Direct-Projected AR Based Interactive User Interface for Medical Surgery

Byung-Kuk Seo, Moon-Hyun Lee,
Hanhoon Park, and Jong-Il Park
Center for Intelligent Surgery System
Department of Electrical and Computer Engineering, Hanyang University
{nwseoweb, fly4moon, hanuni}@mr.hanyang.ac.kr
jipark@hanyang.ac.kr

Young Soo Kim

Center for Intelligent Surgery System
Department of Neurosurgery, Hanyang University
ksy8498@hanyang.ac.kr

Abstract

In the field of computer aided surgery, augmented reality (AR) technology has been successfully used for enhancing accuracy of surgery and making surgeons convenient by visually assisting them in performing a number of complicated and time-consuming medical operations. However, there are still medical operations that do not receive the benefit of AR technology. As a representative one, surgeons still use an ink pen when they mark surgical targets for scheduling an operation. The ink pen is inconvenient because the mark drawn by the foreign matter is not easily modified or deleted. And the ink pen is also unlikely to be sanitary. In this paper, we propose an interactive user interface based on direct-projected augmented reality (DirectAR) technology for handling all these problems with the ink pen and its validity is shown in experimental results.

1. Introduction

DirectAR is a technology that virtual information is directly displayed onto real scenes in 3-D. DirectAR is able to overcome technological limitations (lack in resolution and small field of view) and ergonomic shortcomings (heavy optics and discomfort incidental to wearing) of traditional displays for AR such as video see-through and optical see-through head-mounted displays. As projectors have been recently widespread, various techniques for DirectAR have been developed [1], which include techniques for resolving fundamental problems of projection-based systems: projection images are distorted geometrically and radiometrically on dynamic (or non-planar) and colourful textured surfaces [2, 3, 4, 5, 6, 7, 8, 9].

DirectAR technology has been mainly applied to interactive systems. Rekimoto *et al.* [10] and Shen *et al.* [11] proposed collaborative augmented space for effective interaction between multi-users. To enhance interaction between users and systems, Kjeldsen *et al.* [12] detected pre-defined regions and Oka *et al.* [13] recognized fingertips and gestures in real-time. Zaeh *et al.* [14] also proposed an interactive laser-projection system where tool trajectories and target coordinates are interactively visualized and manipulated in the industrial robot's environment.

Grimson *et al.* [15] showed potentials of DirectAR technology in medical fields by applying it to surgical navigation systems and by enhancing reality visualization of internal anatomical structures overlaid onto surgical targets. Hoppe *et al.* [16] showed the clinical testing results of their projector-based visualization system for intra-operative navigation. Most recently, Schonfelder *et al.* [17] proposed a system that visualizes trocar positions for abdominal minimally invasive surgery by using DirectAR. As another medical application, we propose a DirectAR based interactive user interface, called *Virtual-pen*, for computer aided surgery. Our proposal is aimed to eliminate an ink pen which is currently used to mark surgical targets and to virtually project marking of surgical targets onto human body directly in real-time. The virtual-pen will be more convenient to modify or delete and more sanitary than the ink pen. It can be also used as an interactive device for other medical operations such as visualizing medical imagery on display and controlling the brightness and the position of an operation lamp, and so on. This paper is organized as follows. Section 2 introduces our framework briefly and explains component techniques in detail. Section 3 shows experiment results. The conclusion is drawn in Section 4.

0-7695-3056-7/07 $25.00 © 2007 IEEE
DOI 10.1109/ICAT.2007.42

2. Framework

Our system consists of two devices: a projector-camera system and an optical tracking device (see Figure 1). The projector-camera device has been remarked as good framework to implement DirectAR technology. It estimates relative geometry among camera-, projector-image planes in 2D and surgical targets in 3-D. The optical tracking device is used for tracking the position and orientation of a probe which serves as Virtual-pen. It enables to track the probe robustly in real-time without interfering with operation illumination.

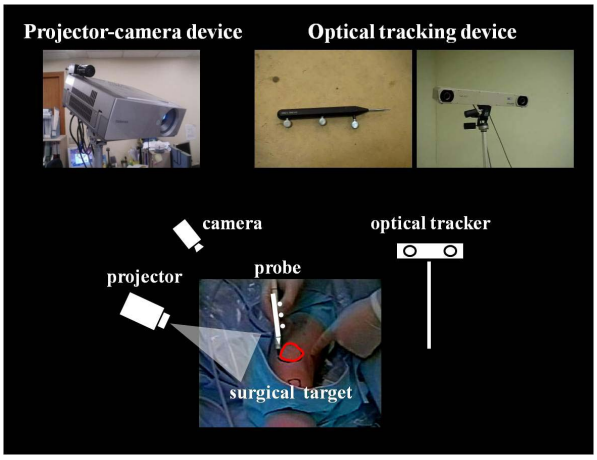

Figure 1. A projector-camera device identifies surface geometry and directly projects trajectory images onto surgical targets. An optical tracking device measures the position and orientation of a probe in real-time.

2.1. System Calibration

To calibrate a projector, a camera, and an optical tracker which have different coordinates respectively, a 'standard' checkerboard pattern which is fixed and planar is used. The camera is calibrated by typically using Zhang's calibration method [18]. For calibrating the projector, we use a modified version of the Zhang's calibration method [4] because the projector can be modelled as inverse projection of the pin-hole camera, which is considered as a perspective view. The optical tracker is calibrated by manually defining three reference vectors of the coordinates on the checkerboard pattern which is same as it used in former calibration steps. Definitely all devices are initially calibrated one time.

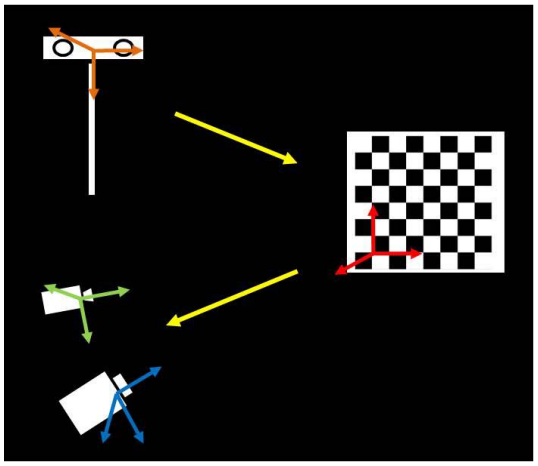

Figure 2. Geometric relationship among the coordinates of a projector (blue), a camera (green), an optical tracker (orange), and a surface (red). Three devices are initially calibrated by using a fixed planar checkerboard pattern.

2.2. Estimation of Surface Geometry

There are many different methods to reconstruct the geometry of 3-D surface. Yang *et al.* [19] categorized different methods based on two orthogonal criteria, passive vs. active, and online vs. off-line. Active methods use external light sources such as beam projectors, while passive methods use the only ambient light. Off-line methods usually interrupt normal operations, but online methods can use while the system is in normal use. In this paper, an imperceptible coded structured light technique, which is one of active online methods, is used for finding correspondences between different image planes without noticing intrusive patterns. The geometry of the 3-D surface with respect to two perspective views is recovered by using triangulation method [20]. In the following sections, we will describe how the geometry of the 3-D surface is estimated in detail.

2.2.1. Imperceptible Coded Structured Light Technique. Coded structured light is a technique based on the projection of patterns onto the measuring surface. Structured light techniques were classified by Salvi *et al.* [21] according to coding strategies. Among structured light techniques, temporal coding methods

are commonly used where projectors project a set of patterns successively onto the measuring surface and cameras take images of patterns. Every pixel has its own codeword formed by the sequence of illumination values for that pixel across the projected patterns and each correspondence (or direct mapping) between image points and points of the projected pattern is easily found without geometrical constraints. When illumination levels are two, a codeword is coded as 0 and 1 (see Figure 3).

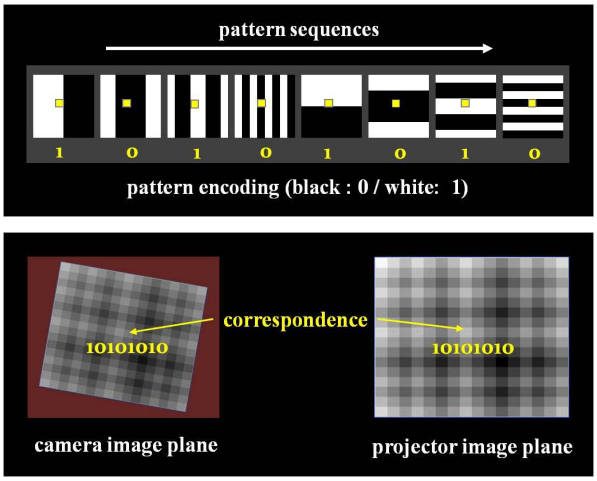

Figure 3. Binary coded structured light patterns and finding correspondences.

Binary coded structured light technique is good for the purpose of real-time estimation and high accuracy because every pixel is precisely encoded by projecting a set of patterns consecutively. However, this technique cannot be directly applied to augmented reality because the projection of patterns visually interferes with performing the normal operation. To resolve this problem, Yasumuro *et al.* [22] used an additional infrared projector to project near-infrared patterns which are invisible to human eyes. Raskar *et al.* [23] proposed a method called imperceptible structured light. They used special engineered digital light projector (DLP) that was able to turn light on and off at a very high rate. Also structured light patterns were inserted on between bit-plane, so it made human eye to be unable to percept between bit-planes include structured light patterns and complementary patterns by switching very fast. Recently, Park *et al.* [5] proposed a method that embeds complementary patterns into augmented reality images based on a simple pixel operation without using sensors or special-purpose projector.

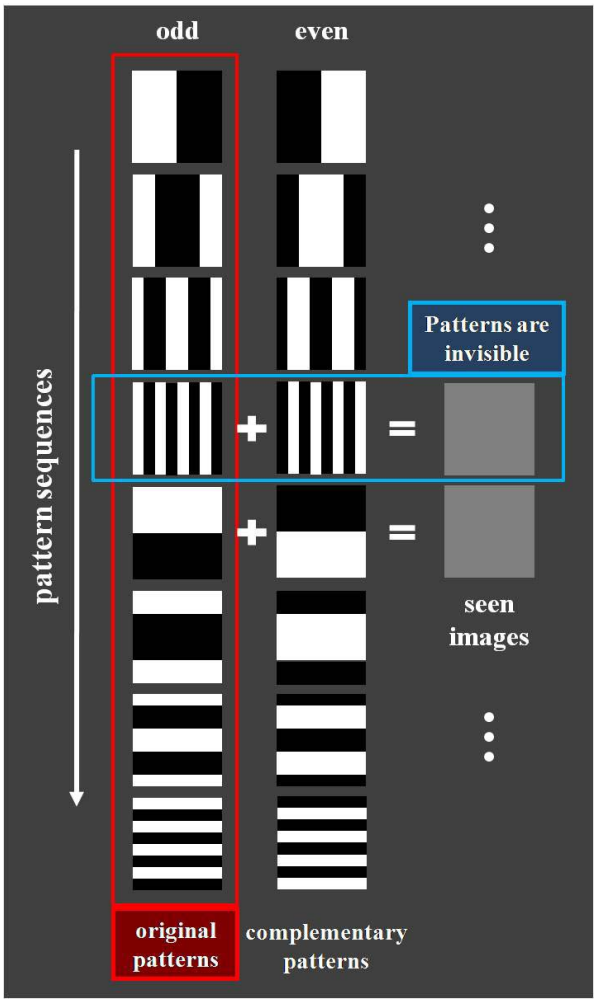

Figure 4. Complementary patterns are inserted every even frame. Patterns of both odd and even frames are projected at a double frame rate and then original patterns are invisible to users.

Imperceptible coded structured light technique gives our system more practical solutions in the real operation environment because it makes surgeons not only to free from interruption of structured light patterns during operation and but also to process in real-time. In our approach, complementary patterns every even frame are inserted. When the projection frame rate of sequence is doubled, original patterns would be externally invisible because increase and decrease in the pixel value are offset by the characteristic of the human vision (see Figure 4). The detection of original patterns is achieved by capturing

every odd frame. In order to detect original patterns every the exact time, a camera should be synchronized in advance. We use an external trigger port of the camera that provides control over timing of the image acquisition. A vertical synchronization signal is extracted from a VGA signal which is input into a projector and is provided to the external trigger port (see Figure 5).

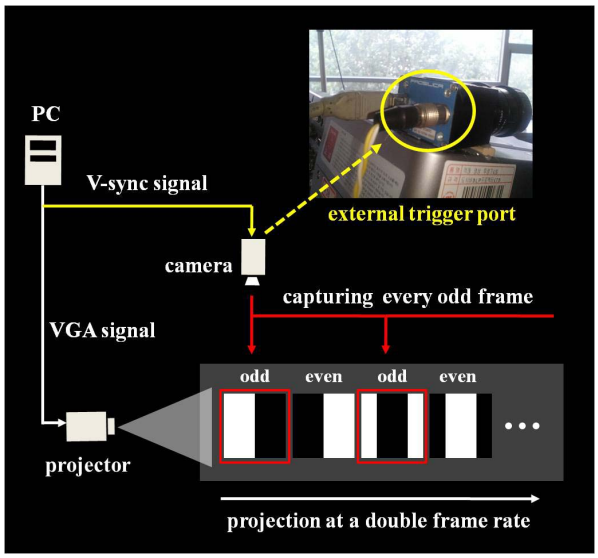

Figure 5. A vertical synchronization signal is extracted from a VGA signal that is connected to a projector and is supplied to an external trigger port of a camera. When the camera is synchronized with the projector that projects patterns at a double frame rate, the camera detects original patterns by capturing every odd frame.

2.2.2. Surface Geometry Reconstruction. The key point of this section is to estimate a point X of a 3-D surface which satisfies the given camera geometry and projector geometry. The relationship between a surface and a camera-, a projector-image plane is respectively presented as

$$x_c = P_c X\ ,\ x_p = P_p X \tag{1}$$

where x_c is a point of a camera image, x_p is a point of a projected image, P_c is a camera projective matrix, P_p is a projector projective matrix, and X is a point of a 3-D surface. Let triangulation is denote by τ, estimation X is written as

$$X = \tau(x_c, x_p, P_c, P_p)\,. \tag{2}$$

There are four unknown factors but the camera projective matrix P_c and the projector projective matrix P_p are computed by calibration and correspondences between points of the camera image and points of the projected image $x_c \leftrightarrow x_p$ are found by coded structured light technique.

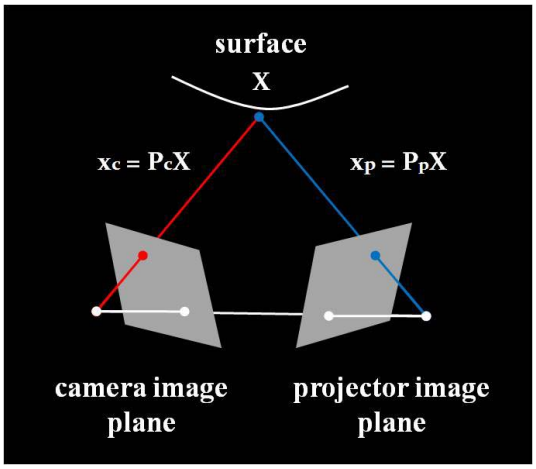

Figure 6. The geometry of two perspective views. The geometry of a 3-D surface is reconstructed by using triangulation method.

Between a point of the camera image and a point of the 3-D surface, $x_c \times (P_c X) = 0$ and written as follows:

$$\begin{aligned} x({p_c}^{3T}X) - ({p_c}^{1T}X) &= 0 \\ y({p_c}^{3T}X) - ({p_c}^{2T}X) &= 0 \\ x({p_c}^{2T}X) - y({p_c}^{1T}X) &= 0 \end{aligned} \tag{3}$$

where $x_c = \{x, y, z\}$, $X = \{X, Y, Z, 1\}$, and p_c^{iT} is rows of P_c. Equations between a point of the projected image and a point of the 3-D surface are obtained in the same way and then an equation of the form $AX = 0$ is composed with

$$A = \begin{bmatrix} xp_c^{3T} - p_c^{1T} \\ yp_c^{3T} - p_c^{2T} \\ x'p_p^{3T} - p_p^{1T} \\ y'p_p^{3T} - p_p^{2T} \end{bmatrix} \quad (4)$$

where $x_p = \{x', y', z'\}$, $X = \{X, Y, Z, 1\}$, and p_p^{iT} is rows of P_p. This is a redundant set of equations, since the solution is determined only up to scale. The geometry of the 3-D surface is presented as the form of a triangular mesh which is piece-wise planar.

2.3. Computation of Trajectory Images

Marking surgical targets are to project trajectory images of a probe onto surgical targets. Trajectory of the probe can be considered as a set of points which a ray that the probe is pointing are intersected with one of triangles of a reconstructed surface mesh. Here is an algorithm to find intersections [24]. The goal of the algorithm is first to estimate the coordinates of intersections with respect to the triangle's vertices and to determine if the ray goes through the triangle. Let a ray which is a direction vector of the probe is $r(t)$, the parametric representation of the ray is

$$r(t) = O + \vec{D}t \quad (5)$$

where O is a origin point of the ray, $\vec{D}$ is a direction vector of the ray, and t is a point on the ray. The implicit representation of the plane is

$$\vec{N} \cdot p + d = 0 \quad (6)$$

where $\vec{N}$ is a normal vector of the plane, p is a point on the plane, and d is constant. The normal of the plane containing the triangle $\vec{N}$ is computed with the cross product (see Figure 7):

$$\vec{N} = \overrightarrow{V_0V_1} \times \overrightarrow{V_0V_2}. \quad (7)$$

The constant d is computed by the dot product:

$$d = -V_0 \cdot \vec{N} \quad (8)$$

The intersection I is computed by calculating the following equation:

$$I = -\frac{d + \vec{N} \cdot O}{\vec{N} \cdot \vec{D}}. \quad (9)$$

Note that following three cases are not considered: (i) when the triangle and the ray are parallel ($\vec{N} \cdot \vec{D} = 0$), (ii) when the intersection is behind the origin of the ray ($I \le 0$), (iii) when a closer intersection has been already found for the ray ($I \ge t_{ray}$).

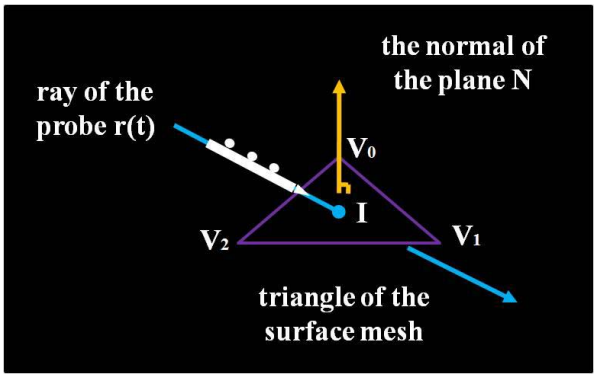

Figure 7. Estimation of ray-triangle intersections.

Next, we should determine whether an intersection is inside of the triangle or not. The intersection I is written by

$$\overrightarrow{V_0I} = \alpha\overrightarrow{V_0V_1} + \beta\overrightarrow{V_0V_2}. \quad (10)$$

Both coefficients α, β are computed by solving the following equations:

$$\begin{aligned} x_p - x_0 &= \alpha(x_1 - x_0) + \beta(x_2 - x_0) \\ y_p - y_0 &= \alpha(y_1 - y_0) + \beta(y_2 - y_0) \\ z_p - z_0 &= \alpha(z_1 - z_0) + \beta(z_2 - z_0) \end{aligned} \quad (11)$$

where $I_0 = \{x_p, y_p, z_p\}$, $V_0 = \{x_0, y_0, z_0\}$, $V_1 = \{x_1, y_1, z_1\}$, and $V_2 = \{x_2, y_2, z_2\}$. We can expect that the intersection I is inside of the triangle if

$$\alpha \ge 0, \beta \ge 0, \alpha + \beta \le 1. \quad (12)$$

Finally, a trajectory image for projecting onto the surgical target is pre-warped as follows:

$$x_{trajectory} = P_p I \quad (13)$$

where $x_{trajectory}$ is a point of a projected trajectory image, P_p is a projector projective matrix, and I is a intersection.

3. Experimental Results and Discussion

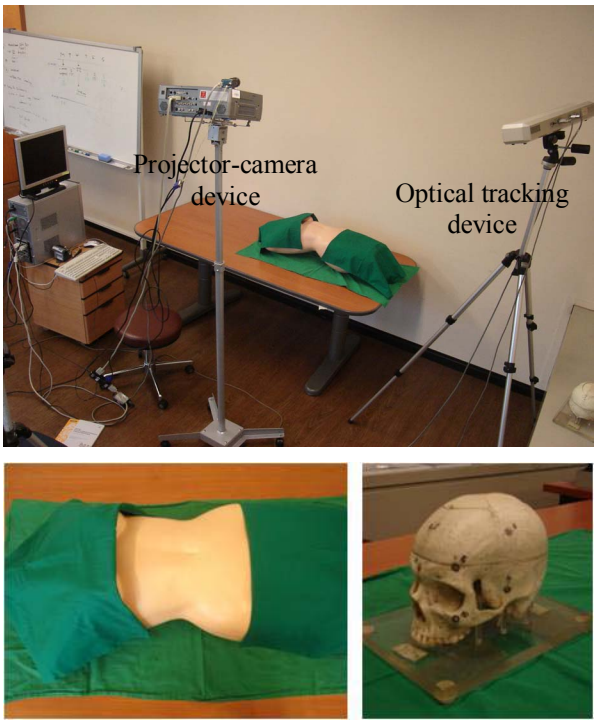

Figure 8. (top) A projector-camera device and an optical tracking device are set up in a laboratory environment. Demonstration models are (bottom-left) a phantom of torso and (bottom-right) a phantom of head.

Our experimental environment was composed of a DLP projector (TOSHIBA, TDP-T350), a camera (Prosillica, EC655C), an optical tracker (NDI, hybrid position sensor), and a probe (Traxtal Technologies, HP005 series) (see Figure 8(top)). The camera was perfectly synchronized with the projector and it could capture original patterns every odd frame while the projector successively projected a set of patterns at a double frame rate. A measurement volume of the optical tracker was set up reliably and the 3-D Root Mean Square (RMS) volumetric accuracy acceptance criterion for the position sensor of the optical tracker was ≤ 0.350 mm. Note that the probe should be located within the set measurement volume to obtain optimum accuracy and its makers should not be shadowed by any objects. Our experiments were demonstrated using a phantom of torso and a phantom of head (see Figure 8(bottom)).

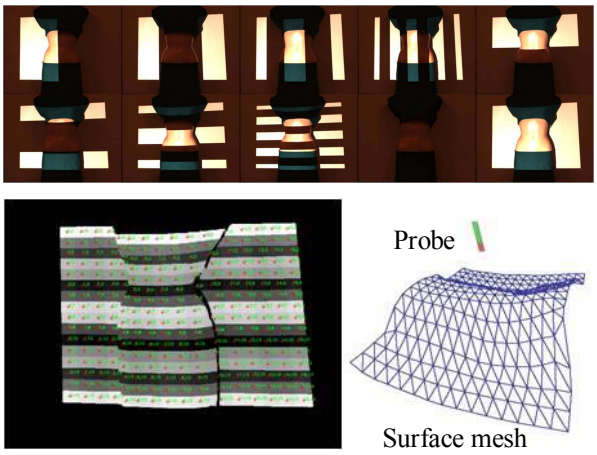

Figure 9. (top) A synchronized camera captures every odd frame. (bottom-left) A codeword map is obtained from the sequence of patterns. (bottom-right) The geometry of the 3-D surface is presented as the form of a triangular mesh.

Figure 9 shows the process for reconstructing a 3-D surface of the phantom of torso based on binary coded structured light technique combined with complementary patterns. Every even frame image was offset by every odd frame image and surgeons could not notice original patterns externally. On the other hand, original patterns were extracted by capturing every odd frame. Each codeword that were obtained from the sequence of patterns was shown as a map in Figure 9(bottom-left). Correspondences between image points and points of the projected image were computed by direct mapping of the codeword respectively. The geometry of the 3-D surface was presented as the form of a triangular mesh in Figure 9(bottom-right).

Figure 10 and Figure 11(bottom) show the results of marking surgical targets using Virtual-pen. The direction vector of a probe was obtained by the coordinates of the probe which was tracked in real-time. The coordinates of points that intersected the direction vector of the probe with a triangle of the surface mesh were calculated mathematically.

Trajectory images of the probe were obtained by the coordinates of intersections and were projected simultaneously while the probe was moving. The drawn lines onto the models were good in both cases and especially Figure 11(bottom), where the line was drawn over an existing line gave an idea of the accuracy. Actually, the drawing accuracy is not a critical factor of our system because marking surgical targets is to inform surgeons the positions of surgical targets, so drawing errors caused by patient's motion (e.g. breathing) are not a problem.

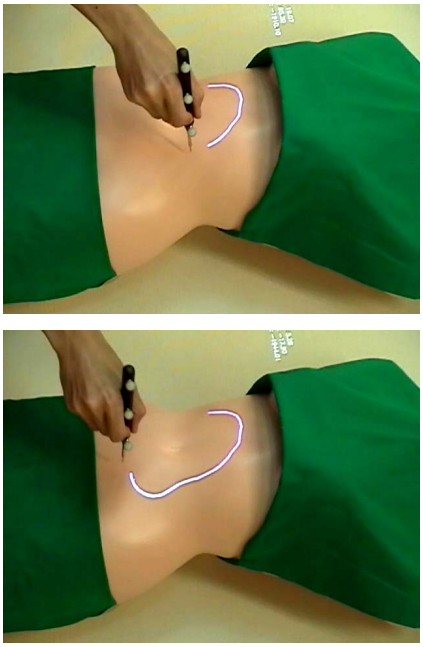

Figure 10. Demonstration onto a phantom of torso. Trajectory images of a probe are directly projected onto the phantom of torso (The procedure is from up to down).

To implement interactive functions of Virtual-pen, the projection area was divided into several regions of which each corresponds to pre-defined actions: drawing and deleting. The functions were activated simultaneously when the probe was laid in regions. The pre-defined menus are useful to surgeons who handle lots of medical operations because it can be extensible to offer various functions, e.g. displaying surgical information and controlling components of surgery systems.

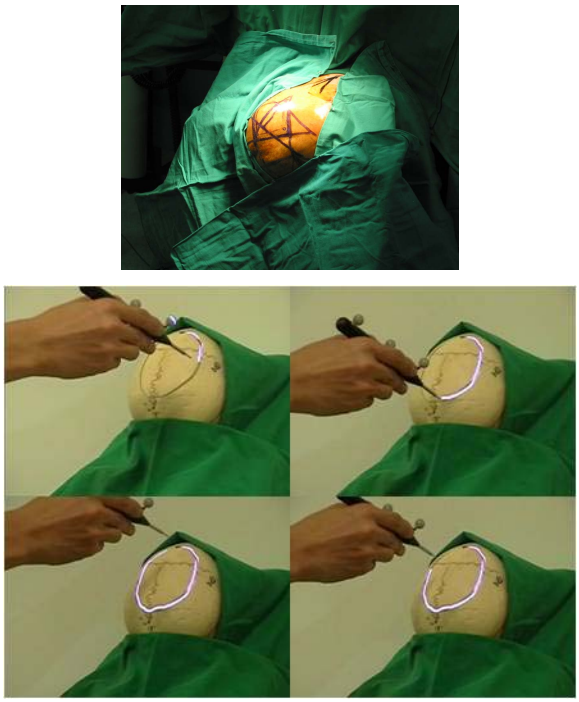

Figure 11. Comparison of (top) marking onto a human head using an ink pen with (bottom) marking onto a phantom of head using Virtual-pen (The procedure is on clockwise).

4. Conclusion

In this paper, it was demonstrated that DirectAR based interactive user interface could completely resolve problems by successfully marking surgical targets without using conventional tools such as an ink pen. By using the projector-camera device and an accurate optical tracking device, 3-D surfaces of surgical targets were reconstructed and surgical targets were marked by directly projecting onto surfaces. In addition, pre-defined menus offered interactive functions of the system dynamically. Therefore our successful results enhanced computer aided surgery and supported image guided surgery. Currently, we are trying to perform the clinical test to verify our computer aided surgery system.

5. Acknowledgements

This study was support by a grant (02-PJ3-PG6-EV04-0003) of Ministry of Health and Welfare, Republic of Korea.

References

[1] O. bimber and R. Raskar, *Spatial Augmented Reality: Merging Real and Virtual Worlds*, A K Peters, 2005.

[2] R. Sukthankar, R.G. Stockton, and M.D. Mullin, "Smarter Presentations: Exploiting Homography in Camera-Projector Systems," In *Proceedings of ICCV*, Nov. 2001, pp. 247-253.

[3] H. Park and J.-I. Park, "Modern Approaches to Direct-Projected Augmented Reality: A Review," In *Proceedings of International Symposium on Ubiquitous Virtual Reality (ISUVR)*, Jul. 2006, pp. 113-114.

[4] H. Park, M.-H. Lee, S.-J. Kim, and J.-I. Park, "Surface-Independent Direct-Projected Augmented Reality," In *Proceedings of ACCV*, Jan. 2006, pp. 892-901.

[5] H. Park, M.-H. Lee, B.-K. Seo, and J.-I. Park, "Undistorted Projection onto Dynamic Surface," In *Proceedings of Pacific-Rim Symposium on Image and Video Technology (PSIVT)*, Dec. 2006, pp. 582-590.

[6] S.K. Nayar, H. Peri, M.D. Grossberg, and P.N. Belhumeur, "A Projection System with Radiometric Compensation for Screen Imperfections," In *Proceedings of IEEE International Workshop on Projector-Camera Systems (PROCAMS)*, Oct. 2003, pp. 193-206.

[7] M. Ashdown, T. Okabe, I. Sato, and Y. Sato, "Robust Content-Dependent Photometric Compensation," In *Proceedings of IEEE International Workshop on Projector-Camera Systems (PROCAMS)*, Jun. 2006, pp. 60-67.

[8] A. Grundhofer and O. Bimber, "Real-Time Adaptive Radiometric Compensation," *IEEE Transaction on Visualization and Computer Graphics*, IEEE Computer Society, Feb. 2007, Vol. 13, No. 5. pp. 1-12.

[9] O. Bimber, A. Emmerling, and T. Klemmer, "Embedded Entertainment with Smart Projectors," *Computer*, IEEE Computer Society, Jan. 2005, Vol. 38, pp. 48-55.

[10] J. Rekimoto and M, Saitoh, "Augmented Surfaces: A Spatially Continuous Work Space for Hybrid Computing Environments," In *Proceedings of SIGCHI*, May 1999, pp. 378-385.

[11] C. Shen, K. Ryall, C. Forlines, A. Esenther, F. Vemier, K. Everitt, M. Wu, D. Wigdor, M.R. Morris, M. Hancock, and E. Tse, "Informing the Design of Direct-Touch Tabletops," *Computer Graphics and Applications*, IEEE Computer Society, Sep./Oct. 2006. Vol. 5, Issue 5, pp. 56-66.

[12] R. Kjeldsen, A. Levas, and C. Pinhanez, "Dynamically Reconfigurable Vision-Based User Interface," *International Journal of Machine Vision and Applications*, 2004, Vol. 16, No 1, pp. 6-12.

[13] K. Oka, Y. Sato, and H. Koike, "Real-Time Tracking of Multiple Fingertips and Gesture Recognition for Augmented Desk Interface Systems," *Computer Graphics and Applications*, IEEE Computer Society, Nov. 2002. Vol. 22, No 6, pp. 64-71.

[14] M.F. Zaeh and W. Vogl, "Interactive Laser-Projection for Programming Industrial Robots," In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR)*, Oct. 2006, pp. 125-128.

[15] W.E.L. Grimson, T. L., W.M. Wells, G.J. Ettinger, S.J. White, and R. Kikinis, "An Automatic Registration Method for Frameless Stereotaxy, Image Guided Surgery, and Enhance Reality Visualization," *IEEE Transactions on Medical Imaging*, Apr. 1996, Vol. 15, No 2.

[16] H. Hoppe, G. Eggers, T. Heurich, J. Raczkowsky, R. Marmulla, H. Worn, S. Hassfeld, and J.L. Moctezuma, "Projector-Based Visualization for Intraoperative Navigation: First Clinical Results," In *Proceedings of Computer Assisted Radiology and Surgery*, Jun. 2003.

[17] C. Schonfelder, L.A. Kahrs, T. Stark, J. Raczkowsky, and H. Worn, "Calculation and Visualization of Trocar Positions for Abdominal Minimally Invasive Surgery," *International Journal of Computer Assisted Radiology and Surgery*, Springer, Jun. 2007, Vol. 2, Supplement 1, pp. S506.

[18] Z. Zhang, "Flexible Camera Calibration by Viewing a Plane from Unknown Orientations," In *Proceedings of ICCV*, Sep. 1999, pp. 666-673.

[19] R. Yang and G. Welch, "Automatic Projector Display Surface Estimation Using Every-Day Imagery," In *Proceedings of Central Europe on Computer Graphics, Visualization and Computer Vision*, Feb. 2001.

[20] R. Hartely and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2003.

[21] J. Salvi, J. Pages, and J. Batlle, "Pattern Codification Strategies in Structured Light Systems," *Pattern Recognition*, Elsevier, Dec. 2004, Vol. 37, No 4, pp. 827-849.

[22] Y. Yasumuro, M. Imura, Y. Manabe, O. Oshiro, and K. Chihara, "Projection-Based Augmented Reality with Automated Shape Scanning," In *Proceedings of the SPIE*, Mar. 2003, Vol. 5564, pp. 555-562.

[23] R. Raskar, G. Wlch, M. Cutts, L. Stesin, and H. Fuchs, "The office of the Future: A Unified Approach to Image-Based Modeling and Spatially Immersive Displays," In *Proceedings of SIGGRPH*, Jul. 1998, pp. 179-188.

[24] D. Badouel, *Graphics Gems*, Academic Press, 1990.