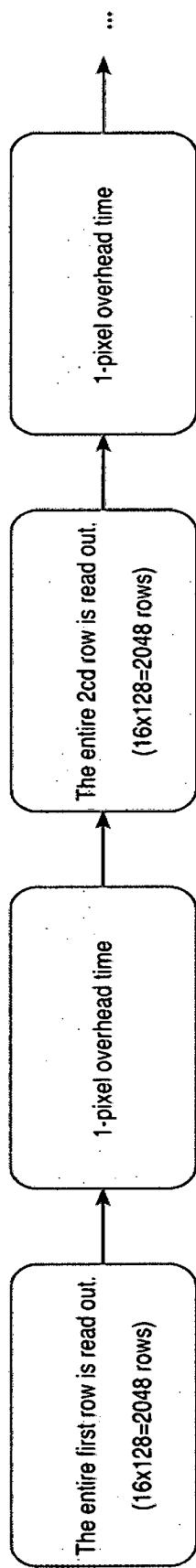
Figure 8
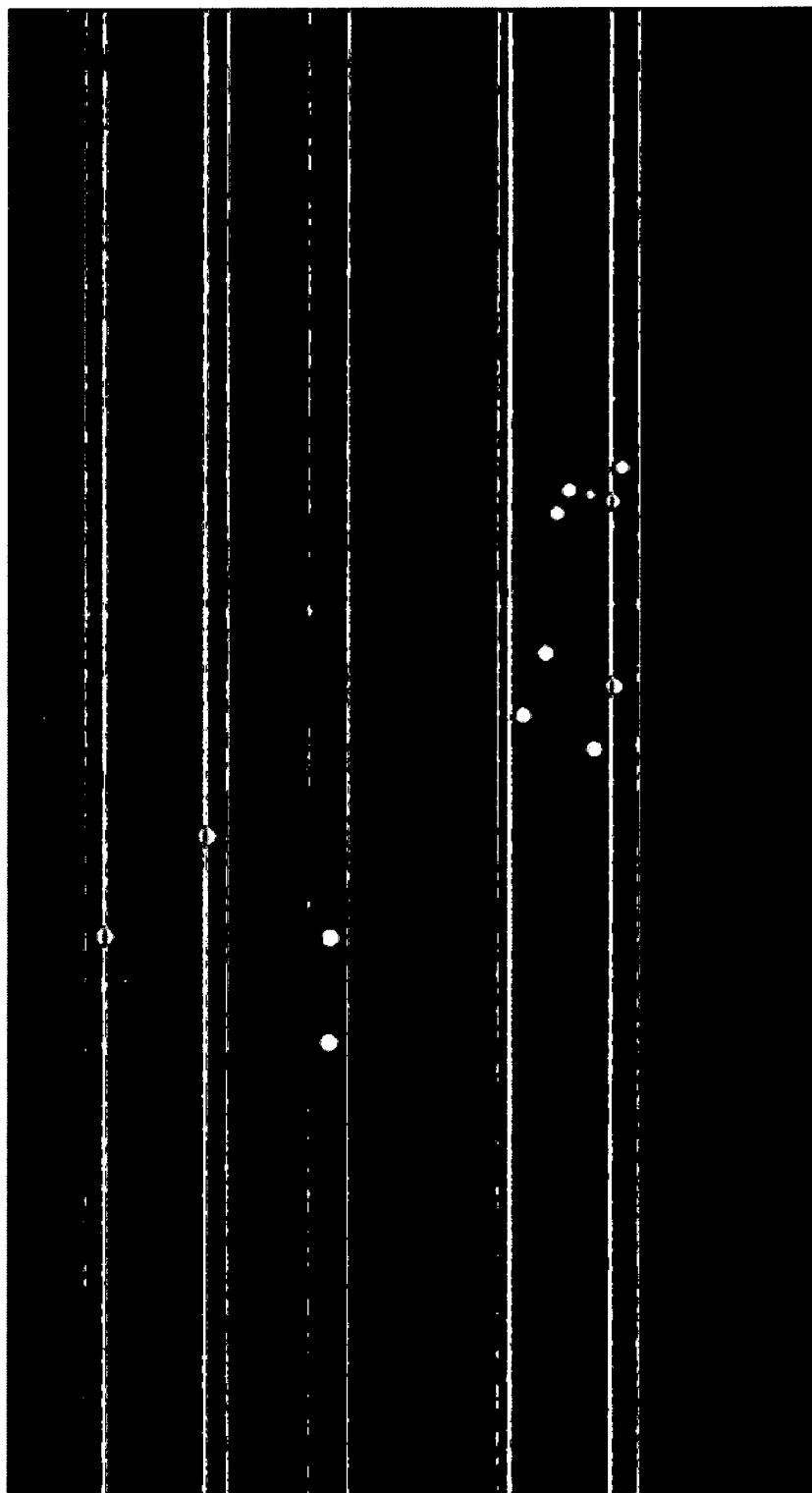
Figure 9
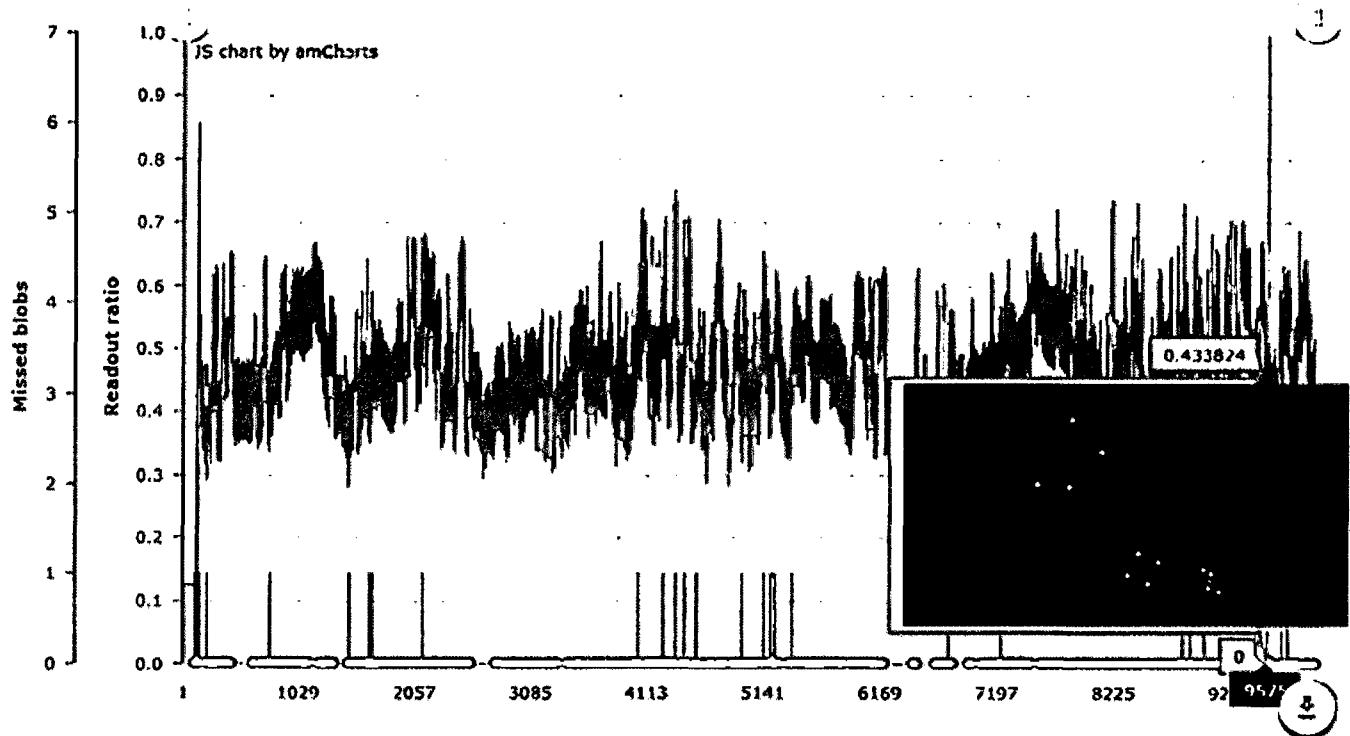
The following graph shows the histogram of the readout ratio:
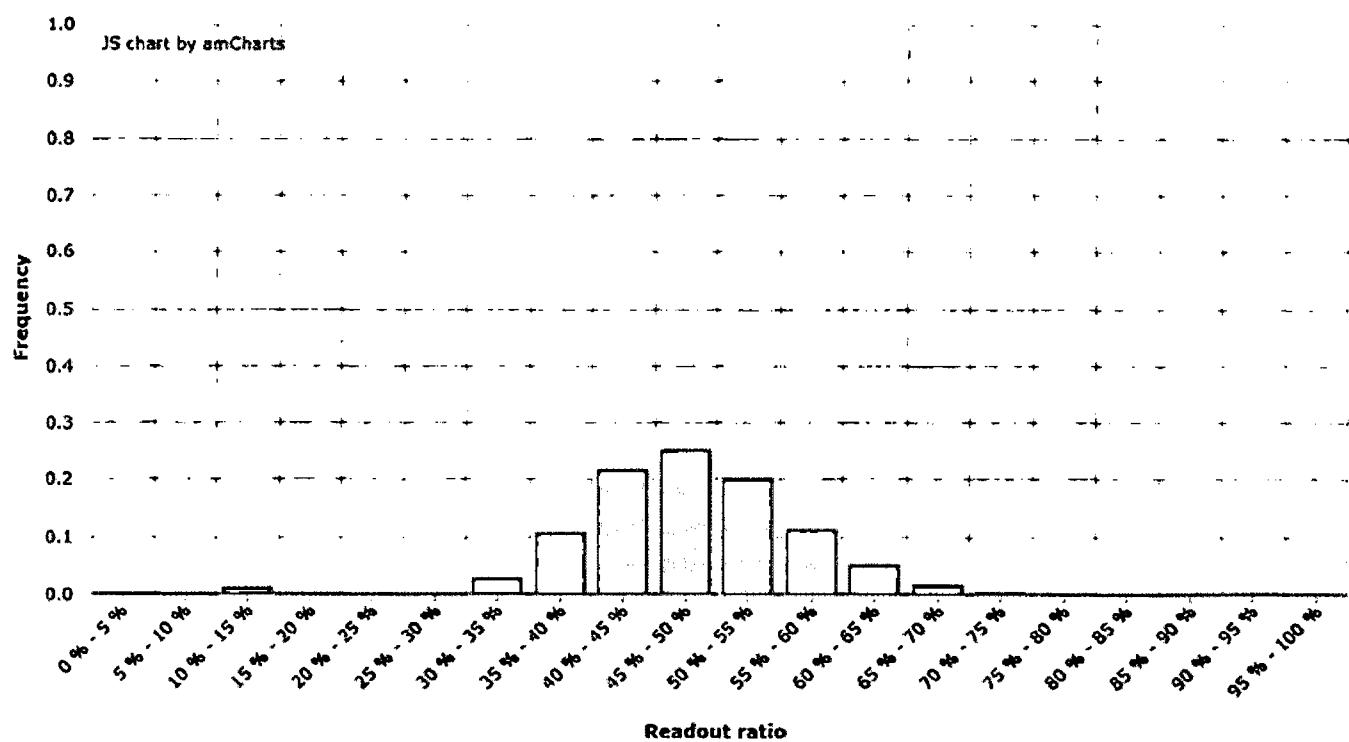
Figure 10

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
INV. A61B34/30   A61B34/20
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

A61B

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2016/139638 A1 (ATRACSYS SÀRL [CH]) 9 September 2016 (2016-09-09)	15-23, 25-28
Y	figures 1,2,7 page 5, line 25 page 13, last paragraph - page 18, last paragraph -----	24
Y	WO 02/076302 A2 (LB MEDICAL GMBH [DE]; LUETH TIM [DE]; BIER JUERGEN [DE]; BIER ANGELIKA) 3 October 2002 (2002-10-03) figures 1,2,8 page 12, last paragraph -----	24
	-/--	

[X] Further documents are listed in the continuation of Box C.   [X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
1 June 2018	13/06/2018

Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Authorized officer Lüddemann, Tobias

C(Continuation).   DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MUEHLMANN U ET AL: "A new high speed cmos camera for real-time tracking applications", ROBOTICS AND AUTOMATION, 2004. PROCEEDINGS. ICRA '04. 2004 IEEE INTERNATIONAL CONFERENCE ON NEW ORLEANS, LA, USA APRIL 26-MAY 1, 2004, PISCATAWAY, NJ, USA,IEEE, US, vol. 5, 26 April 2004 (2004-04-26), pages 5195-5200, XP010768217, DOI: 10.1109/ROBOT.2004.1302542 ISBN: 978-0-7803-8232-9 page 5196 - page 5199; figures 2,3,6,7 -----	15-17,21
X	RIBO M ET AL: "A new optical tracking system for virtual and augmented reality applications", IMTC 2001. PROCEEDINGS OF THE 18TH. IEEE INSTRUMENTATION AND MEASUREMENT TECHNOLOGY CONFERENCE. BUDAPEST, HUNGARY, MAY 21 - 23, 2001; [IEEE INSTRUMENTATION AND MEASUREMENT TECHNOLOGY CONFERENCE. (IMTC):], NEW YORK, NY : IEEE, US, vol. 3, 21 May 2001 (2001-05-21), pages 1932-1936, XP010547287, DOI: 10.1109/IMTC.2001.929537 ISBN: 978-0-7803-6646-6 figures 1,2 the whole document -----	15

2

INTERNATIONAL SEARCH REPORT

Box No. II   Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:   1-14
   because they relate to subject matter not required to be searched by this Authority, namely:

   see FURTHER INFORMATION sheet PCT/ISA/210

2. [ ] Claims Nos.:
   because they relate to parts of the international application that do not comply with the prescribed requirements to such
   an extent that no meaningful international search can be carried out, specifically:

3. [ ] Claims Nos.:
   because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III   Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable
   claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of
   additional fees.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers
   only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this international search report is
   restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest   [ ] The additional search fees were accompanied by the applicant's protest and, where applicable, the
                             payment of a protest fee.

                         [ ] The additional search fees were accompanied by the applicant's protest but the applicable protest
                             fee was not paid within the time limit specified in the invitation.

                         [ ] No protest accompanied the payment of additional search fees.

Form PCT/ISA/210 (continuation of first sheet (2)) (April 2005)

FURTHER INFORMATION CONTINUED FROM   PCT/ISA/ 210

Continuation of Box II.1

Claims Nos.: 1-14

Claims 1-14 relate to a method for treatment of the human or animal body
by surgery, practised on the human or animal body (Rule 39.1 (iv) PCT).
Consequently, no opinion will be formulated with respect to the novelty,
inventive step and industrial applicability of the subject-matter of
these claims (Article 17(2)(a)(i) PCT), see also p.7 l.30 in that the
marker may be screwed on a patient, p.9 last para. with fig.6 specifying
a surgical tool, p.13 l.10 the tool may be a drill or biopsy needle.

International application No
PCT/IB2018/050899

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 2016139638	A1	09-09-2016	EP 3265009	A1	10-01-2018
			US 2018049809	A1	22-02-2018
			WO 2016139638	A1	09-09-2016
WO 02076302	A2	03-10-2002	AT 473688	T	15-07-2010
			CA 2447813	A1	03-10-2002
			CN 1536975	A	13-10-2004
			EP 1392174	A2	03-03-2004
			JP 4439812	B2	24-03-2010
			JP 2005509449	A	14-04-2005
			US 2004157188	A1	12-08-2004
			WO 02076302	A2	03-10-2002