Virtual Knee Arthroscopy Using Haptic Devices and Real Surgical Images

Shahzad Rasool[1], Alexei Sourin[1], Vladimir Pestrikov[2], and Fareed Kagda[3]

[1] Nanyang Technological University, Singapore
{shahzadrasool,asourin}@ntu.edu.sg
[2] Moscow Institute of Physics and Technology, Russia
vipestr@gmail.com
[3] Alexandra Hospital, Singapore
fareed_kagda@juronghealth.com.sg

Abstract. Knee arthroscopic surgery is performed on the knee joint by making small incisions on the skin through which an endoscopic camera (arthroscope) is inserted along with miniature surgical instruments. It demands from the surgeons to acquire special motor-skills. A few commercial simulators are available for arthroscopic surgery training however the area is still very open for research and development. In contrast to the common fully-3D way of simulation of knee arthroscopy, we propose a hybrid image-based approach where real arthroscopic videos are converted to panoramic images which are augmented with 3D deformable models of the tissues as well as 3D models of surgical instruments. The motions of the virtual arthroscope and the instruments are controlled by two desktop haptic devices. The hybrid virtual scene is visualized through a moving circular window, which follows the motion of the virtual arthroscope.

Keywords: virtual arthroscopy, haptics, image-based.

1 Introduction

Arthroscopic surgery procedures are performed on joints in a minimally invasive way by making small incisions on the skin through which a miniature endoscopic camera (arthroscope) is inserted along with special surgical instruments. The arthroscope is a wide-angle oblique-viewing camera with view angles of 30º, 70º or 90º and a light source attached to its end. Arthroscopy demands from the surgeons to acquire special motor-skills while learning complex stepwise tasks including positioning (triangulation) of the instrument in front of the surgical camera (arthroscope), examination of the tissues in the surgical area, and using different instruments for cutting away and removing damaged parts of the tissues. The major challenge for the trainee surgeons is that the surgical area is observed as a 2D image on the video monitor which is physically displaced from the remotely controlled instruments. The specifics of the side-looking arthroscope creates further challenges in its manipulation to be able to see the surgical instruments inserted through the incision made a few centimeters

V.G. Duffy (Ed.): DHM 2014, LNCS 8529, pp. 436–447, 2014.

away from the camera incision point. To achieve these skills, the surgeons have to go through special training programs.

Simulation of arthroscopic surgery for training purposes can be done in virtual reality environments using common personal computers and haptic devices. It is very feasible because the surgeons are detached from the actual three-dimensional surgical scene, see it as a two-dimensional image on a video monitor, which can be replaced by a computer monitor, and operate within the scene with remotely controlled instruments, which can be quite realistically simulated with desktop haptic devices.

In this paper we first survey the works related to virtual arthroscopy done with haptic devices and camera models for oblique-viewing endoscopes (Section 2). Then we address the problem of modeling virtual arthroscopic cameras controlled by desktop haptic devices in a hybrid image-based virtual environment (Section 3). We consider /1/ how to simulate the constraint camera motion which is characteristic for minimally invasive surgery, and /2/ how to simulate photorealistic views corresponding to the actual arthroscopic cameras. The paper is concluded in Section 4.

2 Related Work

Haptics-based simulation has become a growing research topic in arthroscopy simulation. Thus, Sherman et al. [1] developed a virtual environment knee arthroscopy training system where a custom-made force feedback device is attached to the mock instruments to provide haptic feedback to the user. Gibson et al. [2] used volumetric object representations and smoothed these models before surface normal calculation to ensure stability of the haptic algorithm presented. Bayona et al. [3] presented a low-cost arthroscopic simulation system where a commercial laparoscopic interface is employed as the surgical instrument for delivering haptic feedback. Mabrey et al. [4] used commercial haptic devices to interact with the underlying volumetric representation of the knee. Pinto et al. [5] presented an orthopedic surgery simulator with a mixed surface and volumetric models were used for calculating the force feedback. Wang et al. [6] proposed a surgical procedure simulation system for training of arthroscopic anterior cruciate ligament reconstruction, where two specially designed force feedback models were used for the haptic rendering of probing and drilling operations. A few commercial simulators for arthroscopic surgery training have been developed. Among them are ArthroS from VirtaMed [7], ArthroSim from TolTech [8], ARTHRO-Mentor from Simbionix [9], and SIMENDO arthroscopy from SIMENDO [10].

A few camera models have been proposed to incorporate oblique-viewing property of endoscopic cameras. Yamaguchi [11] was the first to formulate a camera model and calibration method for such cameras to be used in an augmented reality system. Based on Tsai's camera model, it establishes the extrinsic parameters of the camera as a function of rotation about the arthroscope axis. Calibration is performed by attaching optical markers to the camera and tracking them. Rotation parameters are measured using a rotary encoder. In [12], the calibration was simplified by tracking the arthroscope rod instead of camera head hence reducing the number of parameters to be estimated. However, two additional optical trackers were used in an attempt to eliminate the need for the rotary encoder. Buck et al. [13] generalized Yamaghuchi's method by incorporating the changes to camera intrinsic parameters and a radial distortion component. Optical trackers are used for calibration but the calibration result is slightly less accurate. A real-time method for calibration and removal of radial distortion is proposed in [14], along with a formulation of relative rotation between endoscope axis and camera head. These methods provide calibration procedures for endoscopic camera so that a mapping between the 3D scene and the resulting 2D image can be extracted.

In this paper, we propose modeling of arthroscopic camera in hybrid image-based virtual environments where 3D reconstruction of the anatomical structures is mostly avoided. Instead, real arthroscopic videos are converted to panoramic images which are augmented with 3D deformable models of the tissues as well as 3D models of surgical instruments as in [15]. The location and orientation of the virtual camera are derived from the combined constraint motion of the arthroscope rod. A new formulation of the camera view direction in terms of arthroscope axis rotation is presented with physical simulation of the incision (pivot) point for the arthroscope and the instrument rods.

3 Modeling the Arthroscopic Camera in Hybrid Image-Based Virtual Environment

We propose a hybrid image-based approach for arthroscopy simulation where real arthroscopy images are mostly used for visual and haptic rendering rather than 3D models of the surgical area. At the preprocessing stage, a panoramic image of the entire knee cavity has to be created by stitching some frames of a real arthroscopy video. This image file then is visualized through a moving circular window which follows the motion of the virtual arthroscopic camera (Fig. 1).

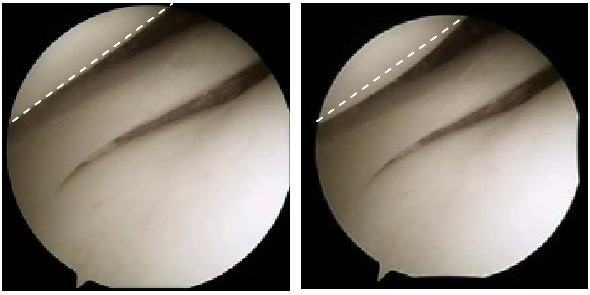
Fig. 1. Panoramic image created by stitching images obtained from the actual surgical video and the hybrid simulated view

The image displayed should correspond to what can be seen for any given location and orientation of the arthroscope. The panoramic backdrop image is then augmented with a few 3D models of deformable tissues as well as models of the surgical instruments which will be seen by the camera. The images displayed on the virtual monitor are then very close to those displayed during the actual surgery.

To haptically render the backdrop image as if it were the actual 3D scene, a depth map is then extracted from the intensity values of its pixels while some noise removal filters are applied to improve the approximation of depth map from pixel intensities, as it was previously reported in [16-17]. This approach can produce believable interaction with much shorter turnout time than that of the full 3D modeling approach provided there is a reliable and quick pipeline for making panoramic images from the surgical videos taken during the operation. A straight-forward but tedious way of doing it is to manually make an image by adding the matching parts from the consecutive images. We have come up, however, with a method allowing for automatic making of panoramic images from the surgical videos. In the rest of this section we will consider the novel issues related to making the panoramic images and the specifics of simulating the arthroscopic camera in this hybrid virtual environment.

3.1 Automatic Generation of Panoramic Images

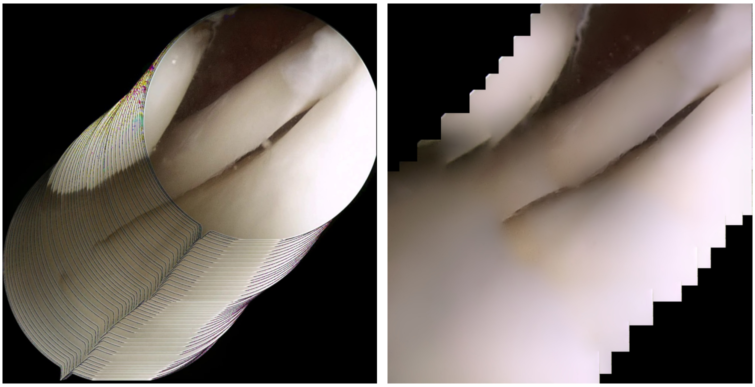
There are a few software tools commonly used for making panoramic images of streets and nature scenes from video clips, such as, Arcsoft Panorama Maker [18], SoftOptics Panoptica [19], and Microsoft Image Composite Editor [20]. However, they cannot be used in case of the arthroscopic videos since their individual frames are lacking clear feature points [21] which can be compared and used for finding the stitching transformations. To identify such feature points in the individual arthroscopic images, we propose to use the methods based on analysis of the gradient of brightness. However, this gradient is rather low in the arthroscopic images. Hence, we increase the image contrast by using image histogram of brightness. To improve the precision of calculations, we need to first equalize the illumination of the objects. For this task, we used SSR (Single Scale Retinex) algorithm. We then use SURF (Speeded-Up Robust Features) algorithm [22] for extracting feature points from the images, while library FLANN (Fast Library for Approximate Nearest Neighbors) [23] was used for comparing the features. The selection of SURF was based on the surveys done in [24] and [25]. Next we have to find a homographic transformation of source images to match the corresponding singular points belonging to different images. We used RANSAC (RANdom SAmple Consensus) method to find the initial approximation. The computed homographic transformation matrix is further refined with the Levenberg-Marquardt method in order to reduce the re-projection error even more. After this, automatic finding of the matching features becomes more reliable, and individual images can be automatically stitched into the panoramic image. However, while applying this algorithm to individual video frames, we noticed that the central parts of the images stich better than the peripheral parts. This happens due to the fish-eye distortions in the individual images. The distortions can be eliminated from each contributing video frame before they are processed for stitching. This requires taking with the arthroscopic camera a calibration image of the Cartesian grid. Knowing the parameters of the camera and the fish-eye projection transformation, the function mapping the original pixels to non-distorted image can be obtained (Fig. 2).



(a) (b)

Fig. 2. Removal of equisolid fisheye distortions for the images taken with "LINVATEC HD4300 4mm" with the focal length 18.5 mm and the angle of view 30º. (a) A Frame from the actual surgical video (b) The respective corrected frame. Notice the differences at the outlined parts.

The final panoramic image obtained from a short arthroscopic video which was first corrected to eliminate fish-eye distorted and then processed by the proposed algorithm followed by additional blending, as in [26], is shown in Fig. 3.



(a) (b)

Fig. 3. Automatic stitching and blending of the whole arthroscopic video (a) into the panoramic image (b)

3.2 Modeling Arthroscopic Camera and Surgical Instruments

Two desktop haptic devices are used to control the virtual arthroscopic camera and the surgical tools. To constrain their motion as that of the surgical camera and instruments pivoting about the point of insertion into the joint, we modified the haptic devices by extending them with actual surgical tools, as shown in Fig. 4. The handle of each device moves in and out and rotates about its fixed pivot point as in the actual surgery. As a result, the constrained motion of the haptic devices reflects the insertion of the camera and the tool through the incisions on the body of the patient.

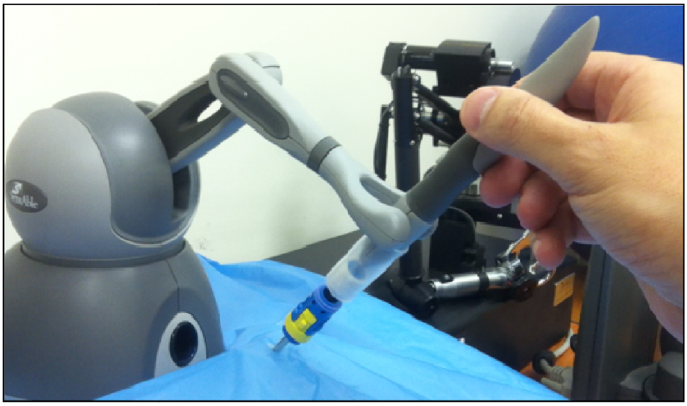


Fig. 4. Geomagic Touch haptic device modified for minimally invasive surgery simulation

The relative positions of the haptic interface point (HIP), the incision point and the virtual camera are shown in Fig. 5.

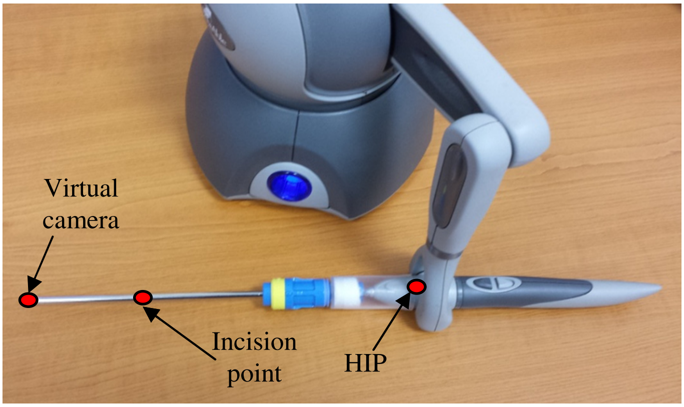


Fig. 5. Location of HIP, virtual camera and incision point on the modified haptic device

First, the virtual camera must be calibrated which refers to the calculation of the incision point coordinates in the device coordinate system. This is implemented by moving the device handle and pressing the primary stylus button at two different handle positions. The HIP position and position of the virtual camera, which is at a specific distance from the HIP defined by the length of the device extension, are recorded at the two instances when the stylus button was pressed. The intersection between the two lines defined by these two pairs of points gives the coordinates of the incision point. Note that the incision point can be located anywhere between the HIP and the virtual camera based on the calibration.

For a fixed distance between the HIP and the incision point, the movement of the HIP is restricted to a virtual spherical surface centered at the incision point. For such motion, the virtual camera traverses along another spherical path, also centered at the incision point. As the device is moved in and out, the radii of the two spherical paths are altered. The location of the HIP is used to calculate the location of the virtual camera in a coordinate system centered at the incision point, as shown in Fig. 6.

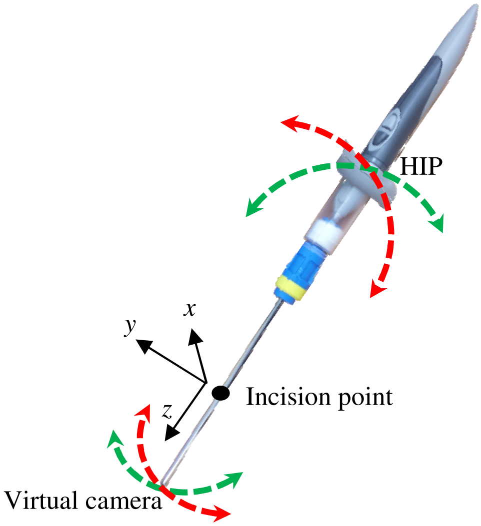


Fig. 6. Corresponding motion of HIP and virtual camera along spherical paths

A similar formulation is then used for the instrument haptic device to compute virtual tool position. The transformation for the virtual tool is such that its incision point and hence the corresponding local coordinate system has its origin on the *x*-axis of the camera coordinate system. The distance between the two insertion points is user tunable and corresponds to the distance between the two incision points on the real joint.

3.3 Visualization of the Surgical Area

A special method has to be used to model visualization with the oblique-viewing arthroscopic camera. Here, we need to display the relevant part of the existing panoramic image following the motion of the virtual camera controlled by the haptic device mimicking the actual surgical camera. Hence, we have to solve a problem of displaying various parts of the image as if they were seen by the actual surgical camera.

To this end, the position of the virtual camera is mapped onto the image coordinates. Let's first consider the case of a forward looking camera. As mentioned previously, the virtual camera moves along a virtual sphere. Hence, for each position of the camera on the sphere, a pixel P in the image is selected as the center pixel of the camera view. This is done by using the latitude and longitude angles formed by the virtual camera with the local coordinate system defined at calibration. Notice that the inward or outward movement of the camera would not alter the pixel at the center of the view as the latitude and longitude angles are constant for such motion. Any rotation of the arthroscope about its axis is then a rotation of the image about this pixel. However, since the actual arthroscopic cameras are usually oblique-viewing, therefore the selected pixel cannot be used as the center pixel of the view. Instead, a pixel *P'* at a distance *r* from *P* is used as the center of the view. Any rotations of the arthroscope about its axis will then translate into the rotation of pixel *P'* about pixel *P*. Thus, the rotation of the arthroscope about its axis results in a set of center pixels that form a circle with radius *r* on the image. The distance *r* between pixels *P* and *P'* is controlled by the inward or outward movement of the device. Thus, as expected, such motion brings the offset pixel *P'* closer to *P*.

In order to validate the motion of the virtual camera, we use a regular grid image and plot the pixels at the center of the camera view for a few frames when the camera device handle is rotated about its axis at different depths of insertion. This is shown in Fig. 7 where the location of these pixels is outlined as the camera device rotates. Figs. 7 (c-d) show the same for a panoramic arthroscopic image stitched together from the frames of the actual surgical video.

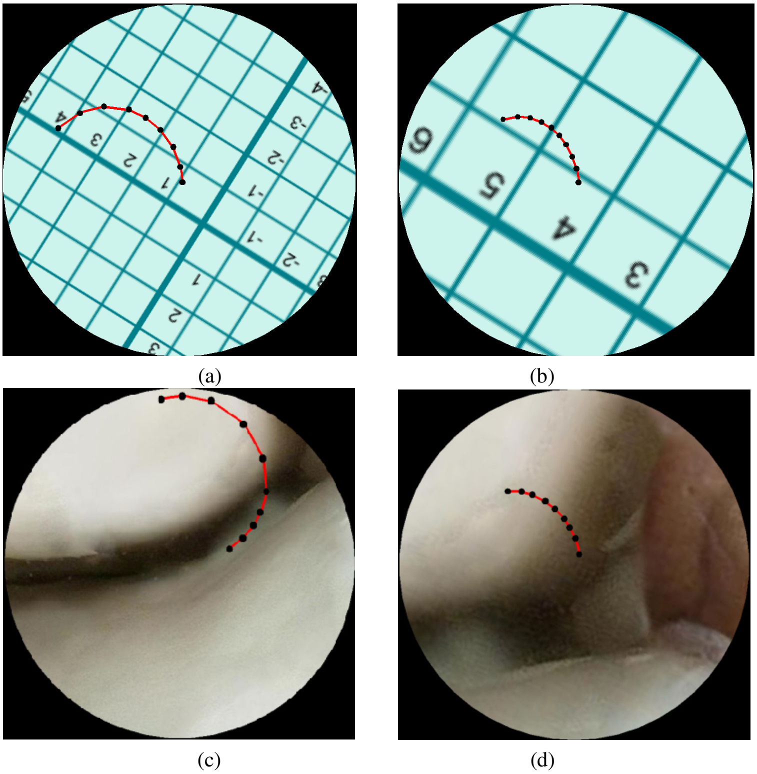


Fig. 6. Pixels at center of the camera view as the camera device is rotated about its axis at different depths of insertion

3.4 Depth Estimation for Haptic Rendering

To haptically render the backdrop image as if it were the actual 3D scene, the depth for collision detection is calculated from the image with pixel precision. Since arthroscopic images are frontally illuminated, image intensity has a direct correspondence to scene depth. However, in order to account for different contributions of various color components to the grayscale value of each pixel, some colors must be filtered. This is because certain muscle and tissue parts may have a higher dominance of one color component as compared to the white bones. We know that the knee cavity does not usually have green or blue colored areas but rather either have white bones, menisci or cartilage, or have some muscle parts with a dominant red component. The original RGB image is thus color-filtered to reduce the dominance of red component and converted to grayscale to use its intensity values as depth map, which is normalized to occupy the entire depth range of the scene. A second filtering pass, for example using a median filter, is applied to remove noise in the extracted depth map as shown in Fig. 8.

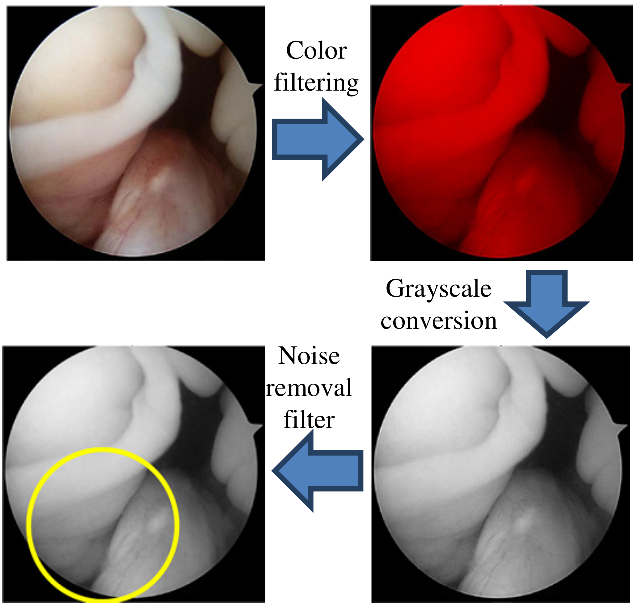


Fig. 7. Extraction of depth information from arthroscopic images

As mentioned in Section 3.2, for each position of the virtual tool a pixel is identified on the image by mapping the instrument device coordinates to image coordinates. The depth for this pixel and its 8-neighbors is then used to calculate the feedback force, as previously proposed in [15-17]. When 3D objects are augmented to the scene for modeling editable tissues in the form of either implicitly defined functions or polygon meshes, collision detection is performed as with fully 3D scenes. The haptic rendering algorithm then branches to common 3D force feedback algorithms for force calculation. Various scenes illustrating simulated surgical procedures are shown in Fig. 9.

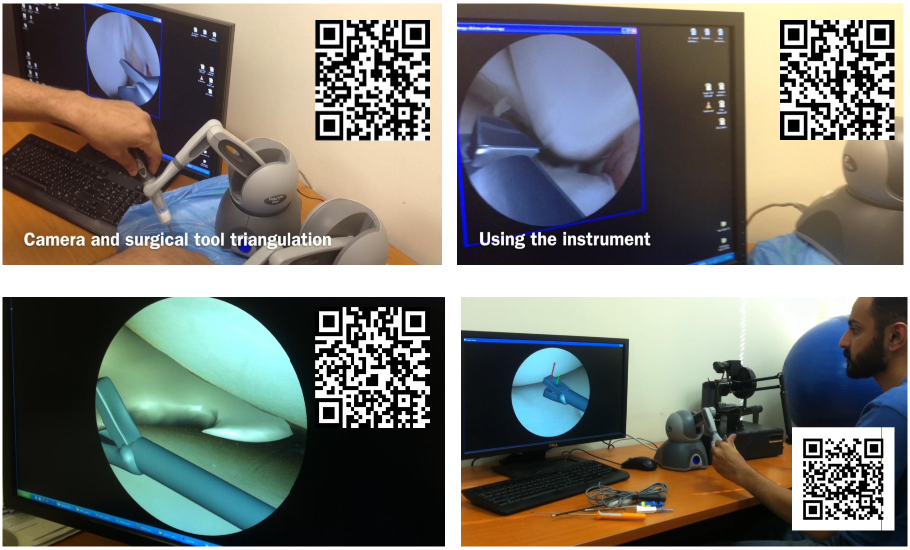


Fig. 8. Various simulations with the QR codes linking to the respective YouTube videos

4 Conclusion

We have presented a way of modeling arthroscopic camera views using haptic devices in hybrid image-based virtual environments. Based on the specifics of the wide-angle oblique-viewing arthroscopic camera, a virtual camera model was presented by formulating the viewing direction as a function of the rotation of the arthroscope axis. Camera position and up vector are two other parameters that define the virtual camera for which physical constraints are proposed to restrict the motion by making extensions of the haptic device. Thus, the haptic device handle mimics the displacements and rotations of the actual surgical camera.

In contrast to the full 3D modeling, in the hybrid image-based virtual environments the problem of simulating views of the virtual camera is rather unusual and more complicated since the challenge is to properly deliver the actual surgical images while navigating the simulated virtual camera. It also involves special methods of making panoramic images from the surgical videos. The challenge in making such panoramas is that common methods and tools cannot be immediately used here since the arthroscopic images are lacking clear feature points needed for automatic image stitching. We therefore proposed our own approach to solve this problem, and it allowed us to achieve very acceptable results that can be used straight away in virtual arthroscopy. The use of a moving hemisphere is proposed to generate undistorted photorealistic views of the surgical area. The proposed approach has been validated by the professional surgeon to be used in the surgical training system.

Acknowledgements. This project is supported by the Singapore MOE Grant "Collaborative Haptic Modeling for Orthopaedic Surgery Training in Cyberspace", Jurong Health Research grant "Photorealistic Virtual Arthroscopy Simulation and Training using Haptic Force Feedback Devices to Achieve Hand-eye Coordination", and by the Russian Foundation for Basic Research Grant 12-07-00157-a. The project is also supported by Fraunhofer IDM@NTU, which is funded by the National Research Foundation and managed through the multi-agency Interactive & Digital Media Programme Office hosted by the Media Development Authority of Singapore.

References

1. Sherman, K.P., Ward, J.W., Wills, D.P., et al.: A portable virtual environment knee arthroscopy training system with objective scoring. Studies in Health Technolgies and Informatics 62, 335–336 (1999)
2. Gibson, S., et al.: Simulating Arthroscopic Knee Surgery using Volumetric Object Representations, Real-Time Volume Rendering and Haptic Feedback. In: Troccaz, J., Grimson, E., Mösges, R. (eds.) CVRMed-MRCAS 1997, CVRMed 1997, and MRCAS 1997. LNCS, vol. 1205, pp. 367–378. Springer, Heidelberg (1997)
3. Bayona, S., Espadero, J., Pastort, L., et al.: A low-cost arthroscopy surgery training system. In: The IASTED International Conference on Visualizatoin, Imaging and Image Processing (VIIP 2003), Benalmadena, Spain (2003)
4. Mabrey, J.D., Gillogly, S.D., Kasser, J.R., et al.: Virtual reality simulation of arthroscopy of the knee. Arthroscopy: The Journal of Arthroscopic & Related Surgery: Official Publication of the Arthroscopy Association of North America and the International Arthroscopy Association 18(6), 28e (2002)
5. Pinto, M.L., Sabater, J.M., Sofrony, J., et al.: Haptic simulator for training of Total Knee Replacement, pp. 221–226.
6. Wang, Y., Xiong, Y., Xu, K., et al.: vKASS: A surgical procedure simulation system for arthroscopic anterior cruciate ligament reconstruction. Computer Animation and Virtual Worlds 24(1), 25–41 (2013)
7. VirtaMed, VirtaMed ArthroS - Virtual Reality training for arthroscopy, VirtaMed AG (2012)
8. TolTech. ToLTech - ArthroSim Arthroscopy Simulator (2013), http://www.toltech.net/medical-simulators/products/arthrosim-arthroscopy-simulator
9. Simbionix, Arthro Mentor - VR training simulator for knee and shoulder arthroscopic procedures, Simbionix Corporation (2012)
10. Simendo, B.V.: SIMENDO Arthroscopy (May 2011), http://www.simendo.eu/products/
11. Yamaguchi, T., et al.: Camera Model and Calibration Procedure for Oblique-Viewing Endoscope. In: Ellis, R.E., Peters, T.M. (eds.) MICCAI 2003. LNCS, vol. 2879, pp. 373–381. Springer, Heidelberg (2003)
12. Chenyu, W., Jaramaz, B.: An easy calibration for oblique-viewing endoscopes. In: Proceedings of the IEEE International Conference on Robotics And Automation, pp. 1424–1429 (2008)

13. De Buck, S., Maes, F., D'Hoore, A., Suetens, P.: Evaluation of a Novel Calibration Technique for Optically Tracked Oblique Laparoscopes. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part I. LNCS, vol. 4791, pp. 467–474. Springer, Heidelberg (2007)
14. Melo, R., Barreto, J.P., Falcao, G.: A New Solution for Camera Calibration and Real-Time Image Distortion Correction in Medical Endoscopy Initial Technical Evaluation. IEEE Transactions on Biomedical Engineering 59(3), 634–644 (2012)
15. Rasool, S., Sourin, A.: Image-driven virtual simulation of arthroscopy. The Visual Computer, pp. 1-12 (2012)
16. Rasool, S., Sourin, A., Kagda, F.: Image-driven haptic simulation of arthroscopic surgery. Studies in Health Technology and Informatics 184, 337–343 (2013)
17. Rasool, S., Sourin, A.: Tangible images. In: SIGGRAPH Asia 2011 Sketches, Hong Kong, China (2011)
18. ArcSoft. ArcSoft Panorama Maker, `http://www.arcsoft.com/panorama-maker/` (September 26, 2013)
19. SoftOptics. panOptica, `http://www.panopticas2.softoptics.co.uk/` (September 26, 2013)
20. Microsoft. Microsoft Research Image Composite Editor (ICE), `http://research.microsoft.com/en-us/um/redmond/groups/ivm/ICE/` (September 26, 2013)
21. Brown, M., Lowe, D.G.: Automatic Panoramic Image Stitching using Invariant Features. Int. J. Comput. Vision 74(1), 59–73 (2007)
22. Bay, H., Tuytelaars, T., Van Gool, L.: SURF: Speeded Up Robust Features. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006, Part I. LNCS, vol. 3951, pp. 404–417. Springer, Heidelberg (2006)
23. Muja, M., Lowe, D.: Fast approximate nearest neighbors with automatic algorithm configuration, pp. 331-340
24. Bauer, J., Sünderhauf, N., Protzel, P.: Comparing Several Implementations of Two Recently Published Feature Detectors. University of Pennsylvania Law Review 154(3), 477 (2006)
25. Juan, L., Gwon, O.: A Comparison of SIFT, PCA-SIFT and SURF. International Journal of Image Processing (IJIP) 3(4), 143–152 (2009)
26. Brown, M., Lowe, D.G.: Automatic Panoramic Image Stitching using Invariant Features. Int. J. of Computer. Vision. 74(1), 59–73 (2007)