3D Surface Reconstruction and Registration for Image Guided Medialization Laryngoplasty

Ge Jin[1], Sang-Joon Lee[1], James K. Hahn[1], Steven Bielamowicz[2], Rajat Mittal[3], and Raymond Walsh[4]

[1]Department of Computer Science
[2]Division of Otolaryngology, School of Medicine
[3]Department of Mechanical and Aerospace Engineering
[4]Department of Anatomy and Cell Biology, School of Medicine
The George Washington University, Washington DC, USA
{jinge,bigdolph,hahn@gwu.edu,sbielamowicz@mfa.gwu.edu, mittal@gwu.edu,anarjw@gwumc.edu}

Abstract. The purpose of our project is to develop an image guided system for the medialization laryngoplasty. One of the fundamental challenges in our system is to accurately register the preoperative 3D CT data to the intraoperative 3D surfaces of the patient. In this paper, we will present a combined surface and fiducial based registration method to register the preoperative 3D CT data to the intraoperative surface of larynx. To accurately model the exposed surface area, an active illumination based stereo vision technique is used for the surface reconstruction. To register the point clouds from the intraoperative stage to the preoperative 3D CT data, a shape priori based ICP method is proposed to quickly register the two surfaces. The proposed approach is capable of tracking the fiducial markers and reconstructing the surface of larynx with no damage to the anatomical structure. Although, the proposed method is specifically designed for the image guided laryngoplasty, it can be applied to other image guided surgical areas. We used off-the-shelf digital cameras, LCD projector and rapid 3D prototyper to develop our experimental system. The final RMS error in the registration is less than 1mm.

Key words: Image Guided Surgery, 3D Reconstruction, Registration

1 Introduction

It is estimated that 7.5 million people in the United States have a voice disorder, and about 1/3 of new patients with voice disorders are diagnosed with vocal fold paresis or paralysis. Vocal cord paralysis and paresis are debilitating conditions leading to difficulty with voice production. The alterations in voice production are usually severe enough to impede the individual's ability to work and to conduct normal social interactions. Medialization laryngoplasty is a surgical procedure designed to restore the voice in patients by implanting a uniquely configured structural support lateral to the paretic vocal fold through a window cut in the thyroid cartilage of the larynx. Currently, the surgeon relies on experience and intuition to place the implant in the desired location, therefore it is subject to a significant level of uncertainty. Window placement errors of up to 5mm in the vertical dimension are common in patients admitted for revision surgery. The failure rate of this procedure is as high as 24% even for experienced surgeons [1]. An intraoperative image guided system will help the surgeon to accurately place the implant at the desired location.

The image guided technology has been successfully applied to various medical domains. However, to our knowledge, image guided techniques have not been applied to the medialization laryngoplasty. The biggest obstacles come from (1) registering the geometry of the delicate anatomy of thyroid cartilage during the surgery to the preoperative 3D CT data (2) introducing minimal intrusion or modifications to the current surgical practices and (3) implementing with only a moderate increase in the additional equipment. In this paper, we will concentrate on the registration of preoperative 3D CT data to the intraoperative 3D surfaces of thyroid cartilage.

Our proposed image guided system will use the anatomical and geometric landmarks and points to register intraoperative 3D surface of thyroid cartilage to the preoperative 3D radiological data. The proposed approach has three phases. First, the laryngeal cartilage surface is segmented out from the preoperative 3D CT data. Second, the surface of the exposed laryngeal cartilage during the surgery is reconstructed intraoperatively using stereo vision and structured light based surface scanning. The surgical area has non-uniform color and textures, so we take one full-lit image and non-lit image to distinguish the shadow from the light receiving areas and calculate the illumination change map. Third, the two geometries are registered using shape priori based ICP matching. Currently the proposed technqiue has only been applied in a laboratory environment on phantom models. The proposed approach has several advantages over alternative approaches: the combination of stereo vision and structured light surface scanning is capable of tracking the fiducial markers, reconstructing the surface of laryngeal cartilage and matching the preoperative and postoperative surfaces for registration purposes. The computer vision based approach can be applied to delicate areas like laryngeal cartilage with no danger of causing physical damage.

2 Background

Registration in image guided procedures can be classified into three categories based on the fiducial markers: extrinsic invasive (bone affixed markers)[2][3][4], extrinsic noninvasive (skin affixed markers)[5] and intrinsic markers [6]. In the case of laryngoplasty, the bone fixed fiducial markers would make potentially damage to the thin laryngeal cartilage. While, the skin affixed markers will move significantly relative to the laryngeal cartilage. Intraoperative medical imaging system can be used for the multi-modal image registration in image guided surgery [7][8]. However, for the medialization laryngoplasty, this will modify the current surgical procedure and increase the medical cost by introducing additional medical equipment. So, in our system, we will use the intrinsic markers for the registration.

The structured light based surface reconstruction system can be classified into three categories: time-multiplexing, spatial neighborhood and direct coding. Spatial neighborhood [9][10] and direct coding [11] methods are relatively fast and capable of measuring dynamic surfaces. However, the bandwidth of projector and quantization error introduced by the CCD camera will make the color and neighborhood based methods less accurate than time multiplexing methods. Time-multiplexing is a way to encode the pixel information in the temporal domain. Posdamer and Altschuler [12] first proposed a 3D surface measurement method with binary coded light pattern. Inokuchi [13] further improved the coding scheme using gray code to make the code-word robust to the noise. Recently, Gühring [14] combined the gray code light pattern and line shifting to reconstruct highly accurate 3D surface model. For our experimental framework, since the primary goal is to reconstruct accurate 3D surface for registration, we used sub-pixel accuracy line shifting method to reconstruct the 3D surfaces.

The global alignment of multiple 3D point sets or surfaces has been well studied in the field of 3D model acquisition area. ICP (Iterative Closest Point) algorithm was introduced to geometrically align two similar geometric models [15][16]. A new geometric transform matrix is calculated by minimizing the MSE (Mean Square Error) between the closest point pairs. Horn [17] described a closed form solution for the quaternion calculation from the closest point pairs. Kd-tree [15], and approximated kd-tree [18] is used to accelerate the closest point searching process. Recently, sub-sampling scheme from the geometric data, closest point searching method, rejection of outliers and error minimization method are used to compare various ICP algorithms [19]. In our case, the number of points from the preoperative CT and structured light based surface scanning are relatively small (about 3000 points), so we used all the point samples during the ICP matching process. For the closest point searching, a balanced kd-tree is used to accelerate the searching speed. We used the closed form solution from [17] to calculate the unit quaternion rotation vector, and rejected the outliers from sample space if the closest distance is longer than 2 times of mean closest distance. For our case, the shape features of the laryngeal cartilage will be a good candidate for fast initial pose estimation. We used two crossing planes to calculate the initial pose for fast shape matching.

3 Image Guided Medialization Laryngoplasty

The work flow of our surface registration process is shown in Figure 1 Left. There are three major steps: surface extraction from preoperative CT data, structured light based intraoperative surface reconstruction and ICP based point clouds registration.

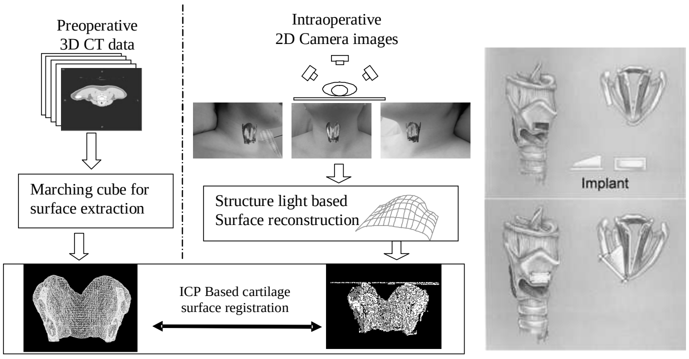

Fig. 1. Left:Work flow of surface registration, Right: Medialization laryngoplasty

3.1 Medialization Laryngoplasty

The medialization laryngoplasty (Figure 1. Right) procedure is the thyroplasty procedure, which is aimed at medializing the membraneous aspects of the vocal fold. A thyroplasty implant is a patient-specific device that must be properly aligned in reference to the underlying vocal fold and have a size and shape such that it medializes the vocal fold and alters the vibratory characteristics of the vocal fold to a state that most closely resembles that of the uninjuried vocal cord.

3.2 Surface Extraction from CT Data and Phantom Model Construction

We used visible human CT data set from NIH for our experiment. The thyroid cartilage surface is extracted using marching cube algorithm [20]. The extracted triangular mesh is rendered in wire frame, flat shading and texture mapping (Figure 2. Left). The extracted 3D surface model is converted to a solid CAD

Fig. 2. Left: Iso-surface extraction from the CT data, Right: Phantom model

model and sent to the 3D prototyping device (Stratasys FDM 3000). The prototyper is capable of constructing a 3D phantom model with the accuracy of 0.1mm (Figure 2. Right).

3.3 Structured Light based Intraoperative Surface Reconstruction

In the surgical environment, the area of scanning has non-uniform color and texture. Threshold based image segmentation could not provide accurate structured light pattern. Similar with photometric calibration of projector and camera, one full-lit image and non-lit image are used to distinguish the shadow from the light receiving areas. Another difficulty in reconstructing laryngeal cartilage surface is the small size of anatomical structure. Usually, the structured light based surface scanning is applied to rather big structures like: human faces, statues and so on. The fully exposed larynx is about 90x50x50 mm. If the distance from camera and laryngeal cartilage is larger than certain distance, a regular camera with standard resolution (640X480) could not provide enough resolution. Since the camera should not disturb the surgical procedures, there is a minimum distance requirement for the surgical environment. With this restriction in mind, a higher resolution camera is required to provide enough accuracy for surface reconstruction.

Structured light based surface reconstruction requires light projection device (LCD projector) and one or more cameras. In our case, we used LCD projector with two cameras. Since the camera to camera calibration has higher accuracy than camera to projector calibration, we only calibrated the camera pairs and used the LCD for illumination purpose. For the camera calibration, we used the planar homography based camera calibration method from [21].

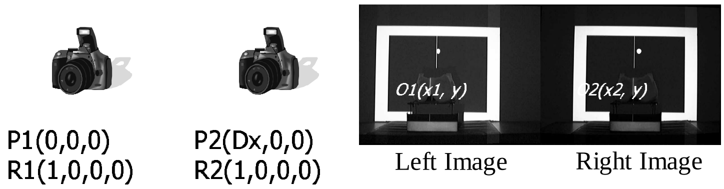

Fig. 3. Camera parameter after rectification and rectified images

After calibration, the images from two cameras are rectified to align the horizontal scan lines. After rectification, the searching of pixel correspondence has been reduced to one dimension. Furthermore, the camera internal and external parameters are simplified. In figure 3, the P_1, P_2 is the camera position vector and the R_1,R_2 is the cam-era rotation matrix represented by the quaternion. In equation 1, the M_1, M_2 is the pinhole camera projection matrix. If we find the pixel correspondence (O_1 and O_2) in left and right images, we can calculate the real 3D position of the pixel in camera coordinate system by solving the linear equations shown on (1).

$$M_1 = \begin{bmatrix} f \; 0 \; C_{x1} \\ 0 \; f \; C_y \\ 0 \; 0 \; 1 \end{bmatrix} \; ; \; M_2 = \begin{bmatrix} f \; 0 \; C_{x2} \\ 0 \; f \; C_y \\ 0 \; 0 \; 1 \end{bmatrix} \; ; \; \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = M_1^{-1} \begin{bmatrix} x_1 \\ y \\ 1 \end{bmatrix} = M_2^{-1} \begin{bmatrix} x_2 \\ y \\ 1 \end{bmatrix} \tag{1}$$

From the above equation, we can easily notice that the sub-pixel accuracy in pixel correspondence is the most critical issue in 3D reconstruction. If we only have the pixel level accuracy, the recovered depth value will not be continuous. So, we experimented with sub-pixel accuracy line shifting method to reconstruct the surface of thyroid cartilage phantom model. First, we searched for the peak intensity along the horizontal scan line. Then, we used the 7 nearby pixels for the sub pixel peak detection and calculated the 2nd order derivative for 5 pixels around the detected peak. The sub-pixel intensity peak is calculated with zero-crossing of 2nd order derivatives. In the sub-pixel accuracy peak detection, the camera shutter speed, film sensitivity and projector focus simultaneously affect the peak detection result. The preliminary experiment has indicated that: the focus of the beam projector should focus on the laryngeal cartilage surface to provide maximum intensity variation and the camera shutter speed needs to be adjusted to capture the sub-pixel illumination change. If the image is over-exposed, the peak of the light strip will spread over several pixels and as a result the detected peak is not accurate.

3.4 ICP based Point Clouds Registration

To register the 3D surface from preoperative CT data and the point clouds from the structured light based surface reconstruction, we need to preprocess the 3D surface from CT. The point clouds from the computer vision are only the front side of the thyroid cartilage. Therefore, we need to remove the back facing polygons from the preoperative CT surface so that the back facing polygons do not affect the registration result. We used the surface normal to separate the front facing and back facing polygons. In order to reduce the searching time for the closest point matching, we used balanced k-d tree. A kd-tree is a space-partitioning data structure for organizing points in a k-dimensional space. It uses splitting planes that are perpendicular to one of the coordinate system axes (Figure 4. right).

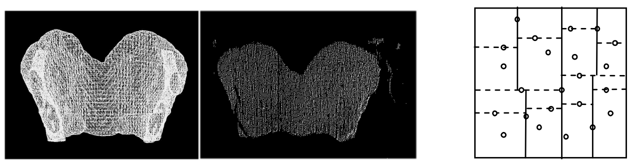

Fig. 4. Left: 3D model from CT and from structured light Right: 2D kd-tree

We used the point to point euclidian distance as our closest point matching criteria. After the calculation of closest point, we rejected the outliers from sample space if the closest distance is longer than 2 times of mean closest distance. The minimization of mean square error is only considered on inliers.

Suppose M and D are 3D point sets from preoperative and intraoprative stages, the goal of ICP algorithm is to find the optimal rotation and translation that minimize equation $E(R,T) = \sum_{i \in M} \sum_{j \in D} \| m_i - (R \cdot d_j + T) \|$. We used unit quaternion Q(q_0,q_1,q_2,q_3) to represent the rotation matrix R.

$$R = \begin{bmatrix} q_0^2 + q_1^2 + q_2^2 + q_3^2 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_1q_2 + q_0q_3) & q_0^2 + q_2^2 - q_1^2 - q_3^2 & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) & q_0^2 + q_3^2 - q_1^2 - q_2^2 \end{bmatrix} \quad (2)$$

The closed form solution from [17]'s work is used to calculate the quaternion vector. To determine the rotation vector, we first subtract center of mass position from each point clouds set. A covariance matrix N is calculated using the equation 3 where $S_{xx} = \sum \sum m^*_{ix} * d^*_{jx}$. The new quaternion vector Q is the eigenvector of largest positive eigen value of N.

$$N = \begin{bmatrix} S_{xx} + S_{yy} + S_{zz} & S_{yz} - S_{zy} & S_{zx} - S_{xz} & S_{xy} - S_{yx} \\ S_{yz} - S_{zy} & S_{xx} - S_{yy} - S_{zz} & S_{xy} + S_{yx} & S_{zx} + S_{xz} \\ S_{zx} - S_{xz} & S_{xy} + S_{yx} & S_{yy} - S_{xx} - S_{zz} & S_{yz} + S_{zy} \\ S_{xy} - S_{yx} & S_{zx} + S_{xz} & S_{yz} + S_{zy} & S_{zz} - S_{yy} - S_{xx} \end{bmatrix} \quad (3)$$

The original ICP algorithm calculates the translation vector using the difference in the center of mass point. This is correct when the center of mass points in preoperative and intraoperative surfaces are close. But, in our case, the surface points from the preoperative CT consist of points that are not exposed to the camera. Furthermore, the structured light based reconstruction stage also consists of noise points. So we separated the translation calculation from rotation calculation stage. For the translation, we used the summed average of displacement vector of matched closest point pairs. Suppose [X_{M1},Y_{M1},Z_{M1}] and [X_{M2},Y_{M2},Z_{M2}] is closest matching point pair from CT and structured light based reconstruction, the new translation vector is calculated using $[T_x, T_y, T_z] = [\sum_n (X_{M1} - X_{M2})/n, \sum_n (Y_{M1} - Y_{M2})/n, \sum_n (Z_{M1} - Z_{M2})/n]$.

The initial pose estimation will greatly affect the convergence speed and the correctness of the final result. Unlike original ICP based shape matching, for the medical image registration, the ground truth target mesh is known. The shape features of the laryngeal cartilage will be a good candidate for fast initial pose estimation. One important observation is that the laryngeal cartilage surface can be approximated by two crossing planes (Figure.5). Point to plane distance ($\frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}$)is used to estimate the plane equation ($ax + by + cz + d = 0$). Minimizing the sum of squared distance from point to plane will provide a plane equation that best fit the point clouds. The center of mass of point clouds is projected to the plane to provide the unique matching point on the plane. The

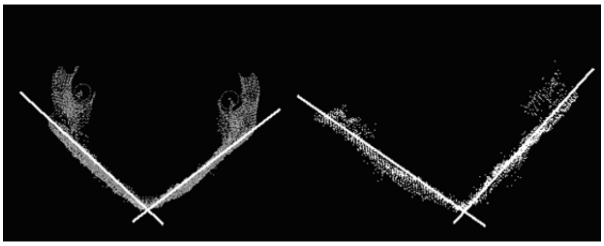

Fig. 5. Approximation of the larynx with two crossing planes

SVD based closed form solution is used to approximate the plane equation. The plane equation is the vector associated with smallest singular value (Equation 4). Geometric description based on initial shape approximation will provide a close initial pose estimation for the ICP method.

$$SS\_Dist \;=\; \sum_{v \in M} \frac{(ax+by+cz+d)^2}{a^2+b^2+c^2};\; SVD:\; D = \frac{1}{N}\sum_{v \in M} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}^T \qquad (4)$$

4 Experiment and Result

We used Intel Xeon 3.2GHz Workstation with 2GB memory for our experiment. For the structured light based surface reconstruction, we have experimented with two Logitec Quickcam cameras, Nikon D70s digital cameras and LCD projector. The surface reconstruction result with sub-pixel accuracy line shifting is show on figure 6 and figure 7 left.

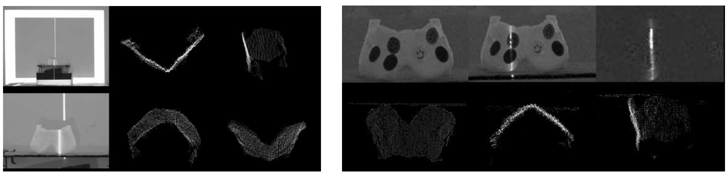

Fig. 6. Surface reconstruction result for phantom model

To mimic the real situation, color dotted phantom model and animal bone are used for the experiment. The illumination change value is calculated by dividing the illuminated image with non-lit image. In ICP based point clouds registration, the computation time for kd-tree construction is 94 ms. Shape priori based ICP matching takes 515 ms to match the two point clouds with RMS error 0.9mm.

The original ICP method with the same RMS error takes 4 sec. The final mean square error in two matched point clouds is 0.899mm and the registration result is shown on figure 7 right.

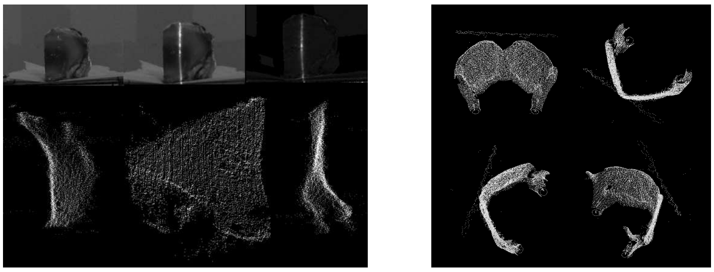

Fig. 7. Left: Animal bone surface reconstruction Right: Shape priori based ICP matching

5 Conclusions and Future Works

In this paper, we proposed an image guided system for the medialization laryngoplasty. To our knowledge, this is the first attempt to apply image-guided techniques to the medialization laryngoplasty. Due to the delicate nature of thyroid cartilage surface, we could not directly use the fiducial marker based optical tracking system for the image registration. Instead, we introduced a structured light based stereo vision system that could be used for 3D surface reconstruction and feature tracking. We used the sub-pixel accuracy line shifting for the 3D reconstruction. To mimic the real situation, color dotted phantom model and animal bone is used for experiment. Instead of using the absolute intensity value, the illumination change map is used for light peak detection. To match the 3D surface from preoperative CT and the point clouds from structured light based reconstruction, we proposed a shape priori based initial pose estimation combined with the ICP algorithm to register two sets of point clouds. The mean square error of ICP based registration is less than 1.0mm. Our experimental framework can be applied to other image guided applications. For the future work, we will use the registration result and the projective texture mapping techniques to render the preoperative thyroid cartilage surface and visualize the important anatomical structures (vocal fold and airway lumens) beneath the thyroid cartilage surface. This work is supported by a grant from the National Institute of Health (No. R01-DC007125-0181) for developing computer-based tools for medialization laryngoplasty.

References

1. T. D. Anderson, J. R. Spiegel, R. T. Sataloff: Thyroplasty revisions: frequency and predictive factors. Journal of Voice, Vol. 17(3): (2003) 442–448
2. P.J Kelly: Computer-assisted stereotaxis: new approaches for the management of intracranial intra-axial tumors. Neurology, Vol. 36(4): (1986) 535–541
3. K. P. Gall, L. J. Verhey, Wagner M: Computer-assisted positioning of radiotherapy patients using implanted radioopaque fiducials. Medical physics. Vol. 20(4): (1993) 1153–1159
4. Maurer CR Jr, Fitzpatrick JM, Wang MY, Galloway RL Jr, Maciunas RJ, Allen GS: Registration of head volume images using implantable fiducial markers. IEEE Trans Med Imaging Vol. 16(4): (1997) 447–462.
5. Darabi K, Grunert P, Perneczky A: Accuracy of intraoperative navigation using skin markers. In: CARS 1997. Berlin, Springer-Verlag, (1997) 920–924.
6. M. I. Miga, T. K. Sinha, D. M. Cash, R. L. Galloway, and R. J. Weil: Cortical Surface Registration for Image-Guided Neurosurgery Using Laser-Range Scanning IEEE Transaction of Medical Imaging, Vol. 22(8): (2003) 973–985
7. Wells, W. M., Viola, P., Atsumi, H., Nakajima, S., Kikinis, R.: Multi-modal volume registration by maximization of mutual information. Medical Image Analysis, Vol. 1(1): (1996) 35–51
8. H. Livyatan, Z. Yaniv, L. Joskowicz: Gradient-Based 2D/3D Rigid Registration of Fluoroscopic X-ray to CT. IEEE Trans. Med. Imaging Vol. 22(11): (2003) 1395–1406
9. K. L. Boyer, A. C. Kak: Color-encoded structured light for rapid active ranging. IEEE Trans. Pattern Anal. Mach. Intell., Vol. 9(1): 1987 14–28
10. J. Salvi, J. Batlle, E. Mouaddib: A robust-coded pattern projection for dynamic 3D scene measurement. Pattern Recogn. Lett., Vol. 19(11): (1998) 1055–1065
11. C. J. Davies, M. S. Nixon: A hough transform for detecting the location and orientation of 3-dimensional surfaces via color encoded spots: IEEE Trans. Systems, Man and Cybernetics, Vol. 28(1): (1998) 90–95
12. J. L. Posdamer, M. D. Altschuler: Surface measurement by space-encoded projected beam systems. Computer Graphics and Image Processing Vol. 18(1): (1982) 1–17
13. S. Inokuchi, K. Sato, F. Matsuda: Range imaging system for 3-D object recognition. In: Proc. of the International Conference on Pattern Recognition, (1984) 806–808
14. J. Gühring: Dense 3-d surface acquisition by structured light using off-the-shelf components. In SPIE Photonics West Videometrics, Vol. 4309: (2001) 220–231
15. Besl P. J., Mckay N. D.: A method for registration of 3d shapes. IEEE Trans. Pattern Anal. Mach. Intell., Vol. 14(2): (1992) 239–256
16. Chen Y., Medioni G.: Object modelling by registration of multiple range images. Image Vision Comput., Vol. 10(3): (1992) 145–155
17. B.K.P. Horn: Closed-form solution of absolute orientation using unit quaternions. Journal of the Optical Society of America, Vol. 4(4): (1987) 629–642
18. Greenspan, M., Yurick, M.: Approximate k-d tree search for efficient ICP. In Proc. 3D Digital Imaging and Modeling, (2003)442–448
19. S. Rusinkiewicz, M. Levoy: Efficient Variants of the ICP Algorithm. In 3rd Int. Conf. on 3D Digital Imaging and Modeling, (2001)
20. W. E. Lorensen and H. E. Cline: Marching cubes: a high resolution 3D surface construction algorithm. In Proc. SIGGRAPH 1987 Vol. 21(4): (1987) 163–170
21. Zhengyou Zhang: Flexible camera calibration by viewing a plane from unknown orientations. In 7th IEEE Int. Conf. on Computer Vision, (1999) 666–673