The Kinect as an interventional tracking system

Article *in* Proceedings of SPIE - The International Society for Optical Engineering · February 2012

DOI: 10.1117/12.912444

5 authors, including:

Philipp J Stolka
ADVITOS GmbH

Emad M. Boctor
Johns Hopkins University

Gregory D. Hager
Johns Hopkins University

Michael A Choti
University of Texas Southwestern Medical Center

The Kinect as an Interventional Tracking System

Xiang Linda Wang[1], Philipp J. Stolka[1], Daniel Carnegie[2], Emad Boctor[1,2], and Michael Choti[2]
1 The Johns Hopkins University, Baltimore MD, USA
2 Johns Hopkins Medical Institutions, Baltimore MD, USA

Abstract

This work explores the suitability of low-cost sensors for "serious" medical applications, such as tracking of interventional tools in the OR, for simulation, and for education. Although such tracking – i.e. the acquisition of pose data e.g. for ultrasound probes, tissue manipulation tools, needles, but also tissue, bone etc. – is well established, it relies mostly on external devices such as optical or electromagnetic trackers, both of which mandate the use of special markers or sensors attached to each single entity whose pose is to be recorded, and also require their calibration to the tracked entity, i.e. the determination of the geometric relationship between the marker's and the object's intrinsic coordinate frames. The Microsoft Kinect sensor is a recently introduced device for full-body tracking in the gaming market, but it was quickly hacked - due to its wide range of tightly integrated sensors (RGB camera, IR depth and greyscale camera, microphones, accelerometers, and basic actuation) - and used beyond this area. As its field of view and its accuracy are within reasonable usability limits, we describe a medical needle-tracking system for interventional applications based on the Kinect sensor, standard biopsy needles, and no necessary attachments, thus saving both cost and time. Its twin cameras are used as a stereo pair to detect needle-shaped objects, reconstruct their pose in four degrees of freedom, and provide information about the most likely candidate.

Keywords

Image-guided; tracking; calibration; Kinect; stereo camera; interventional; low-cost

1 Description of Purpose

1.1 Aims

The main driver for this work was to explore the suitability of low-cost sensors for “serious” medical applications, one of which is the tracking of interventional tools in the OR. Although such tracking – i.e. the acquisition of pose data e.g. for ultrasound probes, tissue manipulation tools, needles, but also tissue, bone etc. – is well established, it relies mostly on external tracking devices such as optical or electromagnetic (EM) trackers. Both approaches mandate the use of special (optical) markers or (EM) sensors that have to be attached to each single entity whose pose is to be recorded, and also require the calibration of those attachments to the tracked entity, i.e. the determination of the geometric relationship between the marker’s and the object’s intrinsic coordinate frames. Not only is this calibration cumbersome and lengthy, but also it has to be strictly maintained, and has to be repeated for each new object that is to be tracked. Furthermore, the attachment has to be in the line of sight of the cameras (for optical tracking) or outside of ferromagnetic interference (for EM tracking) at all times, lest “tracking loss” occur. Last not least, conventional tracking systems carry substantial sticker prices, partly justified through their accuracy, partly through their relatively limited production volumes.

Nevertheless, the benefits often outweigh those drawbacks – knowing the current spatial relationships enables navigation, i.e. the guidance of the operator by means of displaying location information of tools and regions of interest (ROIs), even when direct observation is impossible, such as during endoscopic or other minimally-invasive interventions. Therefore, providing guidance using low-cost solutions might be appreciated not only in the traditional application fields, but also in training and education, OR simulations, and less-developed markets. The Microsoft Kinect sensor is a relatively recently introduced device for full-body tracking originating in the gaming market, but was very quickly hacked – due to its wide range of tightly integrated sensors (RGB camera, IR depth and greyscale camera, microphones, accelerometers, and basic actuation) – and used beyond this application area. As its field of view and its accuracy are within reasonable usability limits, we describe a medical needle-tracking system for interventional applications based on the Kinect sensor, standard tools, and no necessary attachments.

1.2 State of the Art

Different approaches to track surgical tools exist, ranging from robotic needle holders on the high end of the complexity range through optical and EM tracking to mechanical biopsy guides, constraining the needle motions in a pre-defined fashion. In our previous work [Stolka-2011] we described stereo-camera needle tracking coupled with SLS surface reconstruction. We elaborate on this approach in the present paper. Other approaches differ in the details of implementation (e.g. [Rohling-2011] using a single-camera needle tracking system). A more complete overview will follow in the full paper.

2 Methods

2.1 Materials

The proposed tracking system consists of both hardware and software. The hardware includes an off-the-shelf standard Microsoft Kinect and an Apple iMac (Core i7 2.93GHz, 5GB RAM). The Kinect works as a 3D depth sensor and as a stereo vision system which considers the IR camera and optical camera as a stereo pair. The stereo pair needs to be calibrated before use, but this only has to be done once. The software was developed using C++, using the OpenGL and Freenect libraries.

2.2 Algorithms

The system should be able to find needles during standard procedures automatically in real time, which means any needle visible to the sensors on top of the background features shall be detected and its pose in stereo camera coordinates be computed with a frame rate sufficiently high to allow interventional operation. 2D needle detection and 3D needle reconstruction can be achieved using image pairs from the two calibrated Kinect IR and optical cameras. The reason we use the IR/optical camera pair rather than the depth image obtained by Kinect is that needles are invisible in the depth images (Fig.1). Therefore, depth images alone cannot be used to detect needles. The proposed algorithm can be summarized in three steps: (1) find needle candidates, (2) prune candidate set, and (3) reconstruct 3D needle pose.

Before the first use of the system, the stereo system has to be calibrated to obtain the focal lengths, principal points, distortions and rotation/translation vectors between two cameras. The calibration was conducted using the standard OpenCV implementation.

2.2.1 Find needle candidates:

Any needle should be able to be found by looking for prominent, straight-line segments in an image.

Strategies to find those features in RGB images are well established. We apply edge-detection on the grey-scale image and then compute the Hough transform on the resulting binary image. The *n* strongest peaks in the transform are selected as needle candidates. Using the probabilistic Hough transform, one can find the presumed end points of detected 2D lines.

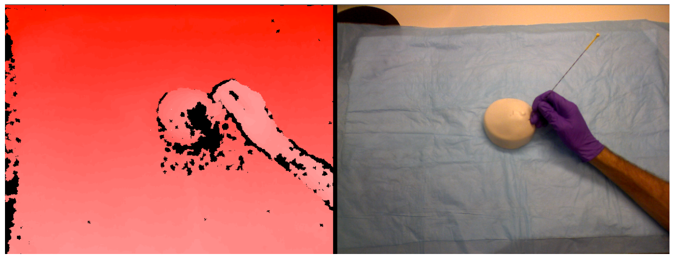
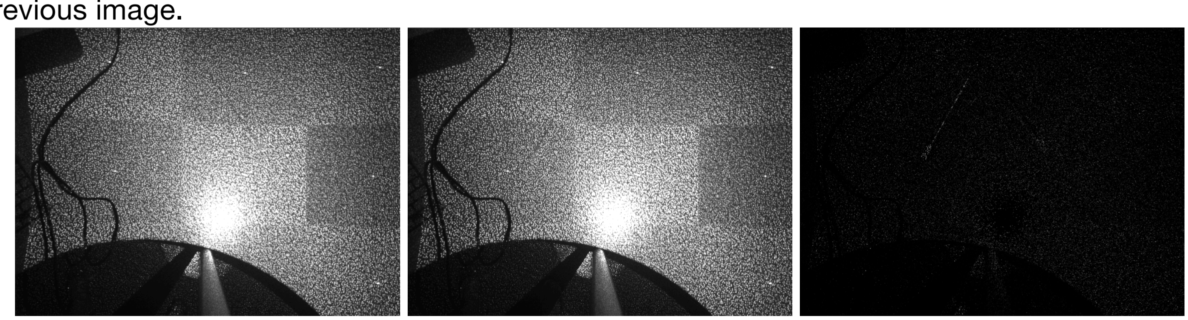


Figure 1. Depth image from Kinect (left); RGB image from Kinect (right)

The situation is more difficult in IR-grey images, which are usually very noisy and have low contrast (Fig.3). Because the IR comes mainly from the IR source on the Kinect, there is always a strong shadow of the needle shown in an IR image, which is often much more visible than the real needle. The projected IR pattern by Kinect can confuse the line detector too. In order to overcome the described problems in IR images, we do not detect the needle directly in IR images, but rather in a difference image obtained by subtracting one IR image from the previous image.

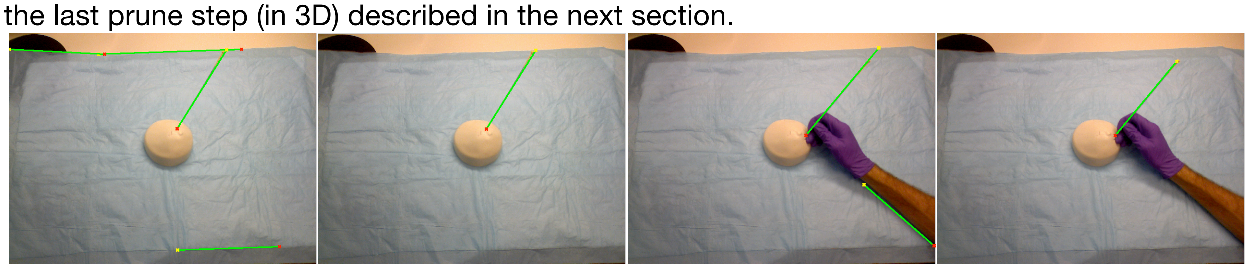


(a) (b) (c)

Figure 3. Subtracting one IR image (a) from the previous one (b) to obtain a less noisy high contrast difference image (c).

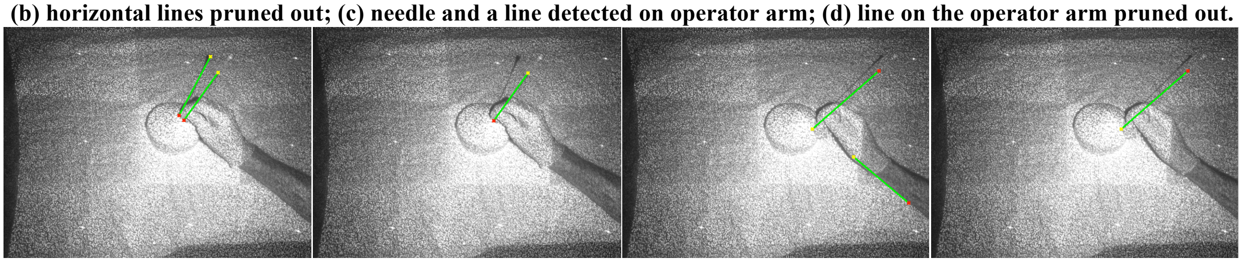
2.2.1 Prune candidate set:

The procedure above results in a potentially large number of line segments all over the two stereo images, most of which do not correspond to any actual needle. Therefore, we need to prune out the false needle lines, and the probable needle line candidates are kept for further processing. We defined several rules, some of which discard candidates in the 2D images, with the last prune step (in 3D) described in the next section.



(a) (b) (c) (d)

Figure 4. Detected 2D needle lines shown as green lines in RGB images: (a) needle, with some horizontal lines detected; (b) horizontal lines pruned out; (c) needle and a line detected on operator arm; (d) line on the operator arm pruned out.



(a) (b) (c) (d)

Figure 5. The detected 2D needle lines are shown as green lines in IR images: (a) both the needle and its shadow detected; (b) the shadow pruned out; (c) needle and a line detected on the operator arm; (d) line on the operator arm pruned out.

2.2.2 3D needle generation:

3D needle lines are generated from all the combinations of 2D needles in stereo image pairs. A line representation is computed using stereo calibration information, providing two distinct points C_{p_upper} and C_{p_lower} on the 3D line. As the needle is detected with 4 degrees of freedom (DoF) – the axis-rotational and axis-translational DoF are not recoverable – this representation is sufficient. There remain false 3D needles to be pruned out. We apply the last pruning rule on all the generated 3D needles which is checking whether the two recovered 3D points on the needle are within a proper depth range (as defined by the geometric setup).

2.3 Experiments

Two sets of experiments were carried out. The first set was designed to test the needle detection method. We placed a needle on a uniform surface and fixed a Kinect vertically above the surface at a distance of about 700mm. In order to test the robustness of the discussed methods, the needle was placed on different surfaces at different positions with different orientations. The second set of experiments was designed to simulate the real situation in an OR. We covered the table with a medical cloth and then placed a breast phantom on it, with the Kinect fixed vertically above at a distance of about 470mm. An operator held a needle in his hand and inserted it at different positions of the phantom with different orientations.

3 Results

The images generated by Kinect are sized 640x480 at 30 fps. The algorithm is able to detect needles in real time. Typical scenes allow the immediate detection of generic needles (using Canny edge detection, Hough line transform, and lots of stereo line combinations) against general backgrounds without preparation. However, in low contrast IR images, additional processing (described in the full paper) may be necessary.

Skin surfaces (phantom and human) can be reliably reconstructed within the Kinect sensor's working range. (In our experiments, this range is between 500mm and 1000mm). The surface is used to prune out the false 3D needles. Under the simulated OR conditions, the needle is reliably detected from both IR and RGB images under different conditions (different position, orientation, surface, lighting, etc.).

4 New or breakthrough work to be presented

Without any modifications to the sensor or the needles, an off-the-shelf Kinect can be used as an interventional tracking system. Use in laboratory conditions is described in this abstract, with realistic use cases described in the full paper.

5 Conclusions

We can use the low-cost Kinect sensor for interventional needle tracking. In clinical or educational practice, this allows "second-opinion" tracking in addition to EM or optical in navigated image-guided interventions; with no calibration or marker attachment necessary and no sterility issues. It is possible to use several Kinects in parallel if lines-of-sight are an issue in the OR. In future work, it may be used to track multiple and different tools as well.

The work has not been submitted for publication or presentation elsewhere.

References

[Stolka-2011]
[Rohling-2011]