

Augmented Reality System for Oral Surgery Using 3D Auto Stereoscopic Visualization

Huy Hoang Tran¹, Hideyuki Suenaga², Kenta Kuwana¹, Ken Masamune¹,
Takeyoshi Dohi¹, Susumu Nakajima², and Hongen Liao¹

¹ The University of Tokyo, Tokyo 113-8656, Japan

² The University of Tokyo Hospital, Tokyo 113-8655, Japan

Abstract. We present an augmented reality system for oral and maxillofacial surgery in this paper. Instead of being displayed on a separated screen, three-dimensional (3D) virtual presentations of osseous structures and soft tissues are projected onto the patient's body, providing surgeons with exact knowledge of depth information of high risk tissues inside the bone. We employ a 3D integral imaging technique which produce motion parallax in both horizontal and vertical direction over a wide viewing area in this study. In addition, surgeons are able to check the progress of the operation in real-time through an intuitive 3D based interface which is content-rich, hardware accelerated. These features prevent surgeons from penetrating into high risk areas and thus help improve the quality of the operation. Operational tasks such as hole drilling, screw fixation were performed using our system and showed an overall positional error of less than $1mm$. Feasibility of our system was also verified with a human volunteer experiment.

1 Introduction

Image guided systems have been supporting surgeons in a wide range of clinical applications from diagnostic to postoperative evaluation [1,2]. In oral and maxillofacial surgery, the prerequisite for a successful operation is the exact knowledge of high risk tissues such as blood vessels and dental nerves so that these inferior tissues can be avoided [3]. One challenging problem in current navigation systems is the lack of depth information of the visualized structures due to the limitation of conventional flat displays. In addition, as the visual information is often separated from the surgical site, surgeons have to make a hand-eye transformation in order to match visual information to the real scene.

In order to give the surgeons an intuitive grasp of the depth information of a surgical scene, several approaches have been presented [4,5]. Among those, the image overlay system using Integral Videography (IV) [6,7] is a promising solution for the simplicity in implementation and the ability to produce motion parallax in both horizontal and vertical directions. In such system, users can see the 3D internal structure superimposed onto the actual anatomy through a half-silvered mirror [8]. However, the low rendering speed and the lack of the ability to visualize enriched information make it difficult to use such system in navigation.

For example, one would expect to interactively change the transparency of a surface to see the blood vessel inside, or get a real-time visual feedback about the progress of the operation.

In this study, we present an augmented reality system for oral surgery using IV image. We develop a GPU based rendering algorithm to accelerate the IV images generation. Our system also offers an intuitive 3D based user interface that allows surgeons to interact with the image in 3D environment. Feasibility study of the system is carried out with a set of experiments including a human volunteer experiment.

2 Materials and Methods

2.1 3D IV Display and Rendering

The basis of IV comes from the principle of Integral Photography [9]. To reproduce 3D image of an object in real space, a micro lens array is placed in front of a high density LCD display. Each pixel in the portion behind a micro lens emits a light ray in a unique direction that connects the pixel and the center of the lens. From a specific viewpoint, one can only see the set of light rays emitted from the pixels that lie on the straight lines connecting the viewpoint and the lenses. Consequently, different aspects of an object can be observed from different directions, giving the observer a sense of depth (Fig. 1(a)).

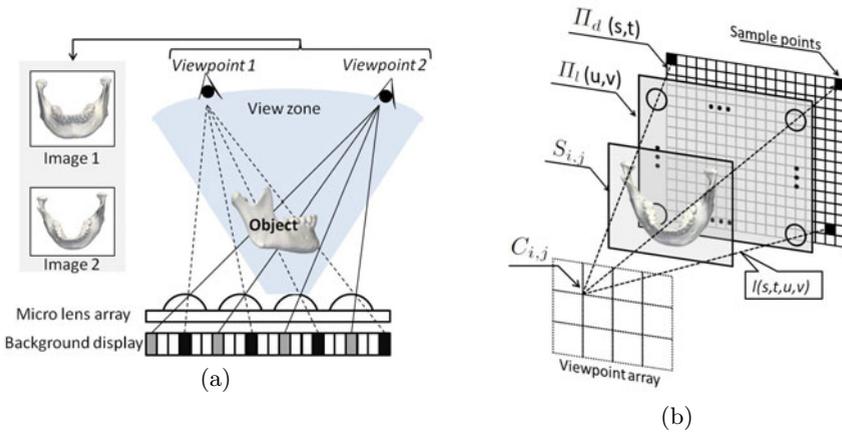


Fig. 1. (a) Principle of Integral Videography. (b) Sampling an image using Π_l and Π_d

In order to create an IV image, every pixel on the background display need to be computed. There are two main IV rendering methods: volume ray-tracing and surface based rendering. In the former, value of each pixel is depends on how the corresponding light ray intersects with the object. This method requires computation for every light ray but creates very accurate 3D images. The latter uses CG surface rendering to render an array of multi-view images from a set

of predefined viewpoints. These images are then mapped backward to the background image using pixel distribution algorithm [10]. This method has all the advantages of surface rendering (shading, texture mapping, etc.). However, due to the fact that the spatial arrangement and shape of the lens in general differ from those of the pixels in an image, apparent distortions are observable. For that reason, surface based method is often used for other purposes than surgical navigation.

In order to render accurate IV image while maintaining the desired visual effects, we introduce a light field [11] approach to the surface based method. The image $S_{i,j}$ rendered from view $C_{i,j}$ is projected onto the lens array plane Π_l and the display plane Π_d with the center of projection located at $C_{i,j}$. Applying the transformation for all i and j will result in a densely sampled light space in which each light ray is defined as a four parameters function $l(s, t, u, v)$ ((s, t) from Π_d and (u, v) from Π_l)(Fig. 1(b)).

Value of a pixel located at $\mathbf{P} = (s, t)^T$ in Π_d is computed as follow.

$$I(s, t) = \int \int_{\Pi_l} \Theta(s, t, u, v) l(s, t, u, v) dudv, \quad (1)$$

where $\Theta(s, t, u, v)$ is a weight function that defines the contribution of the light ray to the final result such that

$$\int \int_{\Pi_l} \Theta(s, t, u, v) dudv = 1. \quad (2)$$

The design of the weight function is beyond the scope of this paper. However, for a well-manufactured lens array, contribution of light rays from lenses located far away from the query pixel can be neglected. Therefore a Dirac delta function in which the base width equals to the diameter of one lens can be a good choice.

$$\Theta(s, t, u, v) \stackrel{def}{=} \delta_R(d) = \frac{1}{R\sqrt{\pi}} e^{-\frac{d^2}{R^2}}, \quad (3)$$

where d is the Euclidean distance between $(s, t)^T$ and $(u, v)^T$, R is the radius of a micro lens.

$l(s, t, u, v)$ can be computed by resampling Π_l and Π_d , this is done through several interpolations using the sampled points. These interpolations are in general time consuming tasks that slow down the rendering speed significantly. Fortunately, recent GPUs provide a very effective way to perform these interpolations through their naive texture accessing function. To do so, all of the images are rendered and stored in the GPU's texture memory. Next, a rectangle is rendered as a proxy geometry where textures are loaded. At the final stage of the graphics pipeline, each fragment (i.e. a potential pixel) is fed with the pixel value from equation 1. Per fragment processing is implemented using a GPU shading language, the C for Graphics [12]. Our algorithm does not require external CPU memory so we can get rid of the bottlenecks in data transmission between GPU and CPU.

2.2 IV Augmented Reality System

A half-silvered mirror is attached to the IV display. Through the half mirror, surgeons can see the reflected IV image superimposed onto the patient (Fig. 2(a)). We use a fly eyes type lens array with lens diameter of 1.001 mm, which is fixed to a 6.4 inches high resolution LCD (203 dpi). In order to track the movement of the patient, a marker frame is fixed to the lower jaw. The marker frame is calibrated using a model of exactly the same size with the lower jaw.

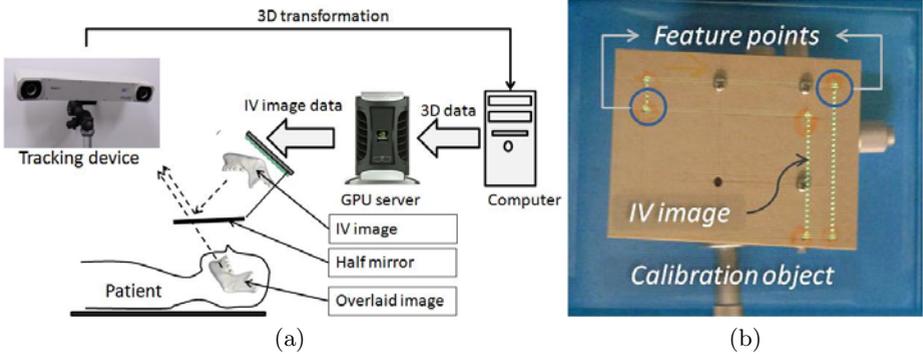


Fig. 2. (a) IV augmented reality system configuration. (b) IV image overlaid with calibration object.

Although the geometry of the IV display-half mirror combination is known, a small displacement of the lens array with respect to the display would result in a large displacement of the IV Image. To accurately register the IV image to the world coordinate system, we use a calibration object which consists of six feature points. A model of the object is created inside a computer and its IV image is then overlaid with the real object. Manual adjustment in position and orientation is made to the object until all of the feature points are aligned completely as shown in Fig. 2(b).

In order to support surgeons in operational tasks such as hole drilling, bone resection and screw fixation, in addition to the osseous structures, IV images of the instrument, surgical path, soft tissues are utilized in our system. Surgeons can verify the progress of the operation, i.e. current position of the instrument, status of the bone and the risk tissues visually or through text based information. The reason why we add text based information is that it improves the surgeons' confidence in situations where small differences (less than 1mm) are difficult to recognize visually.

3 Experiments and Results

3.1 Image Generation Speed

The prerequisite for our system is that movement of the instrument as well as the patient has to be updated in real-time. In addition, surgeons should have the

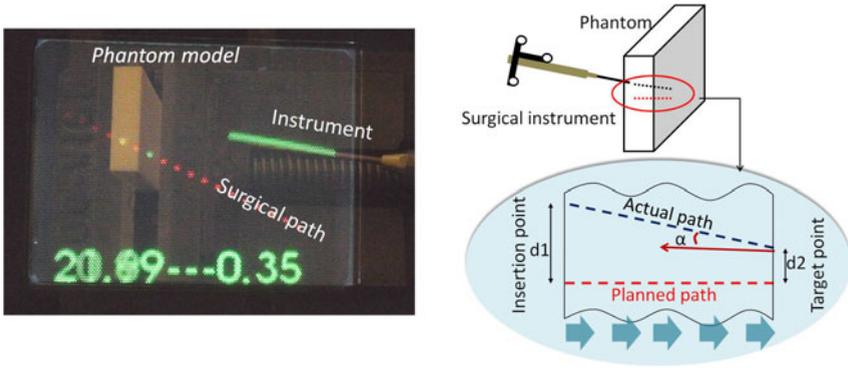


Fig. 3. Drill experiment using image overlay system

ability to change the properties of the displayed objects to obtain the desired visual effect. Therefore, we evaluated the rendering speed of each individual object. The specifications of our computer system are as follow.

CPU: Pentium Core 2 Duo, 2.4 GHz

GPU: NVIDIA Quadro Plex model 4 (Quadro FX 5600 \times 2)

Elemental IV image size: 1024 \times 768 pixels

The rendering system can render IV image of a surgical instrument at 80 frames per second (fps). For a complex anatomical model, e.g. the whole lower jaw, rendering speed is 5 fps. Therefore, the minimum updating speed of our system is around 5 fps. In most cases, our system shows a rendering speed of more than 10 fps, which is sufficient for surgical navigation.

3.2 Accuracy Evaluation

In order to evaluate the accuracy of our system when used in a real clinical case, a surgeon was asked to performed operational tasks such as hole drilling using real surgical instruments. The drilling task was performed on a phantom model imitating a bone surface. For the sake of simplicity in error estimation, simple models with known geometry were used rather than complex anatomical models. The objective of the operation is to drill through the model along a predefined path. During the operation, the surgeon was asked to keep the instrument's axis as close as possible to the surgical path, otherwise the operation would be considered a failure.

The drill path was set to be not perpendicular to the model's surface so that the surgeon would not get a hint from the model's surface. As shown in (Fig. 3), the drill path was inclined at approximately 45° to the model's surface.

When the instrument is aligned properly to the surgical path, color of the instrument is automatically changed to alert the surgeon that it is safe to proceed. The surgeon will then start his drilling task. The operation ends when the drill

reaches the target point on the other side of the model. By measuring the position of the drill holes in both sides, we can compute the positional and angular errors of the operation.

After 10 trials, positional error of the insertion point (d_1) was $0.7 \pm 0.4mm$, the error of the target point (d_2) was $0.5 \pm 0.3mm$ and the angular error (α) was $3.3 \pm 1.1^\circ$. Error of the target tends to be smaller than that of the insertion point. The reason is that even the first approach to the model's surface was not good, the surgeon was still able to correct the orientation of the instrument during the drilling task.

3.3 Feasibility Evaluation

Screw fixation using model. Another task that is often performed during oral surgery is screw fixation. We use the anatomical model of a lower jaw in this experiment. The prerequisite for a successful operation is that blood vessels and nerves must not be damaged. For that reason, a nerve channel underneath the bone surface is highlighted (Fig. 4(a)). During operation, an alert is sent to the surgeon when the instrument matches the planned path, or when the nerve channel is likely to be touched. Again, a surgeon was asked to perform the screw fixation under the guidance of our system. In order to mimic as closely as possible the clinical situation, the surgeon used the instrument set (screw and screw driver) that was manufactured for medical use.

After the fixation was completed, we verified that the screw was fixed in the proper place as planned in the planning stage, with no threat of damaging the nerve (Fig. 4(b)).



Fig. 4. (a) IV image of the bone with blood vessel highlighted. (b) screw fixation result.

Human volunteer experiment. A problem should come up when utilizing our system in real clinical case is the limited work space due to the presence of IV display and half mirror in the surgical area. Another problem should be the response speed of the system when patient make a movement. A volunteer experiment was carried out by the same surgeon to evaluate the influence of those factors in real clinical applications.

Fig. 5 (left) shows the settings of this experiment. The half mirror was placed at about 25 cm above the surgical area, giving the surgeon a sufficient work space. A marker frame was attached tightly to the patient's lower jaw for a proper position tracking. Fig. 5 (right) shows the surgical area as viewed from the surgeon's point of view. The IV image was superimposed with the patient correctly, giving the surgeon a fast and accurate understanding of the surgical area. Any movement made by the patient (movement of the head, the lower jaw, etc.) was properly detected and IV images are updated with a short time delay.

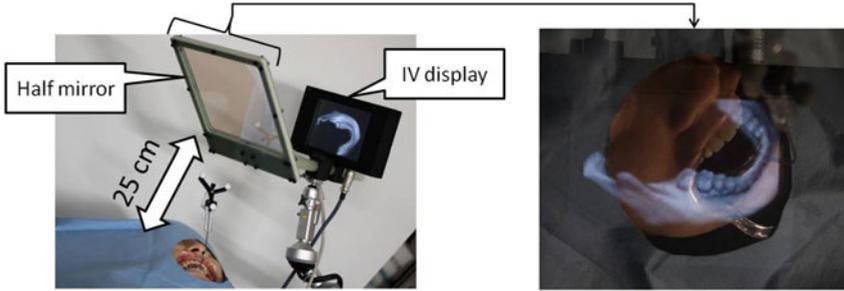


Fig. 5. Settings of volunteer experiment (left), IV image overlaid onto surgical area (right)

4 Discussion and Conclusion

In our experiments, IV image gave surgeons a fast and intuitive estimation of the target structures in real space. However, as each micro lens becomes a logical pixel, when the instrument was closed to the surgical path (the distance is smaller than one lens), surgeons tended to rely more on textual information than visual information because displacement smaller than one lens was difficult to recognize. For that reason, the system proposed in this paper has better accuracy than previously developed systems ($0.7 \pm 0.4mm$). This accuracy met the requirement for use in clinical application and thus the potential of applying our system in real clinical cases has been proven. In addition, the GPU based rendering algorithm has enabled real-time tracking of the patient movement. Although a short time delay can be noticed, this would not become a problem because the patient movement is not so dynamic during operation.

Despite the installation of the IV display and the half mirror, surgeons still have sufficient work space for the surgery. Moreover, it can be easily removed from the surgical area when AR representation is not needed. This feature makes our system very flexible and pose no restriction to the operation workspace.

In conclusion, we have developed an augmented reality system to support surgeons in oral and maxillofacial surgery. Our system provides an on-site, fast and accurate 3D virtual representation of the surgical area. A number of experiments mimicking real clinical interventions and showed promising results. In the

future, by improving some of the technical aspects, particularly the registration method, we intend to carry out clinical trials on real patients.

Acknowledgment. This work was supported in part by Research Fellowship of the JSPS (21.9746), Grant for Industrial Technology Research (07C46050) of NEDO, Japan, Grant-in-Aid for Scientific Research on Innovative Areas (22103505) of the MEXT in Japan, and KAKENHI (222400061, 22659366).

References

1. Germano, I.M., Villalobos, H., Silvers, A., Post, K.D.: Clinical use of the optical digitizer for intracranial neuronavigation 45, 261–269 (1999)
2. Plaweski, S., Cazal, J., Rosell, P., Merloz, P.: Anterior cruciate ligament reconstruction using navigation: a comparative study on 60 patients. *The American Journal of Sports Medicine* 34, 542–552 (2006)
3. Hassfeld, S., Muehling, J.: Computer assisted oral and maxillofacial surgery: A review and an assessment of technology. *Int. Oral Maxillofac Surg.* 30, 2–13 (2001)
4. Blackwell, M., Nikou, C., M.DiGioia, A., Kanade, T.: An Image Overlay system for medical data visualization. *Medical Image Analysis* 4, 67–72 (2000)
5. Traub, J., Sielhorst, T., Heining, S.M., Navab, N.: Advanced Display and Visualization Concepts for Image Guided Surgery. *Journal of Display Technology* 4, 483–490 (2008)
6. Liao, H., Nakajima, S., Iwahara, M., Kobayashi, E., Sakuma, I., Yahagi, N., Dohi, T.: Development of Real-Time 3D Navigation System for Intra-operative Information by Integral Videography. *Journal of Japan Society of Computer Aided Surgery* 2, 245–252 (2000)
7. Liao, H., Iwahara, M., Hata, N., Dohi, T.: High-quality integral videography using a multi-projector. *Optics Express* 12, 1067–1076 (2004)
8. Liao, H., Inomata, K., Sakuma, I., Dohi, T.: 3-D Augmented Reality for MRI-guided Surgery using Integral Videography Autostereoscopic-image Overlay. *IEEE Transactions on Biomedical Engineering* 57, 1476–1486 (2010)
9. Lippmann, M.G.: Epreuves reversibles donnant la sensation du relief. *J. de Phys.* 7, 821–825 (1908)
10. Liao, H., Nomura, K., Dohi, T.: Autostereoscopic Integral Photography Imaging using Pixel Distribution of Computer Graphics Generated Image. In: *ACM SIGGRAPH* (2005)
11. Levoy, M.: Hanrahan, p.: Light field rendering. In: *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, pp. 31–42 (1996)
12. Fernando, R., Kilgard, M.J.: *The Cg Tutorial*. Addison-Wesley Professional, Reading (2003)