

The Development of Optical See-through Display Based on Augmented Reality for Oral Implant Surgery Simulation

Yen-Kun Lin¹, Kun-Lun Tsai² and Hong-Tzong Yau³

¹National Chung Cheng University, <u>arb1130@cad.me.ccu.edu.tw</u> ²National Chung Cheng University, <u>andy001134@cad.me.ccu.edu.tw</u> ³National Chung Cheng University, <u>imehty@ccu.edu.tw</u>

ABSTRACT

The aim of this paper is to develop an optical see-through display (OSTD) system applied to the oral implant training. The proposed system includes the calibration, virtual objects and augmented reality (AR) visualization. For the image on the mircodisplay is projected through the lens and located between the lens and the focal point. The calibration is used to determine the intrinsic camera parameters. In addition, the transformation among the human's eye, the tracking camera and the magnification of the lens can be computed by the geometry calibration. Finally, the virtual objects can be superimposed upon an existing scene and registered correctly during the tracking. For AR visualization, the useful information, such as the nerve canal, can be viewed during the training. Finally, the proposed system can provide an efficient, stable and realistic method for users in the oral implant operations.

Keywords: oral implant training, optical see-through display, augmented reality. **DOI:** 10.3722/cadaps.2012.111-120

1 INTRODUCTION

Tooth decay and periodontal disease are major dental problems. Worse still, missing teeth could be happened if people neglect continuously their oral problems. In addition, missing teeth is also caused by some unpredicted accident. When the time for missing teeth is lengthen, the alveolar bone is gradually atrophic and becomes low. Moreover, people's chewing function and naturalness in appearance are affected seriously by the atrophy of alveolar bone. For the reason given above, most people pay attention to missing teeth.

For the treatment of missing teeth, there are many restorative options such as dentures, partials, bridges and implants. However, none have proven to be as functionally effective and durable as implants [1]. The main reason is that a dental implant is an artificial tooth root placed into patient's jawbone and lasts a lifetime. Then, the root is not loose and unstable and thus Alveolar bone can be preserved.

It is noteworthy that the surgeon's experience and ability are important factors when making the pilot hole used to place the implant into the jawbone. For a successful surgery, the surgeon must consider not only the quality and quantity of the jawbone, but also the distance between the artificial root and the nerve canal. Moreover, the excessive pressure and friction are generated during making

the pilot hole and screwing the implant. These will lead to injure the surrounding bone and thus the bone will die. Finally, the failure of a dental implant occurs and this is called osteonecrosis. Therefore, above additional information plays an important role in the successful surgery.

There is no disagreement on this point that the real training is the most important key before the surgery. Traditionally, students in the dental training use the plastic teeth or pig's teeth to practice the dental implant operations without any additional information. In the other hand, haptic virtual reality (VR) simulators for dental surgery have shown a strong potential in dental education, recently. The advantages of these simulators are interactive, safe and reusable patient environment for novices.

However, some important functions are still missing in current VR dental simulators. First, the realism of graphics and haptic rendering is limited. In the real world, the real part is more dominant than the virtual one. Many others provide only a spherical shaped tool for simplicity of collision detection and haptic rendering. Secondly, the kinematic feedback of drilling and screwing isn't exactly computed. Nevertheless, few of these dental simulators give a detailed description about the fidelity of the feedback force and real visualization when drilling and screwing on jawbone. Most of us would accept that unrealistic haptic feedback and visualization in the training will mislead the surgeons in the real surgery.

Therefore, augmented reality visualization and the traditional training are integrated into the proposed system to solve the above problems. The AR visualization doesn't only superimpose the 3D virtual objects onto live video image of the patient's jawbone, but also provides the addition information during the training. Then, a practical approach with high fidelity of feedback force is presented. Our training system proposed in this paper has the following advantages:

- Build a optical see-through display (OSTD) system
- Superimpose the virtual objects upon an existing scene
- Execute the registration of the virtual objects using AR technique
- Combine the visualization and touch in the proposed oral implant training system

2 RELATED WORK

The following section provides an overview on related work regarding systems using OSTD and AR for medical purposes.

2.1 Optical See-through Display

The OSTD has attracted much interest in recent years. The head-mounted display (HMD) [2]-[4] and the head-mounted projection display (HMPD) [5], [6] are extended from OSTD. This technology was pioneered by Fisher [2] and Kijima and Ojika [4]. They showed OSTDHMDs using CRT, lens and beam splitter. Then, Robinson [4] replaced the original mirror by a concave to eliminate the Seidel aberrations of lens. Then, Cakmakci and Rolland [7] showed a summary of the human visual system to the design of HMDs. The same concept can also be applied to OSTD. Therefore, the luminance and type of the micro display, field-of-view (FOV) requirements and Seidel aberrations of lens must be considered. It is noteworthy that the FOV between 40 and 60 degree is applied to the scientific visualization, simulation, training and medical training. For the microdisplay of HMDs, LCoS [8] doesn't only have better reflective ability and higher contrasted image than LCD and OLED, but also less expensive. In order to achieve compact and lightweight OSTD, the microdisplay necessitates a magnification requirement because the microdisplay typically is too small to view with the unaided eye. Hence, lens design is used for the magnified virtual image of the microdisplay at a fixed distance and needs to consider the aberrations.

2.2 AR Technique and Application

Recently, the potential of AR technology in the medical education has been the topic of intensive research in the last decades [9]. In AR based system, there are many issues, such as calibration, tracking, visualization and user interaction. For calibration, the intrinsic camera parameters, such as focal length and principal point, need to be determined to model the geometry of a camera. This can be obtained using well known camera calibration techniques [10], [11]. In the other hand, the extrinsic transformation for tracking also can be computed efficiently by [12]. The data representation in AR

scene includes the slicing rendering, surface rendering and volume rendering. Additionally, there are various AR based dental applications. Rhienmora et al. [13] developed a dental training system utilizing AR technique and haptic device. The users wear a video see-through HMD and use haptic device replacing the real tool to do the crown preparation in real scene. The dental implant surgery, on the other hand, is considered as a potential field in AR technique. The first scholar to give much attention to AR based dental implant surgery was Wansehitz et al. [14]. They developed an augmented operating binocular to enable variable zoom and focus as well as customizable eye distance. Furthermore, Y. Satoshi [15] used a vision base tracking system off-the-shelf to obtain the pose of the user, drill and patient. Then, the user wear an optical see-through HMD and then see the virtual implant objects superimposed upon the real model using the marker.

3 SYSTEM OVERVIEW

We integrate the OSTD, VR and AR to develop an oral implant training system. The system includes several components. An overview of our system architecture is illustrated in Fig. 1. The system can be divided into four main modules: an OSTD module, Calibration module, an AR module and a virtual objects module.



Fig. 1: Flowchart of the dental implant training system.

For OSTD module, the microdisplay, liquid crystal on silicon (LCoS), is used to project image in this module and the lens is designed to enlarge image without the aberration and the distortion using ASAP. The tracking camera (Logitech C905) is used to capture the real scene during image acquisition. The computer can detect the marker within the image from the tracking camera and then compute the transform matrix between the tracking camera and the marker. Furthermore, the view of the human's eye called the virtual eye is captured via other camera (Logitech C905).

In order to register correctly the virtual objects onto the marker, the calibration is necessary. One has to keep in mind that the calibration only has to be carried out once after assembly of the setup. Moreover, the AR module also needs to be integrated into the calibration module. The calibration module resolves itself into the following two points. First, the intrinsic camera parameters for the tracking camera are calculated by presenting printed checkerboard images of known dimensions to the camera. From these known images, the camera model can be iterated until the parameters are estimated to the sufficient precision. After the intrinsic calibration among the human's eye, the tracking camera and the magnification of the lens must be considered via the geometry calibration. The method is based on [12] and the stereo vision concept. The transformation between the marker and the virtual eye can be calculated by the stereo vision concept. Finally, there should be a scale relationship between the image from the virtual eye and the microdisplay due to the same FOV

assumption. The geometry calibration is used to find the scale relationship and modify the original transformation of virtual objects.

After finishing the calibration module, the AR module is used to obtain the transformation between the tracking camera and the marker interactively during the tracking. Once above transformations in calibration module and AR module, all objects can be located in the virtual eye coordinate in our system.

As the last component of the system, a patient RP jawbone has been included which is augmented with some virtual objects such as nerve canal, surgical guide and auxiliary information.

4 CALIBRATION AND TRACKING TECHNIQUE

This section describes the calibration method and the tracking technique. The intrinsic camera parameters and the extrinsic transformation among the marker and the tracking camera and the virtual eye are computed. This proposed method is extended by [12].

4.1 Intrinsic Camera Parameters

The relationship between a 3D point and the corresponding 2D image point for the commonly used pinhole camera model and can be written as

$$p_i = P p_w \tag{1}$$

where p_w and p_i are the 3D and image points represented by homogeneous 4-vector and 3-vector. The 3×4 homogeneous matrix *P*, which is unique up to a scale factor, is called the camera perspective matrix. It can be further decomposed into the intrinsic camera parameter matrix and the relative pose of the camera, as shown in Eq. (2).

$$P = K[Rt] \tag{2}$$

The 3×3 matrix *R* and 3×1 vector *t* are the relative orientation and translation with respect to the world coordinates, respectively. The intrinsic camera parameters *K* of the camera is a 3×3 matrix and usually model as

$$K = \begin{bmatrix} s_x f & \gamma & u_0 \\ 0 & s_y f & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$
(3)

where (u_0, v_0) is the principle point of the camera (the intersection of the optical axis with the image plane), γ is a skew parameter related to the characteristic of the CCD array, *f* is the focal length and s_x and s_y are the scale factor [pixel/mm] in direction of x and y axis.

In this paper, the two step calibration method [12] is adopted. The method includes the distortion calibration and the camera calibration. The distortion calibration is used to solve the uneven spacing between the dots in the camera image.

4.2 Extrinsic Transformation

After obtaining the intrinsic matrix, the relation from the camera coordinate to the camera screen coordinates can be built. Then, the relative orientation and the translation from the marker coordinates to the camera coordinates must be computed in real-time. p_m is a 3D point in the marker coordinates and p_c is a 3D point in the camera coordinates, as shown in Eq. (4).

$$p_c = [Rt]p_m \tag{4}$$

Then, the camera perspective matrix can be obtained by multiplying the intrinsic and extrinsic matrixes. Initially, a size-known square marker is used as a base of the coordinates, as shown in Fig. 2.



Fig. 2: The establishment of the coordinate systems.

4.3 Geometry Calibration

Most calibrations for OSTD system with AR are based on [16], [17]. In addition, two parts in the proposed system are different from the previous literatures. First, the virtual eye and the tracking camera are the same machine. Hence, it can be assumed that the FOV of the virtual eye and the tracking camera are identical. Secondly, the microdisplay(LCoS) in the proposed system is smaller than others. Although the image projected from microdisplay is magnified via the specific lens, the image still can't cover the virtual eye. Hence, there is not correspondence between the depth of the marker viewed by the virtual eye and detected by the tracking camera. There is a linear scale due to the same FOV for the virtual eye and the tracking camera.

Fig. 3 illustrates the main components, the associated coordinate systems and the extrinsic transformations of the OSTD system in the virtual world. All the coordinate systems used in this paper are right-handed. In the virtual world, the key references are the marker coordinates $X_m Y_m Z_m$, the tracking camera coordinates $X_c Y_c Z_c$, the virtual eye coordinates $X_e Y_e Z_e$, the tracking camera screen coordinates $X_{cu} Y_{cv}$, the virtual eye screen coordinates $X_{cu} Y_{ev}$ and the OSTD screen coordinates $X_{ou} Y_{ov}$.



Fig. 3: illustration of the virtual world components and transformations.

Given a 2D projection p_c on the viewing plane of the tracking camera screen coordinates, its 3D point P_c in the tracking camera coordinates can be computed by K_c . Then, the $[Rt]_{c-m}^{-1}$ times P_c makes P_m

Once P_m is obtained, its 3D projection P_e in the virtual eye coordinates can be computed by $[Rt]_{e \leftarrow c}$. Its 2D projection p_e on the viewing plane of the OSTD screen coordinates is given by

$$p_o = k_o [Rt]_{e \leftarrow c} [Rt]_{c \leftarrow m} P_m \tag{5}$$

The $[Rt]_{e^{-c}}$ is used to transform the 3D point from the tracking camera coordinates to the virtual eye coordinates. Then, the k_o is used to map the 3D point in the virtual eye coordinates to the OSTD screen coordinates. The Eq. (6) represents the stereo vision concept.

$$\begin{cases}
P_e = R_{e \leftarrow m} P_m + t_{e \leftarrow m} \\
P_c = R_{c \leftarrow m} P_m + t_{c \leftarrow m} \\
P_e = R_{e \leftarrow c}^T (P_c - t_{e \leftarrow c})
\end{cases}$$
(6)

Thus, the $[Rt]_{e_c}$ can be calculated by Eq. (6) as given in Eq. (7).

$$\begin{cases} R_{e\leftarrow c}^{T} = R_{e\leftarrow m} R_{c\leftarrow m}^{-1} \\ t_{e\leftarrow c} = t_{c\leftarrow m} - R_{e\leftarrow m}^{-1} R_{c\leftarrow m} t_{e\leftarrow m} = t_{c\leftarrow m} - R_{e\leftarrow c} t_{e\leftarrow m} \end{cases}$$
(7)

The next step is to compute the exact K_o . Compared to the previous method, a simplified method is proposed to solve this problem. The W_{eu} and the W_{ou} are the widths that the edge of the marker is projected to the virtual eye screen coordinates and the OSTD screen coordinates respectively, as described in Fig. 4.



Fig. 4: The edge of the marker is projected to the virtual eye screen coordinates and the OSTD screen coordinates respectively.

Therefore, the scale factor can be computed as:

$$\frac{W_{eu}}{W_{ou}} = \frac{Z_2}{Z_1} = Scale \quad factor \tag{8}$$

In order to magnify the exact size for virtual model, the scale factor is used to modify the original transformation

$$\begin{cases}
M = [Rt]_{e \leftarrow c} [Rt]_{c \leftarrow m} = M_2 \times M_1 = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \\
M_1 = [Rt]_{c \leftarrow m} = \begin{bmatrix} R_1 & t_1 \\ 0 & 1 \end{bmatrix} \\
M_2 = [Rt]_{e \leftarrow c} = \begin{bmatrix} R_2 & t_2 \\ 0 & 1 \end{bmatrix}
\end{cases}$$
(9)

where *M* represents the transformation from the marker coordinates to the virtual eye coordinates. $M_{_{I}}$ and $M_{_{2}}$ are the transformations from the marker coordinates to the tracking camera coordinates and from the tracking camera coordinates to the virtual eye coordinates respectively. Then, the t_{1z} multiplied by the scale factor is the modified depth. The *t*_i is replaced by the modified depth as follows:

$$M' = \begin{vmatrix} t_x \\ R & t_y \\ t_{1z} \times Scale & factor \\ 0 & 1 \end{vmatrix}$$
(10)

M' is the modified transformation. After obtaining the M', the image projected from microdisplay can cover the virtual eye screen coordinates. Hence, Eq. (5) can be modified as

$$p_o = k_e M' P_m = k_c M' P_m \tag{11}$$

The K_e is the same as the K_e due to the same FOV for the tracking camera and the virtual eye. Finally, the virtual objects can be superimposed precisely onto the real objects. In the future, the FOV for the tracking camera can be adjusted manually if the virtual eye is replaced by the human's eye.

5 IMPLEMENT AND RESULT

In this paper, the oral implant training system has been implemented in C++ and Visual Toolkit (VTK). Some tests were run on an AMD Athlon 64 X2 Core Processor 5200+, 2.61 GHz CPU with 2GB memory and a NVidia GeForce 7300GT graphic card with 256MB memory. The type of the microdisplay is LCoS (Kopin), the size is 0.26in, the resolution is 430x240 and the bright is 150fL. The webcam we use is *Logitech*[®] Webcam *C905*.

5.1 System Architecture

The OSTD system is composed of the microdisplay, lens, beam-splitter, tracking camera and virtual eye, as shown in Fig. 5(a). As you can see, the eye relief is set 12mm. The front and back views of the OSTD system are illustrated in Fig. 5(b) and (c). Generally, the tracking camera can be located in artificial pose depending on the user's need. In addition, the illumination can be adjusted depending on the environment or the user's need and the maximum is 150fL.



Fig. 5: The OSTD system: (a) Architecture, (b) Front view, (c) AR visualization.

5.2 Calibration

The related parameters can be obtained after the calibration, as described in Tab. 1. In the proposed system, the tracking camera and visual eye are located in a bar. Thus, the rotation matrix of the extrinsic transformation between the tracking camera and visual eye is approximated an identified matrix. Then, the initial registration of the virtual objects and real scene is set manually. The intrinsic

camera parameters can be used to set the FOV for the tracking camera and the space for virtual eye. After obtaining the intrinsic camera parameters, the extrinsic transformation can be computed by the calibration. Then, the scale factor also can be calculated by the geometry calibration. Finally, the magnified virtual object from the microdisplay can registered interactively in the real scene.

Name	Parameters				
FOV	54.462(H), 53.673(V)				
Intrinsic	[520.7812	-0.194	4 330.9274]
camera		0	520.610	07 237.5057	
parameters		0	0	1	
Extrinsic	[1	- 8.673	6e-019	0	-121.241
transformation	2.6020e-0	018	1	-1.1652e-017	2.0144
between two	0	1.1926	5e-017	1	- 20.8211
cameras	0		0	0	1
Scale	4.71428				

Tab. 1 The parameters for OSTD system.

5.3 Efficiency

In the proposed OSTD system, the input includes the surface and volume data. The surface data is composed of the nerve canal and the surgical guide, and the volume data includes the CT or MRI data. As you can see, the resolution and the volume data can be downed obviously for the efficiency in Fig. 6. The main reason is that the 3D texture rendering, the transfer function and the shading are used to show a more real effect for volume data. Moreover, the above computation can increase when the resolution is raised. In the other hand, the surface data can't affect the efficiency too much.



Fig. 6: The efficiency in the OSTD system.

5.4 Error Estimation

Generally, the distance between the surgeon and the patient is about 350 mm. Hence, the distance between the marker and virtual eye is set 400mm in the OSTD system and the size of the marker is set 5cm. Then, an accuracy XY table is used to move left or right. We defined the corner of the marker as A, B, C and D, as shown in Fig. 7.



Fig. 7: The simulation in the proposed OSTD system.

The Initial position for 4 corners of the marker is set 0. Once the XY table is executed, the left direction is negative and the right direction is positive. The error estimation in 400mm is illustrated in Fig. 8(a) and (b).



Fig. 8: The error estimation in the OSTD system: (a) The error in x direction, (b) The error in y direction.

The reasons include the environmental light, lens distortion and the error from the calibration. For environmental light, the better image processing can be a solution. The more robust calibration, in the other hand, can be a solution for the lens distortion. Furthermore, the customized lens design is also can be used to solve the lens distortion. Finally, the error from the calibration can be computed more times or replace by another algorithms [10], [11] to improve the precision.

5.5 Result

The OSTD system can optically superimpose the virtual objects upon the existing scene. The tracking camera is used to detect the marker interactively. Thus, the transformation among the marker, the tracking camera and the virtual eye can be obtained by the calibration. The transformation can be used to locate the visual objects such as the nerve canal, as described in Fig. 9.



Fig. 9: The tracking effect for nerve canal (yellow): (a) Normal, (b) Over work space.

6 CONCLUSION

In this paper, we integrated the hardware design, the calibration, the tracking technique and AR visualization to develop an oral implant training OSTD system. In addition, a geometry calibration isn't only proposed to simplify the previous research, but also is able to have the same visualization.

Only the virtual implants are visualized in jawbone to guide the surgeon to do the surgery or training in [14], [15]. In contrast, novices and surgeons can see the virtual objects, such as implant, the nerve canal and the jawbone, superimposed upon the real model using the marker by the OSTD system. Therefore, the users can only pay attention to one screen during the training or surgery. The screen in the proposed OSTD system can't only show completely the auxiliary information for patient, but also visualize directly the depth perception between the implant and the nerve in the AR visualization.

However, the error should be improved to satisfy the surgical requirement. Moreover, some parallel processing tools, such as OpenMP and GPUs, can be used to increase the efficiency. In the

future, we plan to design an optical see-through HMD. The OSTDHMD can be integrated into the oral implant training or surgery.

REFERENCES

- [1] Misch, C. E.: Contemporary implant dentistry, 2nd ed., Mosby, St Louis, 1998.
- [2] Fisher, R.: Head-mounted projection display system featuring beam splitter and method of making same, U.S. Patent 5 572 229, Nov. 5, 1996.
- [3] Kijima, R.; Ojika, T.: Transition between virtual environment and workstation environment with projective head-mounted display, IEEE Virtual Reality Annu. Int. Symp., 1997, 130-137. DOI:10.1109/VRAIS.1997.583062
- [4] Robinson, J. D.; Franxisco, S.; Gatos, M. L.: Video headset, U. S. patent 5 696 521, 1996.
- [5] Hua, H.; Gao, C.; Biocca, F.; Rolland, J. P.: An ultra-light and compact design and implementation of head-mounted projective displays, IEEE Virtual Reality, 2001, 174-182. DOI: 10.1109/VR.2001.913784
- [6] Hua, H.; Brown, L.; Gao, C.: SCAPE: supporting stereoscopiccollaboration in augmented and projective environments, IEEEComputer Graphics and Application, 24(1), 2004, 66-74. DOI:10.1109/MCG.2004.1255811
- [7] Cakmakci, O.; Rolland, J.: Head-Worn Displays: A Review, IEEE/OSA Journal of Display Technology, 2(3), 2006, 199-216. DOI:10.1109/JDT.2006.879846
- [8] Cakmakci, O.; Ha, Y.; Rolland, J.: A compact optical see-through head-worn display with occlusion support, International Symposium on Mixed and Augmented Reality, 2004, 16-24. DOI:10.1109/ISMAR.2004.2
- [9] Sielhorst, T.; Feuerstein, M.; Navab, N.: Advanced medical displays: a literature review of augmented reality, IEEE Transactions on Medical Imaging, 4(4), 2008, 451-466. DOI:10.1109/JDT.2008.2001575
- [10] Tsai, R.: A versatile camera calibration technique for high accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses, IEEE Int. J. Robot. Autom., 3(4), 1987, 323-344. DOI:10.1109/JRA.1987.1087109
- [11] Zhang, Z.: A flexible new technique for camera calibration, IEEE Trans. Pattern Anal. Mach. Intell., 22, 2000, 1330-1334. DOI:10.1109/34.888718
- [12] Hirokazu, K.; Mark, B.: Marker tracking and HMD calibration for a video-based augmented reality conferencing system, Workshop on Augmented Reality, 1999, 84-93. DOI:10.1109/IWAR.1999.803809
- [13] Rhienmora, P.; Gajananan, K.; Haddawy, P.; Suebnukarn, S.; Dailey, M.; Supataratarn, E.; Shrestha, P.: Haptic Augmented Reality Dental Trainer with Automatic Performance Assessment, International Conference on Intelligent User Interfaces, 2010. DOI:10.1145/1719970.1720054
- [14] Wanschitz, F.; Birkfellner, W.; Figl, M.; Patruta, S.; Wagner, A.; Watzinger, F.; Yerit, K.; Schicho, K.; Hanel, R.; Kainberger, F.; Imhof, H.; Bergmann, H.; Ewers, R.: Computer-enhanced stereoscopic vision in a head-mounted display for oral implant surgery, Clinical Oral Implants Research, 13(6), 2002, 610-616. DOI:10.1034/j.1600-0501.2002.130606.x
- [15] Satoshi, Y.; Takafumi, O.; Hirofumi, Y.; Taiji, S.: Augmented reality system for dental implant surgery, Lecture Notes in Computer Science, 5662, 2009, 633-638. DOI:10.1007/978-3-642-02771-0_70
- [16] Owen, B. C.; Zhou, J.; Tang, A., Xiao, F.: Display-relative calibration for optical see-through headmounted displays, IEEE and ACM International Symposium on Mixed and Augmented Reality, 2004, 70-77. DOI:10.1109/ISMAR.2004.28
- [17] Tuceryan, M.; Navab, N.: Single point active alignment method (SPAAM) for optical see-through HMD calibration for Augmented Reality, IEEE and ACM International Symposium on Augmented Reality, 11(3),2002, 149-157. DOI:10.1162/105474602317473213