10

Intraoperative Visual Guidance and Control Interface for Augmented Reality Robotic Surgery

Rong Wen, Chee-Kong Chui and Kah-Bin Lim
National University of Singapore,
Singapore

1. Introduction

Minimally invasive surgery (MIS) with its small artificial incisions imposed on the patient skin for operation is becoming a preferred surgical treatment comparing with the conventional open surgery. The open surgeries usually create large traumas on the patient bodies to expose their internal anatomy for operational access. The MIS offers its distinct advantages of small injuries, less blood loss and pain and faster recovery during the surgery procedures (Puangmali et al., 2008). However, the indirectly accessed operating causes problems such as restricted vision and difficult hand-eye coordination through the small key holes. The recent development of the modern optics, computer graphics, computer vision and robotics provides possible chances to resolve the problems described above.

Virtual Reality (VR) is a technology providing users an artificial environment to simulate the physical world and objects for displaying, training or gaming etc. Users are allowed to interact with the VR environment through human-computer interface that consists of computer graphic models and various sensors. The major applications of virtual reality in surgery can be divided into three areas: virtual humans for training, virtual telemedicine shared decision environments for training of multiple players and the fusion of virtual humans with real humans for performing surgery (Rosen et al., 1996). VR can create individualized treatment needs and training protocols by providing an environment in which the intensity of feedback and training can be systematically manipulated and enhanced (Anderson et al., 2002; Wang et al., 1999; Cai et al., 2003). VR simulators also allow users to compare their performance with that of their peers (Pearson et al., 2002; Watterson et al., 2002). VR assisted surgical planning facilitates preoperative and postoperative diagnosis with sufficient details of 3-D medical models which are commonly constructed from the computed tomography (CT) or magnetic resonance (MR) images. For example, surgical planning was conducted using a three-dimensional VR based interface to provide a quantitative osteotomy-simulated bone model and prediction of postoperative appearance with photorealistic quality (Xia, et al., 2001). Additionally, there were medical researches based on VR technology specifically conducted to assist surgeons during neurologic, orthopedic, arthroscopic, and urologic surgery (Akay & Marsh, 2001; Burdea et al., 2000; Heng et al., 2004; Kuo et al., 2001; Chui et al., 2006). Recently, there is an increasing

emphasis on data fusion involving integration of virtual patient with real patient as a navigational aid in surgery (Rosen et al., 1996; Wen et al., 2010).

Augmented reality (AR) does not rely solely on artificially created virtual environments but expands the virtual computer-generated models with real world environment. AR is characterized as being a fusion of real and virtual data within a real world environment (Milgram et al., 1994). Medical augmented reality is proposed to create the virtual scenes on the related physical tissues, organs or patient skin to assist surgeons visualizing the internal anatomy or invisible surgical planning such as operational distance, trajectories etc.. Due to the advancement in tracking, visualization and display technology, computer-aided medical procedure based AR solutions was examined in the context of minimally invasive surgery (Navab et al., 2007). AR system has been developed to enhance the endoscopic or laparoscopic view to extend surgeons' visibility of hidden anatomical structures that were beneath the then surgical scene. This augmentation facilitates the surgical operation to avoid risk regions like arteries and nerves. On the other hand the pathology regions are easily to be reached. The needle's insertion happened within the target organ was guided to reach the tumor volume (Konishia et al., 2005; Thoranaghatte et al., 2008; Lee et al., 2010). Thus a more intuitive visual assistance during the surgery procedure is provided for surgeon's response. Integration of laparoscopic and 3D-ultrosound images that provided anatomical information inside the target organs was studied by Konishia et al. (Konishia et al., 2005). With head-mounted display (HMD) based AR interface, numerous studies have demonstrated the potential efficiency of MRI-guided tumour extraction and therapy (Liao et al., 2010). Projector-based visualization system for clinically intraoperative navigation was reported (Hoppe, 2003). Grimson et al. (Grimson et al., 1996) showed its potentials to surgical navigation systems that enhance physical scenes with the augmentation of overlaying virtual internal structures. Another medical application of marking the surgical object with a virtual-pen was based on direct AR interface (Seo et al., 2007). It was proposed to supersede the ink pen which was widely used in the surgical environment. With the virtual-pen, the surgical markings were achieved by directly projecting them onto human body. The introduction of AR to surgical treatment creates a virtual medium between preoperative plan and intraoperative environment.

In order to enhance manual dexterity and accuracy of instrument manipulation, robotic assistance has been introduced in the MIS to solve operation constraint problems. The da Vinci™ Surgical System, a master–slave integrated robot, has been applied for treating a range of surgical conditions improving operational flexibility and accuracy (Guthart & Salisbury, 2000). For the sophisticated minimally invasive operations like coronary artery bypass and mitral valve repair, robot assisted surgery is expected to achieve satisfying clinical outcomes (Devernay et al., 2001). Surgical robot also demonstrated its distinct advantages in speed, accuracy and consistency for the larger liver tumor treatment with RFA. Yang et al. investigated design and implementation of a novel robotic system for RFA considering both kinematics and clinical requirements (Yang et al., 2010). Other advantages offered by the robotic RFA include minimizing radioactive exposure and overcoming visual restriction as well as spatial constraint for minimally invasive approaches like the laparoscopic or percutaneous surgical procedure. Integrating the imaging guidance and the surgical robot assistance into minimally invasive operation brings the preoperative planning and intraoperative manipulation into perspective.

2. Spatial AR interface for robot assisted MIS

Depending on different ways of augmenting the environment, there are three approaches in AR display devices: Optical See Through devices, Video See Through devices, and Direct Augmentations. Direct Augmentations is also called spatial augmented reality (SAR) using projector-camera system to present the virtual information directly in the 3D real world environment (Raskar et al., 2001; Bimber & Raskar, 2004). Based on the above different approaches of augmentations, four kinds of AR platform are popularly used in present surgical simulation and treatment: head-mounted display (HMD) based AR (Cakmakci & Rolland, 2006), semi-transparent mirror based AR (Liao et al., 2010) and projector based AR (Bimber & Raskar, 2004).

Comparing with the other augmentation approaches, SAR offers the distinct advantages that allow users not wearing heavy helmet, overcoming the limited vision area, sharing the AR intraoperative information and communicating freely with their fellow users. The projector-based AR, which is considered to be the most natural way combining the virtual environment and real environment together, is expected to serve as an effective integrated interface to overcome the aforementioned constraints of the MIS. In addition, the surgical robot and interactive hand gesture operation, considered as dynamic components in the real environment, can be easily integrated in this AR interface for robot assisted MIS surgery.

The general work flow using this advanced user interface is illustrated in Fig. 1. It first involves the handling of 3D models that can be derived from the CT images and the preoperative surgical planning. The construction of the static 3D AR environment includes

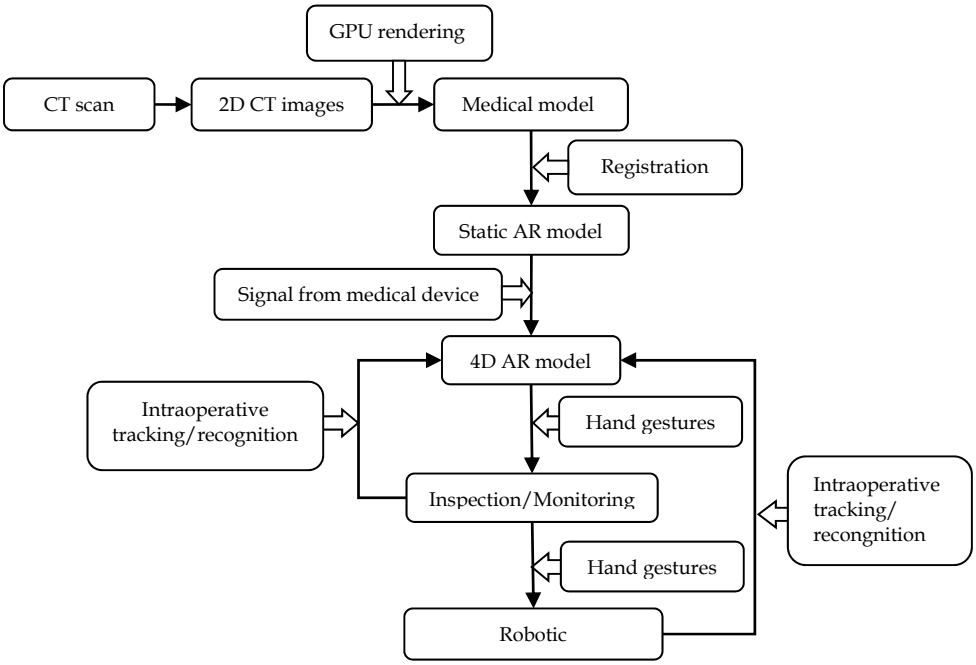

Fig. 1. Work flow of the advanced interface for augmented reality surgery.

registration, geometric and radiometric correction. There is an iterative process with AR real-time model rendering and feedback from the real world environment, including hand gesture interaction with the AR model, surgical robot and the feedback of signal from related medical devices during the surgical operation. When the AR model, displayed over the real organ of the patient by the projector-camera system, receives any "stimulus" from the real environment, there will be changes responding in real time on the model to guide the surgeons' decisions and operation.

3. Registration

Registration is a main challenge in AR technology (Pablo et al., 2010). Correct and consistent registration method plays an important role in estimating the position of either the virtual objects relative to the object in the real world or of the virtual objects relative to a common coordinate system. A number of registration methods have been developed ranging from analysis of the real world environment to evaluations of intentionally introduced fiducials (Massoptier & Casciaro, 2008; Teber et al., 2009).

Marker based methods are commonly used to aid intraoperative registration in robot assisted surgeries (Yang et al., 2010). A marker based registration method for projector-camera system is introduced in this chapter. With the markers attached onto the patient's skin, spatial relationship between the markers and target organ can be derived from CT image of the patient. As shown in Fig. 2, three markers M_1, M_2, M_3, are attached onto the patient body to construct a space reference fame T. This marker based frame is used to construct the spatial relationship that connects the stereovision frame, world coordinate and CT image frame. The spatial relationship enables the virtual model displayed by the projector-camera system to register with the target organ within the patient body. The two steps to overlay a virtual model of the target organ precisely on the patient with correct position, scale and angle in the projector-camera system are described as follows.

The first step involves establishing the spatial relationship between the markers M_i (i=1, 2, 3) and the target organ O. CT imaging of the patient with three markers attached is performed prior to the surgery. Appropriate markers deployment is important. The markers should facilitate the registration and avoid inconveniencing the surgery. Their deployment is dependent on the following considerations. Firstly, the surgical tools' motion should not be obstructed by the markers. The surgical path or motion is given a higher priority in preoperative planning compared to that of markers' deployment. Since the markers are used primarily for constructing the space reference frame T, this can be achieved as long as the stereovision system can detect the markers during registration. Preoperative planning including obstruction prevention will be discussed in section 5.1. Secondly, the markers should be placed beyond the expected projection region to avoid violating the projected virtual model.

For the second step, the initial calibration of projector-camera system has been described by Bimber & Raskar (Bimber & Raskar, 2004). It establishes the correspondence between the pixels in camera and projector as a lookup table. With the stereo vision system, the reference frame T consisted of three markers is achieved from the estimation of space coordinate of the three markers. As mentioned above the spatial relationship between the markers and the target organ based on the image coordinate is derived from the CT images. In this case, the spatial relationship between target organ and the projected virtual model is established as Eq. (1):

$$R_0^T = R_{m_i}^T \times R_o^{m_i} \ . \tag{1}$$

Since the coordinate of the top-left corner L_2 of the projected image is easily detected, the position of L_2 can be adjusted to its expected coordinate. Assume that the L_1 and L_3 are the top-left corner of the source image and the captured image respectively, and $L_1'(x, y)$ and $L_3'(x, y)$ are the expected position of the top-left corner of the source image and the captured image respectively. The expected coordinates of the projected image can be acquired by calculating the spatial relationship with the homography between the camera and the world coordinate. The relationship between $L_1'(x, y)$ and $L_3'(x, y)$ is shown as Eq. (2) and Eq. (3):

$$L_3 = \mathbf{M_A} L_3', \tag{2}$$

$$L_1' = \mathbf{H_p H_c^{-1}} L_3', \tag{3}$$

where $\mathbf{M_A}$ is a affine transformation. A testing projection is conducted to detect the differences between L_3 and L'_3. With the piecewise linear method described above, $\mathbf{M_A}$ and $\mathbf{H_p H_c^{-1}}$ can be replaced by the $\mathbf{M_{A_i}}$ and $\mathbf{H_{p_i} H_{c_i}^{-1}}$ for corresponding piecewise regions.

Suppose that the patient is lying on a surgical bed and the camera is placed along the vertical axis from the target organ relative to the surgical bed. A virtual organ model with the same size and angle located as the real one is expected to be projected on the patient body. This bird-eye view of the image of the virtual model can be derived from the CT images. Finally, the geometric correction for the projected image is conducted to register the virtual model with the real target organ.

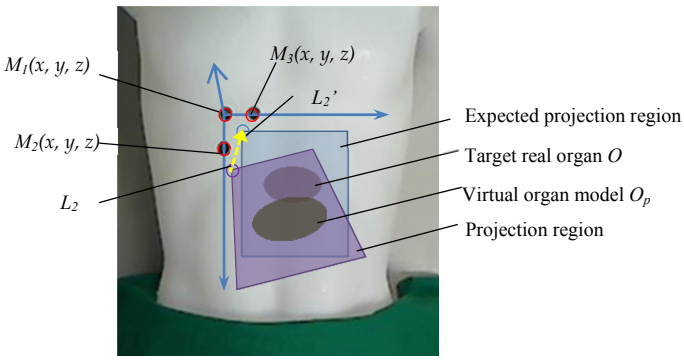

Fig. 2. Marker based registration for SAR (M_1, M_2, M_3 are three markers attached onto the phantom's body.)

4. Correction for projection

In order to create an overall projected image that appears correct and seamless on arbitrary surfaces in the real world, geometric correction plays an important role in construction of the SAR environment. This task is challenging for projector-based displays because of the following problems:

1. Perspective distortion arising from projection.
2. Geometric distortion arising from lens distortion, arbitrary physical surfaces, etc..
3. Real-time geometric correction. This is a challenging problem and it may occur during the dynamic projection in spatial space in which screen surface is changing in real-time because of variance of projector's location or angle.

In addition, the display quality of the projected image is modulated by the colored and textured arbitrary surface sacrificing its photometric quality. Human eyes are sensitive to this artifact. Radiometric correction can be used to relax this restriction with the SAR system by compensating the spatially varying reflectance.

4.1 Geometric correction

Assume the screen surface is planar and projection light is perpendicularly to the surface. The projection on the screen surface appears hardly distorted to the observers in front of the screen. When an image is projected on a planar screen surface that is not perpendicular to it, the projective distortion corrupts the projected image. When an image is projected onto a screen surface with irregular unevenness, geometric factor of the screen surfaces also affects distortion of the projected image.

For a planar projection, homography is used to represent the transformation from the source image to the projected image in the world coordinate, from the projected image to the captured image in the camera as Eq. (4) and Eq. (5):

$$\begin{pmatrix} p_x \\ p_y \\ 1 \end{pmatrix} = \mathbf{H_p} \begin{pmatrix} w_x \\ w_y \\ 1 \end{pmatrix}, \tag{4}$$

$$\begin{pmatrix} c_x \\ c_y \\ 1 \end{pmatrix} = \mathbf{H_c} \begin{pmatrix} w_x \\ w_y \\ 1 \end{pmatrix}, \tag{5}$$

$(w_x, w_y, 1)^T$ is an image point projected on a planar surface. $(p_x, p_y, 1)^T$ and $(c_x, c_y, 1)^T$ are the corresponding points located in the projector's and camera's respectively. From H_p and H_c, a relation from the source image to the captured image can be established in Eq. (6):

$$\begin{pmatrix} p_x \\ p_y \\ 1 \end{pmatrix} = \mathbf{H_p H_c^{-1}} \begin{pmatrix} c_x \\ c_y \\ 1 \end{pmatrix}, \tag{6}$$

The geometric correction process is illustrated in Fig. 3. There are two types of correction models: viewer dependent and viewer independent. For viewer-dependent correction, the source image should be pre-wrapped as $T^{-1}S$, where S is the source image and T is the mapping function between the source image and the captured image. For viewer-independent correction, the source image should be wrapped as $P^{-1}S$, where P is the mapping function between the source image and the projected image. Eq. (7) and Eq. (8) illustrate the correction process**:**

$$PC\left(T^{-1}S\right) = PC\left(C^{-1}P^{-1}\right)S = S\,, \tag{7}$$

$$P\left(P^{-1}S\right) = P\left(CT^{-1}\right)S = S\,. \tag{8}$$

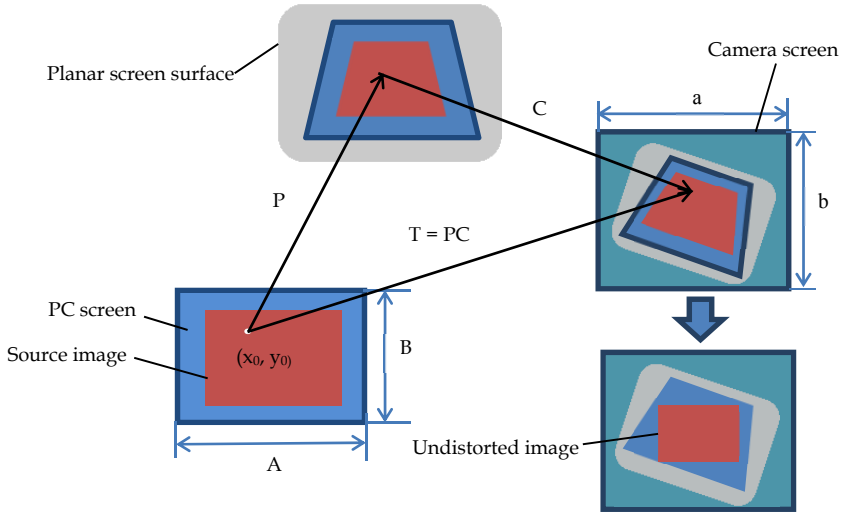

Fig. 3. Viewer-dependent geometric correction.

Although Eq. (6) is not directly applicable when the screen is non-flat, the function $\mathbf{H_p H_c^{-1}}$ can be found easily. An idea of replacing $\mathbf{H_p H_c^{-1}}$ by piecewise linear mapping function that relates the camera and the projector is proposed in this research. With the piecewise linear method, the source image can be marked piecewise by defined feature points. Consequently, the source image can be operated "piece" by "piece", in which the correction of projected image on non-planar screen surface is approximated by connecting all these piece regions. Quadrilateral piece is used as unit region to compute the homography between source image and projected image, captured image and source image. The region connection could be compensated by bilinear interpolation.

The point-correspondence between the image pairs could be acquired from the result of registration. For every piece region, four feature points are established with the respective relationship defined by Eq. (9), (10) and (11). i is the ith piecewise region in the source image, projected image and the captured image. The other parameters have the same meaning as that of the equation (4), (5), (6).

$$\begin{pmatrix} p_{i_x} \\ p_{i_y} \\ 1 \end{pmatrix} = \mathbf{H}_{\mathbf{p}_i} \begin{pmatrix} w_{i_x} \\ w_{i_y} \\ 1 \end{pmatrix}, \tag{9}$$

$$\begin{pmatrix} c_{i_x} \\ c_{i_y} \\ 1 \end{pmatrix} = \mathbf{H}_{\mathbf{c}_i} \begin{pmatrix} w_{i_x} \\ w_{i_y} \\ 1 \end{pmatrix}, \tag{10}$$

$$\begin{pmatrix} p_{i_x} \\ p_{i_y} \\ 1 \end{pmatrix} = \mathbf{H}_{\mathbf{p}_i} \mathbf{H}_{\mathbf{c}_i}^{-1} \begin{pmatrix} c_{i_x} \\ c_{i_y} \\ 1 \end{pmatrix} \tag{11}$$

An example of piecewise geometric correction for chessboard projection is shown as Fig. 4. From Fig. 4, we can observe that the number and location of the selected feature points play important role on the final result.

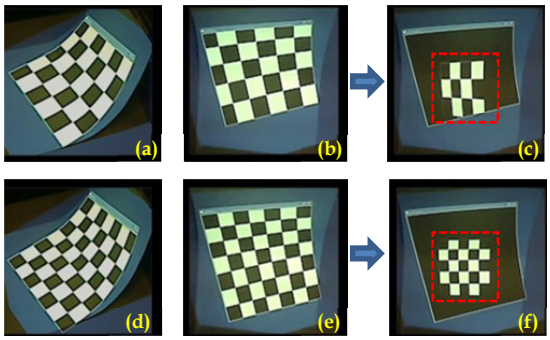

Fig. 4. Geometric correction on the curved surface with a chessboard pattern ((a) and (d) are side views. (b) and (e) are front views. (c) and (f) are the corrected projection for (a) and (d) respectively from the front views).

4.2 Radiometric correction

Radiometry distortion is due to angle of projection, screen texture and the environment lightings. Photometric models of interaction with spectra reflectance between projector and camera were proposed by S. K. Nayar et al. (Nayar et al., 2003) and K. Fujii et al. (Fujii et al., 2005). Based on the photometric models, a practical method to apply radiometric compensation by the projector-camera system is presented in this section. Image $\mathbf{C} = [C_R \;\; C_G \;\; C_B]^T$ acquired by camera from the projected image $\mathbf{P} = [P_R \;\; P_G \;\; P_B]^T$ is affected by surface reflectance and environment lighting. Assume projector and camera each have three color channels (R, G, B). The equation (12) illustrates the radiometric transformation from projector to camera via the screen surface based on the above photometric models:

$$\mathbf{C} = \mathbf{F}(\mathbf{E} + \mathbf{MP}), \tag{12}$$

where :

$$\begin{bmatrix} C_R \\ C_G \\ C_B \end{bmatrix} = \begin{bmatrix} F_R E_R \\ F_G E_G \\ F_B E_B \end{bmatrix} + \begin{bmatrix} F_R M_{RR} & M_{RG} & M_{RB} \\ M_{GR} & F_G M_{GG} & M_{GB} \\ M_{BR} & M_{BG} & F_B M_{BB} \end{bmatrix} \begin{bmatrix} P_R \\ P_G \\ P_B \end{bmatrix}.$$

The vector **F** is the reflectance of surface. **E** is the contribution of the environment lighting to the captured image. The color mixing matrix **M** describes the radiometric interaction between the projector and the camera.

In comparison with the brightness of environment lighting, projector is considerably brighter when performing the projection on the textured and colored surface (Fujii et al., 2005). The color mixing matrix **M** can be derived by projecting two gray images with gray values of 255 and 0 respectively for monotonic radiometric calibration. To compute the **M**

for the color case, normalization of the elements in the **M** is conducted to eliminate the influence of the surface reflectance. The diagonal elements in the **M** are evaluated to unity, $M_{dd} = 1$(d = R, G, B). By comparing of images with differences in one of the three channels, the other elements in the M_{ij} are acquired for the corresponding channel:

$$M_{Rk_1} = \Delta C_{k_1} / \Delta C_R \quad (k_1 = G, B), \tag{13}$$

$$M_{Gk_2} = \Delta C_{k_2} / \Delta C_G \quad (k_2 = R, B), \tag{14}$$

$$M_{Bk_3} = \Delta C_{k3} / \Delta C_B \quad (k_3 = R,G). \tag{15}$$

Based on the piecewise linear method described in the section A, the matrix **F**, **E** and **M** can be replaced by $\mathbf{F}_i$, $\mathbf{E}_i$ and $\mathbf{M}_i$ as (9), (10) and (11).

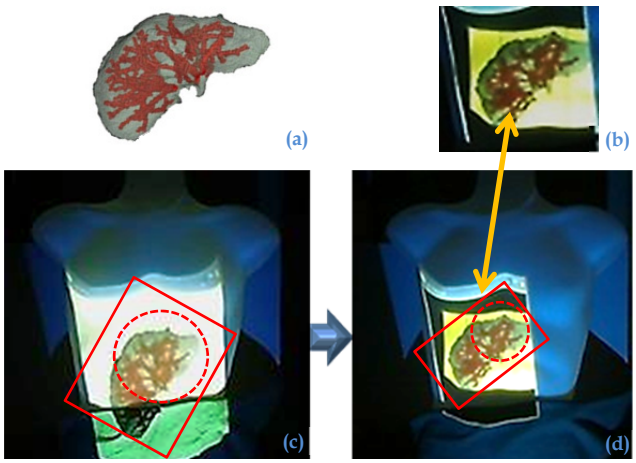

Fig. 5. Projected correction for Augmented Reality environment construction ((a) is a source image of liver model. (c) is the distorted image projected on a phantom body. (b) and (d) are the corrected images.)

5. Intraoperative surgical intervention based on the AR interface

5.1 Therapeutic model design

The dynamic augmented virtual model displayed on the AR interface during surgery is derived from the therapeutic model developed in the preoperative phase. The preoperative plan is based on medical diagnosis, and includes the therapeutic model and surgical planning. The therapeutic model is a stereoscopic cyber-graphic model reconstructed from CT images with relevant organs and tissues segmented. Besides the three dimensional profile of the target tumour and blood vessels, other important circumjacent tissues and organs around the interested region are included in the model as specified by the surgeons. Additionally, markers deployment should also be examined for this therapeutic model design to make sure they are beyond the surgical working path.

For the therapeutic model design based on the robot assisted surgery, optimization algorithms are usually applied to enhance the efficiency of robotic implementation and to decrease complexity during the surgery. For example, an ablation model with a "Voxel Growing" algorithm (Yang et al., 2009) is applied for a prototyped robot assisted radiofrequency ablation (RFA) surgery. This algorithm automatically adapts the different tumour's irregular 3-D profile to produce the ablation points. In order to further reduce the number of required ablation and minimize burning of healthy tissue, the "grown region" is optimally aligned to best fit the geometry of the target region. The input for the ablation model is the surface data of the tumour, which are acquired from CT scan in advance. Surgical path planning is another important consideration for the preoperative planning. This method involves the construction of spherical elements defined by inscribed cubes. Each sphere represents the ablation area, and its inscribed cube represents a voxel element. Voxel elements propagate layer by layer along defined directions according to its surface edges of the present layer, covers the entire tumour area coinciding with its contour. As voxel elements grow, the coordinates of centres of every voxel elements are produced and being assigned as the ablation coordinates. After the therapeutic model and the surgical plan are prepared, the model is projected onto the patient in the spatial AR environment.

5.2 Intraoperative tracking and hand gesture recognition

Intraoperative tracking is to establish the real-time spatial awareness of medical instrument navigation and hand motion necessary for visual guidance and intraoperative control during the surgery. Intraoperative tracking of the operating trajectory of the medical instrument provides surgeons real-time feedback of the operation procedure. This real-time feedback information is usually analysed and dealt with in consideration of the intraoperative process as well as the previous medical preplanning. For example in RFA surgery, when the RFA needle is inserted into the patient body, the distal end of needle is tracked to detect the needle's trajectory. The AR environment is updated in real–time producing a virtual needle part, which is projected on the patient surface, following the needle's treatment plan. The surgeon can visualize the operation within the patient's body and surgical processes.

Numerous tracking methods have been developed for surgical environment including vision based tracking and magnetics based tracking. (Dutkiewicz et al., 2004; Felfoul et al., 2008). After taking into consideration the consistent surgical luminance, distinct background content and usability, a vision-based approach based on Continuously Adaptive Mean-Shift (CamShift) algorithm (Ling et al., 2010) is adopted for intraoperative tracking. CamShift algorithm is built upon the Mean-Shift algorithm to track objects based on probability distribution (Wen et al., 2010; Mustafa & Kalajdziski, 2007). Back projection is a way to find the extent to which the pixel distribution of a region coincides with a histogram model. It can be used to compute the probability distribution image from the original image with the colour histogram. The pixel values in the input image are scaled using the values in the corresponding bin of the color histogram and stored in the probability distribution image. The function computes the probability of each element value according to the empirical probability distribution represented by the histogram. If C is the color of the pixel and F is the probability that a pixel belongs to the object being tracked, then the probability map gives $P(C|F)$. The probability distribution image is used in the algorithm for tracking.

The procedure of the Mean-Shift tracking algorithm is described as follows::

1. An initial location of a search window is chosen.
2. Compute the mean location or centroid in the window based on the probability distribution image. The centroid (x_c, y_c) is calculated as follows:
   Firstly, the zeroth and first moments of the image are calculated as:

$$M_{20} = \sum_x \sum_y x^2 I(x,y), \tag{16}$$

$$M_{02} = \sum_x \sum_y y^2 I(x,y), \tag{17}$$

$$M_{11} = \sum_x \sum_y xy I(x,y), \tag{18}$$

   where *I(x, y)* is the probability value at coordinate (x, y) in the image, ranging over the entire search window. The coordinates of the centroid are then computed from the moments as Eq. (19) and Eq. (20).

$$x_c = M_{10} / M_{00}, \tag{19}$$

$$y_c = M_{01} / M_{00}, \tag{20}$$

3. Move the center of the window to the location derived from the previous step.
4. Return to Step 2 until convergence is achieved.

The CamShift algorithm also has the ability to update the size of the search window to fit the object being tracked as the frames change (Ling et al., 2010).

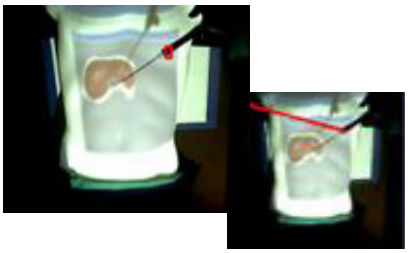

Fig. 6. RFA needle is tracked based on an AR environment.

For intraoperative interaction, hand gesture is adopted to control display of the AR model and robotic device. Surgeon's hand gestures are detected by a stereovision system over or beside the AR environment. This interface has advantage on sterilization which is an important requirement in operating room. The vision based hand gesture can be integrated into the AR system using projector-camera system or conventional tracking device. The AR model projected on the patient surface can be scaled and rotated using hand gestures. Additionally, hand gestures can also be used to guide robotic execution.

In order to enhance the speed and accuracy of hand gesture recognition, an innovative 3D model based method is introduced. This method is based on the analysis of model features

produced by a space palm plan and 3D convex hull points. Users are allowed to pose the defined hand gestures without wearing markers or data glove. Due to human hand's flexibility, stereovision system is used to resolve ambiguity on recognizing the hand features.

As shown in Fig. 7, the 3D model is constructed including the space palm plan which is relatively stable, and a brace model of five fingers with the key joints and finger tips. Before starting the hand gestures recognition, the 3D hand model is initialized by characterizing the model with the specific hand parameters such as palm size, fingers layout of different users. This initialization can be achieved by directly measuring the key parameters of the user's hand under the vision system with a defined pose. These key parameters include hand length and width, the ratio between the palm and the fingers, and the palm's location relative to the stereovision system. Palm plan can be acquired by regression of the hand palm coordinates shown as Fig.7 (b). There are several possible ways of tracking the user's hands (Mahmoudi & Parviz, 2006). In consideration of the background of hand operation based on this AR interface is covered by the surgical cloth which is green in color. CamShift based method used for medical instrument tracking can also be integrated with hand tracking. When the hand tracking is implemented, the palm plan is invoked to change its pose in the space coordinate. Following with the hand tracking, the 3D hand model is attached to the user's hand wherever the hand moves. In this case the parameters of the 3D hand model are analysed in real-time and the hand gesture recognition is achieved in real-time.

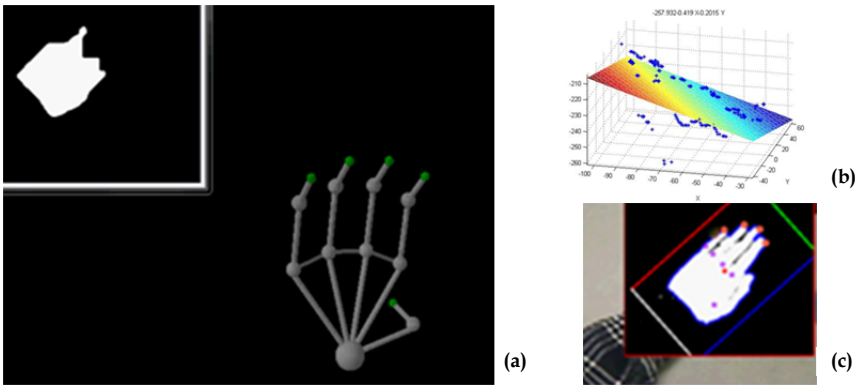

Fig. 7. 3D model based hand gesture recognition.

5.3 Robotic implementation

While the introduction of AR in computer aided surgery enhances visualization of preoperative surgical plan, it alone does not necessarily translate to precision and accuracy in intraoperative execution. In the case of minimally invasive surgery, the surgeon is not only deprived of direct visual information but also subjected to dexterous constraints in the manoeuvring of laparoscopic tools or percutaneous ablation devices. Performing such operation manually even with AR guidance can be non-trivial and often expose to uncertainties and inconsistent outcomes. For instance, the challenges in large tumor treatment with percutaneous ablation technique have been discussed (Abolhassani et al.,2007; Choi, et al., 2000; Goldberg, et al., 1998). This lack of intraoperative consistency poses a serious bottleneck for the effectiveness surgical outcome. Robotic assistance

provides the necessary augmentation in dexterity to provide precision and accuracy even in space constraint situation (Hyosig & Wen, 2001; Kuo & Dai, 2009; Stoianovici et al., 1998; Taylor & Stoianovici, 2003; Taylor, 2008). The Transcutaneous Robot-assisted Ablation-device Insertion Navigation System (TRAINS) is developed for ablation treatment in liver tumor (Chang et al., 2011; Yang et al., 2010). Technical details and demonstrate the implementation of robotic modules for AR robotic system are presented as follows.

The integration of the robotic modules requires establishment of the spatial relationship between the multi-body system and surgical field. It begins with the construction of a kinematic model for the manipulator and frame assignments to the various entities for the AR applications. Frame assignment of the multi-body robotic mechanism can be established based on the Denavit-Hartenberg (D-H) convention described by (Fu, 1987). Fig. 8 (left) illustrates the D-H frame assignments for TRAINS prototype. The joint variables and link parameters tabulated in Table 1.

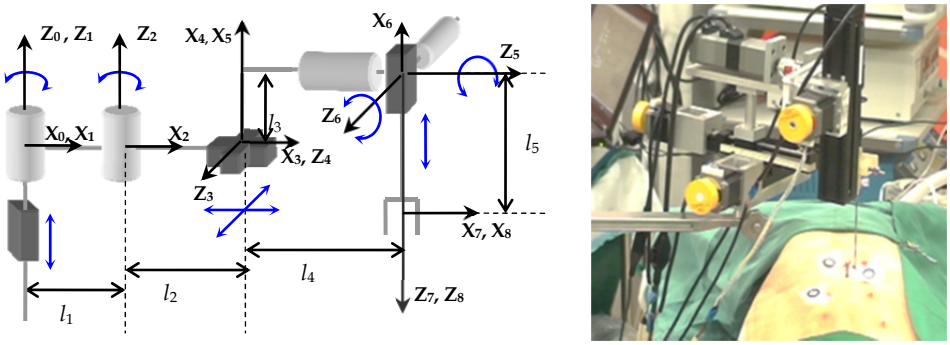

Fig. 8. Denavit-Hartenberg frame assignments for TRAINS prototype (left); Robotic needle insertion for ablation treatment (right).

	θ	D	a	a
${}^{0}T_1$	0	q_1	0	0
${}^{1}T_2$	q_2	0	l_1	0
${}^{2}T_3$	q_3	0	l_2	90°
${}^{3}T_4$	90°	q_4	0	90°
${}^{4}T_5$	0	q_5	0	0
${}^{5}T_6$	q_6	l_4	-90°	-90°
${}^{6}T_7$	q_7	0	0	90°
${}^{7}T_8$	0	q_8+l_5	0	0

Table 1. Denavit-Hartenberg table for manipulator (q_i stands for ith joint variable).

The kinematic model can subsequently be expressed in homogenous transformation matrix as shown in Eq. (21). which represents the orientation and position of the end effector as a function of joint variables from their respective axis of control.

$$ {}^{o}T_c = \begin{bmatrix} {}^{0}R_8 & {}^{0}P_8 \\ \bar{0} & \bar{1} \end{bmatrix}, \tag{21} $$

where

$$ {}^{o}R_8 = \begin{bmatrix} s_{23}s_6c_7 - c_{23}s_7 & s_{23}c_6 & s_{23}s_6s_7 + c_{23}c_7 \\ -c_{23}s_6c_7 - s_{23}s_7 & c_{23}c_6 & -c_{23}s_6s_7 + s_{23}c_7 \\ c_6c_7 & -s_6 & c_6s_7 \end{bmatrix}, \tag{22} $$

$$ {}^{o}P_8 = \begin{bmatrix} s_{23}s_6s_7\left(q_8 + l_5\right) + c_{23}[c_7\left(q_8 + l_5\right) + l_4] + q_5c_{23} + q_4s_{23} + l_1c_2 + l_2c_{23} \\ -c_{23}s_6s_7\left(q_8 + l_5\right) + s_{23}[c_7\left(q_8 + l_5\right) + l_4] + q_5c_{23} - q_4s_{23} + l_1s_2 + l_2s_{23} \\ c_6c_7\left(q_8 + l_5\right)l_3q_1 \end{bmatrix}, \tag{23} $$

where s_j=sin (q_i), c_{ij}=cos(q_i), sij=sin (q_i+q_j) and c_{ij}=cos(q_i+q_j).

Direct vision is usually absence in minimally invasive surgery or percutaneous interventional procedure. Hence registration between diagnostic medical images, surgical field and its entities has to be done. Preoperative image coordinates can be readily transformed to spatial coordinates based on the system configuration of the imaging module. Intraoperative spatial information however, requires a tracking and navigation system. TRAINS uses a vision-based system for tracking and navigation purpose. Light emitting fiducial markers are used for the registration of entities in the real world to the interventional plan. Fig. 8 (right) depicts the robotic setup in a porcine surgery.

6. Discussion

The use of the intraoperative visual guidance and AR control interface in surgical procedures provide surgeons with a very exciting viewable integrated information platform. With this platform, the surgeon can concurrently conduct an intraoperative diagnosis, surgical operation and collection of patient's medical information on a single interactive augmented interface. This augmented reality based interactive interface allows surgeons to not only experience a new type of technology that extends their visibility without looking at any other screen or wearing any head-mounted device during the surgery, but also perform the operation with the combination of preoperative and intraoperative medical information. With regards to the surgical robot, not only does it contribute to higher accuracy of the surgical operation and optimizes the procedural flow, it is also easy for this intraoperative AR interface to incorporate the surgical robot assistance into the surgery, which includes the display of the robotic operation inside the patient and interactive control by the surgeons.

Although interactive spatial AR interfaces have been used for some applications such as games and advertisements, the intraoperative visual guidance and interactive AR interface for surgery is still a challenging problem. The reasons for this are as followed. First of all, an accurate display of the human organs and tissues is essential in surgery. The visual corrections, including geometric corrections of a projected virtual organ model, are important steps in constructing the spatial augmented reality environment, especially when projecting on an arbitrary surface. Other than the influence of irregular geometric surfaces, the projection is also affected by dynamic motions in the AR environment such as the patient's aspiration, unexpected surface motions or occlusions etc. Secondly, registration of an AR model with the real organ is especially difficult because of the non-rigid soft tissue. Thirdly, in usability tests with the medical tool, the ablation needle is as assumed to be a rigid needle without bending during the insertion. However, this assumption may not always be true. Therefore, other than

vision-based tracking technology, other types of intraoperative feedback such as real-time CT scanning etc. are needed to generate more reliable surgical information. Additionally, in order to realize a dynamic observational perspective from any location around the AR environment, more projectors and cameras are required.

7. Conclusion and future work

The intraoperative visual guidance and control AR interface is an attempt to integrate the preoperative diagnostic information with real-time augmented visual information produced by the patient's reactions, surgeon's decisions, and robotic implementation. This interface provides surgeons with a mean to extend a person's visual limitations during the MIS and to engage interactively during the surgical robot assisted surgery. Augmented reality technology has the potential as a therapeutic intervention which combines the virtual augmented physical model with locally detected real-time medical information including geometric variance and respiration rate. The interface is coupled with the augmented physical model, surgeon's operation, and robotic implementation through vision based tracking and hand gesture recognition.

This intraoperative AR interface will benefit from a stereoscopic vision system with higher accuracy and higher frame rate which plays an important factor in vision based tracking and recognition. While the current tracking algorithm worked satisfactorily at the workstation used for testing, users still had problems with robustness when operating in an environment with a colourful background. In addition, this interface could be expanded to assimilate with more medical devices and equipments such as ultrasound that would integrate more real-time augmented medical information. In this scenario, the interface is a powerful integrated medical AR platform centralized in the network of all the other medical devices and equipment in the operating room.

8. References

Abolhassani, N.; Patel, R. & Moallem, M. (2007). Needle insertion into soft tissue: A survey. *Medical Engineering & Physics*, Vol.29, No.4, pp. 413-431, ISSN 1350-4533.

Akay, M. & Marsh, A. (2001). Neuro orthopedic rehabilitation and disability solutions using virtual reality technology information technologies in medicine. *Rehabilitation and Treatment*, Vol. 168, No. 2, pp.1–18, ISSN 1530-437X.

Anderson, J.H.; Chui, C.K.; Cai, Y.; Wang, Y.P.; Li, Z.R.; Ma, X.; Nowinski, W.L.; Solaiyappan, M.; Murphy, K. & Gailloud, P. (2002). Virtual reality training in interventional radiology: The Johns Hopkins and Kent Ridge Digital Laboratory Experience. *Seminars in Interventional Radiology*, Vol. 19, No. 2, pp. 179-185, ISSN 0739-9529.

Bimber, O. & Raskar, R. (2004). Spatial Augmented Reality: merging real and virtual worlds. ISBN 9781568812304, Wellesley, MA : A K Peters, 2004.

Burdea, G.; Popescu, V.; Hentz, V. & Colbert, K. (2000). Virtual reality-based orthopedic telerehabilitation. *Rehabilitation Engineering, IEEE Transactions*, Vol.8, No.3, pp.430–432, ISSN 1063-6528.

Cai, Y.; Chui, C.K.; Ye, X. & Anderson, J. (2003). VR simulated training for less invasive vascular intervention. *Computers & Graphics*, Vol. 27, No.2, pp. 215-221.

Cakmakci, O. & Rolland, J.(2006). Head-worn displays: a review. *Display Technology*, Vol. 2, No.3, pp.199-216, ISSN 1551-319x.

Chang, S.K.Y.; Hlaing, W.W.; Yang, L., & Chui, C.K. (2011). Current technology in navigation and robotics for liver tumours ablation. *Ann Acad Med Singapore*, Vol.40, No.5, pp.231-236, May, 2011.

Choi, D.; Lim, H.K.; Kim, S.H.; Lee, W.J.; Jang, H.J. & Lee, J.Y. (2000). Hepatocellular carcinoma treated with percutaneous radio-frequency ablation: usefulness of power doppler us with a microbubble contrast agent in evaluating therapeutic response. *Preliminary Results1. Radiology*, Vol. 217, No.2, pp.558-563.

Chui, C.K.; Ong, JSK.; Lian, Z.Y.; Yan, C.H.; Wang, Z.; Ong, SH.; Teo, JCM.; Zhang, J.; Wang, SC.; Wong, HK.; Teo, C.L. & Teoh, S. H. (2006) Haptics in computer mediated simulation: training in vertebroplasty surgery. *Simulation & Gaming*, Vol. 47, No. 4, pp. 438-451, ISSN 1552-826X.

Devernay, F.; Mourgues, F.& Coste-Maniere, E. (2001). Towards endoscopic augmented reality for robotically assisted minimally invasive cardiac surgery. *Proceedings of the International Workshop on Medical Imaging and Augmented Reality (MIAR '01)*, pp.16-20, ISBN 0-7695-1113-9, Shatin, N.T., Hong Kong, Jun. 2001.

Dutkiewicz, P.; Kielczewski, M. & Kowalski, M. (2004). Visual tracking of surgical tools for laparoscopic surgery. *Proceedings of the Fourth International Workshop on Robot Motion and Control,* pp.23-28 , ISBN 83-7143-272-0, Zielona Gora Univ., Poland, June, 2004.

Felfoul, O.; Mathieu, J.-B.; Beaudoin, G.; Martel, S. (2008). In vivo mr-tracking based on magnetic signature selective excitation. *IEEE Transactions on Medical Imaging,* Vol.27 , No.1, pp. 28 – 35, ISSN 0278-0062.

Fu, K.S. (1987). Robotics: control, sensing, vision, and intelligence, Vol. 280227, McGraw-Hill, New York, 1987.

Fujii, K.; Grossberg, M.D. & Nayar, S.K. (2005). A projector-camera system with real-time photometric adaptation for dynamic environments. *Computer Vision and Pattern Recognition,* Vol. 2, pp.1180, ISSN 1063-6919.

Goldberg, S.N.; Gazelle, G.S.; Solbiati, L.; Livraghi, T.; Tanabe, K.K. & Hahn, P.F.(1998). Ablation of liver tumors using percutaneous RF therapy. *AJR Am J Roentgenol*, Vol. 170, No.4, pp.1023-1028 , ISSN 1023-1028.

Grimson, W.E.L.; L. T.; Wells, W.M.; Ettinger, G.J.; White, S.J. & Kikinis, R. (1996). An automatic registration method for frameless stereotaxy, image guided surgery, and enhance reality visualization. *IEEE Transactions on Medical Imaging*, (April 1996), Vol. 15, No. 2, pp.129-140. , ISSN 0278-0062.

Guthart, G.S. & Salisbury, J.K.Jr. (2000). The intuitive™ telesurgery system: Overview and application, *Robotics and Automation*, pp.618-621, ISBN 0-7803-5886-4, San Francisco, CA , USA, Apr.2000.

Heng, P.A.; Cheng, C.Y.; Wong, T.T; Xu, Y.; Chui, Y.P.; Chan, K.M. & Tso, S.K. (2004). A virtual-reality training system for knee arthroscopic surgery. *Information Technology in Biomedicine, IEEE Transactions*, Vol. 8, No.2, pp. 217-227, ISSN 1089-7771.

Hoppe, H.; Eggers, G.; Heurich, T.; Raczkowsky, J.; Marmulla, R.; Worn, H.; Hassfeld, S. & Moctezuma, J.L. (2003). Projector-based visualization for intraoperative navigation: first clinical results. *Proceedings of the 17th International Congress and Exhibition*, pp.771, ISBN 0-444-51731-6, London, Jun. 2003.

Hyosig, K. & Wen, J.T. (2001). Robotic assistants aid surgeons during minimally invasive procedures. *Engineering in Medicine and Biology Magazine*, IEEE, Vol.20, No.1, pp.94-104, ISSN 0739-5175.

Konishia K.; Hashizumeb M.; Nakamotod M.; Kakejib Y.; Yoshinoc I.; Taketomic A.; Satod Y.; Tamurad S. & Maehara Y. (2005). Augmented reality navigation system for endoscopic surgery based on three-dimensional ultrasound and computed

tomography: Application to 20 clinical cases. *CARS 2005: Computer Assisted Radiology and Surgery*, Vol. 1281, pp. 537-542. Berlin, Germany, Jun. 2005.

Kuo, C.H. & Dai, J.S. (2009). Robotics for minimally invasive surgery: a historical review from the perspective of kinematics. *International Symposium on History of Machines and Mechanisms*, ISBN 978-1-4020-9484-2, pp.337-354.

Kuo, R.L.; Delvecchio, F.C.& Preminger, G.M. (2001). Virtual Reality: Current Urologic Applications and Future Developments. *Journal of Endourology*, Vol. 15, No. 1, pp. 117-122, ISSN 0892-7790.

Lee, S.L.; Mirna, L.; Vitiello, V.; Giannarou, S.; Kwok, K.W.; Visentini-Scarzanella, M. & Yang, G.Z. (2010). From medical images to minimally invasive intervention: Computer assistance for robotic surger. *Computerized Medical Imaging and Graphics*, Vol. 34, No. 1, pp. 33-45, ISSN .

Liao, H; Inomata, T; Sakuma, I & Dohi, T (2010). Three-dimensional augmented reality for mri-guided surgery using integral videography auto stereoscopic-image overlay. *IEEE Transactions on Biomedical Engineering*, Vol.57, No.6, pp.1476-1486, ISSN 15582531.

Ling, Y.; Zhang, J. & Xing, J. (2010). Video object tracking based on position prediction guide CAMSHIFT. P*roceeding of Advanced Computer Theory and Engineering (ICACTE)*, pp. V1-159 - V1-164, ISBN 978-1-4244-6539-2, Chengdu, China, Aug. 2010.

Mahmoudi, F.& Parviz, M. (2006). Visual hand tracking algorithms. *Proc. of GMAI 2006, IEEE Computer Society press*, ISBN 0-7695-2604-7, pp.228–232, London, UK, July 2006.

Massoptier, L. & Casciaro, S. (2008). A new fully automatic and robust algorithm for fast segmentation of liver tissue and tumors from CT scans. *European Radiology*, Vol.18, No.8, pp.1658-1665, ISSN 1432-1084.

Milgram, P.; Takemura, H.; Utsumi, A.; & Kishino, F. (1994). Augmentedreality:a class of displays on the reality-virtuality continuum in proc. *Telemanipulator Telepresence Technology*, Vol. 2351, pp.282–292.

Mustafa, B. & Kalajdziski, S. (2007). Spherical mapping based descriptors for 3D object matching. *The International Conference on Computer as a Tool*, pp. 361-368, ISBN 978-1-4244-0813-9, Warsaw, Poland, Sept. 2007.

Navab, N.; Traub, J.; Sielhorst, T.; Feuerstein, M. & Bichlmeier, C. (2007). Action- and workflow-driven augmented reality for computer-aided medical procedure. *Computer Graphics and Applications*, Vol.27, No.5, pp. 10–14, ISSN 0272-1716.

Nayar, S.K.; Peri, H.; Grossberg, M.D. & Belhumeur, P.N. (2003). A Projection System with Radiometric Compensation for Screen Imperfections. *Proceeding of IEEE Int'l Workshop Projector-Camera Systems*, pp.97-108 , ISBN 1077-2626, Nov.2007.

Pablo L.; Wajid A.; Alicia C.; Jordi C.; Jerome D.; Ole J.E.; Adinda F.; Hugo F.; Denis K.; Edvard N.; Eigil S.; Patricia S.G.; Francisco M.S.M.; Dieter S.; Mauro S.; Thomas S.; Jos V.S. & Enrique J.G. (2010). Augmented reality for minimally invasive surgery: overview and some recent advances. *Augmented Reality*, pp.1-26, In: InTech, Available from: http://www.intechopen.com/articles/show/title/augmented-reality-for-minimally-invasive-surgery-overview-and-some-recent-advances.

Pearson, A.M.; Gallagher, A. G.; Rosser, J. C.; & Satava, R. M. (2002). Evaluation of structured and quantitative training methods for teaching intracorporeal knot tying. *Surgical Endoscopy*, Vol. 16, No. 1, pp.130–137, ISSN 1432-2218.

Puangmali, P.; Althoefer, K.; Seneviratne, L.D.; Murphy, D. & Dasgupta, P. (2008). State-of-the-art in force and tactile sensing for minimally invasive surgery. *Sensors Journal*, Vol. 8 , No. 4, pp. 371 – 381, ISSN 1530-437X.

Raskar, R.; Welch, G. & Low, K.L. (2001). Shader Lamps: Animating Real Objects With Image-Based Illumination. *Eurographics Workshop on Rendering Techniques*, ISBN 3-211-83709-4, London, Jun.2001.

Rosen, J.M.; Soltanian, H.; Redett, R.J. & Laub, D.R. (1996). Evolution of virtual reality [Medicine] *Engineering in Medicine and Biology Magazine*, Vol.15 , No.2, pp. 16-22, ISSN 0739-5175.

Seo, B.K.; Lee, M.H.; Park, H.; Park, J.I. & Kim, Y.S. (2007). Direct-projected ar based interactive user interface for medical surgery. *Artificial Reality and Telexistence,* ISBN 0-7695-3056-7, Esbjerg, Jylland, Nov.2007.

Stoianovici, D.; Whitcomb, L.; Anderson, J.; Taylor, R. & Kavoussi, L. (1998). A modular surgical robotic system for image guided percutaneous procedures. *Medical Image Computing and Computer-Assisted Interventation – MICCAI'98,* pp.404-410, ISBN 978-3-540-65136-9, Chicago, USA, October 1998.

Taylor, R.H. & Stoianovici, D. (2003). Medical robotics in computer-integrated surgery. *IEEE Transactions on Robotics and Automation*, Vol. 19, No.5, pp.765-781, ISSN 1042-296X.

Taylor, R.H. (2008). Medical robotics and computer-integrated surgery. *32nd Annual IEEE International Computer Software and Applications*, pp.1-1, ISBN 978-0-7695-3262-2, Turku, Finland, July-August 2008.

Teber, D.; Baumhauer, M.; Simpfendoerfer, T.; Hruza, M.; Klein, J. & Rassweiler, J. (2009). Augmented reality a new tool to improve surgical accuracy during laparoscopic partial nephrectomy. *European Urology Supplements*, Vol. 7, No. 3, (March 2008), pp. 258, ISSN 1569-9056.

Thoranaghatte, R.U.; Giraldez, J.G. & Zheng, G. (2008).Landmark based augmented reality endoscope system for sinus and skull-base surgeries. *Engineering in Medicine and Biology Society*, 2008, ISBN 978-1-4244-1814-5, Vancouver, BC, Aug.2008.

Uranues, S.; Maechler, H.; Bergmann, P.; Huber, S.; Hoebarth, G.; Pfeifer, J.; Rigler, B.; Tscheliessnigg, K.H. & Mischinger, H.J (2002). Early experience with telemanipulative abdominal and cardiac surgery with the Zeus™ Robotic System. *European Surgery*, Vol. 34, No. 3, pp. 190–193, ISSN 1682-4016.

Wang, Y.P.; Chui, C.K.; Lim, H.L.; Cai, Y. & Mak, K.H. (1999). Real-time interactive simulator for percutaneous coronary revascularization procedures. *Computer Aided Surgery,* Vol. 3, No. 5, 1999, ISSN 1097-0150.

Watterson, J.D; Beiko, D.T.; Kuan, J.K.; & Denstedt, J.D. (2002). Randomized prospective blinded study validating acquisition of ureteroscopy skills using computer based virtual reality endourological simulator. *International braz j urol official journal of the Brazilian Society of Urology,* Vol.168, No.5, pp.1928–1932, ISSN 00225347.

Wen, R.; Yang, L.; Chui, C.K.; Lim, K.B. & Chang, S. (2010). Intraoperative visual guidance and control interface for augmented reality robotic surgery. *IEEE International Conference on Control and Automation*, pp.947-952, ISBN 978-1-4244-5195-1, Xiamen, China, Jun. 2010.

Xia, J.; Ip, H.H.S.; Samman, N.; Wong, H.T.F.; Gateno, J.; Wang, Dongfeng; Yeung, R.W.K.; Kot, C.S.B. & Tideman, H. (2001). Three-dimensional virtual-reality surgical planning and soft-tissue prediction for orthognathic surgery. *Information Technology in Biomedicine, IEEE Transactions*, Vol. 5, No. 2, pp. 97- 107, ISSN 1089-7771.

Yang, L.; Chui, C.K. & Chang, S. (2009). Design and Development of an Augmented Reality Robotic System for Large Tumor Ablation. *International Journal of Virtual Reality*, Vol. 8, No.1, pp.27-35, ISSN 1741-1882.

Yang, L.; Wen, R.; Qin, J.; Chui, C.K.; Lim, K.B. & Chang, S.K.Y. (2010). A robotic system for overlapping radiofrequency ablation in large tumor treatment. *IEEE/ASME Transactions on Mechatronics*, Vol. 15, No. 6, pp.887-897, ISSN 1083-4435.