Active Illumination based 3D Surface Reconstruction and Registration for Image Guided Medialization Laryngoplasty

Ge Jin[*a], Sang-Joon Lee[a], James K. Hahn[a], Steven Bielamowicz[b], Rajat Mittal[c], Raymond Walsh[d]
[a]Department of Computer Science;
[b]Division of Otolaryngology, School of Medicine;
[c]Department of Mechanical and Aerospace Engineering;
[d]Department of Anatomy and Cell Biology, School of Medicine;
The George Washington University, Washington DC, USA 20052

ABSTRACT

The medialization laryngoplasty is a surgical procedure to improve the voice function of the patient with vocal fold paresis and paralysis. An image guided system for the medialization laryngoplasty will help the surgeons to accurately place the implant and thus reduce the failure rates of the surgery. One of the fundamental challenges in image guided system is to accurately register the preoperative radiological data to the intraoperative anatomical structure of the patient. In this paper, we 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, a structured light based stereo vision technique is used for the surface reconstruction. We combined the gray code pattern and multi-line shifting to generate the intraoperative surface of the larynx. 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. 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.

Keywords: Image Guided Intervention, 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 fold 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 voice in patients with glottal insufficiency due to incomplete vocal fold adduction during voicing. The underlying objective of the procedure is to implant a uniquely configured (i.e. patient specific) structural support lateral to the paretic vocal fold through a window cut in the thyroid cartilage of the larynx. The implant provides vocal fold support by placing the vocal fold into a more medial position, i.e. medialization. The vocal folds are located deep to the thyroid cartilage, anatomical landmarks of the external laryngeal skeleton do not predict the exact location of the underlying vocal folds. The location is approximated by pre- and intraoperative study of the patient's CT scans, visual inspection of the thyroid cartilage, and the geometry of the vocal folds as seen through a trans-nasal endoscopic image. Therefore, it is subject to a significant level of uncertainty. Window placement errors of up to 5mm in the vertical dimension are not uncommon in patients admitted for revision surgery. The objective of this interdisciplinary research is to resolve the issue of accurate implant placement by providing computer assisted surgery tools to improve the outcome of the procedure and to reduce the revision rate. The intraoperative image-guided system is developed to allow the surgeon to accurately place the implant at the desired location. During the operation, the preoperative laryngeal surface from 3D CT data is registered with the intraoperative surface of the larynx.

[*] jinge@gwu.edu; phone 1 202 994-9492; fax 1 202 994-4875;

Medical Imaging 2007: Visualization and Image-Guided Procedures, edited by Kevin R. cleary, Michael I. Miga,
Proc. of SPIE vol. 6509, 650908, (2007) . 1605-7422/07/$18 - doi:10.1117/12.709612

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. To reduce the surface scanning time, we combined the gray-code pattern with multi-line shifting approach. To resolve the ambiguity in multiple line-shifting, we used the dynamic programming method to maximize the sequence matching cost. Third, the two geometries are registered using shape priori based ICP matching. Currently the proposed technique 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. RELATED WORKS

Efforts to correct the voice deficits by surgical manipulation of the laryngeal framework were initially conceived in 1974 by Dr. Nobuhiko Isshiki[1], whose pioneering work revolutionized the management of vocal fold paralysis. Laryngeal framework surgery was modified and popularized in the U.S. by Koufman[2] who coined the term medialization laryngoplasty. While various surgical manipulations fall under the umbrella of medialization laryngoplasty, the two-fold objective of these procedures is 1) to physically approximate the two vocal folds during phonation and 2) to alter the vibratory characteristics of the injured vocal cord. The most common medialization laryngoplasty procedure is the thyroplasty procedure, which is aimed at medializing the membraneous aspects of the vocal fold. The vocal fold is constrained within the laryngeal framework by means of a dense ligamentous attachment, the vocal ligament. In the adult larynx, the laryngeal cartilages have ossified into a rigid framework which is minimally deformable[3]. The rigidity and the distinctive feature of laryngeal cartilage surface inspired us to perform surface based registration in the image guided system. The restoration of voice production and the level of patient satisfaction with a thyroplasty are variable. While revision rates vary, the most recent study published in 2003 indicates an open revision rate of 24%[4]. Optimal voice outcomes are most dependent on the exact placement of the implant relative to the position of the underlying vocal fold and suboptimal voice outcomes and high revision rates reflect the significant challenges inherent in the thyroplasty procedure. The major challenge is determination of the optimal implant configuration and accurate placement of implant during surgery.

Registration in image guided procedures can be classified into three categories based on the fiducial markers: extrinsic invasive, extrinsic noninvasive and intrinsic markers. Extrinsic invasive markers (i.e. stereotactic frames and bone screwed markers) are usually fixed to a patient's large bones. The reported accuracy of extrinsic invasive markers is less than 1mm[5-7]. Extrinsic noninvasive markers are attached to the patient's skin. The accuracy of noninvasive markers is between 2 to 4mm[8]. Commercial systems have been used for a variety of procedures, including neurosurgery, spinal-surgery, and orthopedic surgery. The traditional approach is to affix multi-modal fiducial markers before a CT or MRI is taken of the patient and then register this dataset with the patient using the same markers intraoperatively. Most commercial image-guided systems use optical tracking equipment with extrinsic fiducial markers. Intrinsic markers are anatomical and geometric landmarks, points and surfaces in the preoperative CT/MR dataset and that of the intraoperative medical image[9,10]. In this approach, surfaces, lines, or points in the CT/MR dataset are registered to corresponding features on the patient using intraoperative imaging modalities. The commonly used intraoperative imaging modalities are X-ray, ultrasound, 2D optical images and 3D laser scanning. Ultrasound imaging has been successfully applied to breast, liver and prostate, where the ultrasound energy can easily penetrate the anatomical structure. X-ray imaging has been widely applied for image-guided spinal and orthopedic surgery. The surface based registration method extracts bone surfaces from CT data and contours from X-ray image. The registration is done by minimizing the distance between the projected surface and the contours[11,12]. Pixel similarity based approach directly matches the X-ray image with Digitally Reconstructed Radiography (DRR) image from CT data. The location of CT data relative to the X-ray image is estimated by optimizing the similarity between X-ray and DRR image[13]. Registering the surface from preoperative CT data and point clouds from 3D laser scanning has been applied in image-guided neurosurgery. The accuracy of the Point Based Registration (PBR) and Iterative Closet Point (ICP) methods has been reported to be 1.0-2.9mm and 0.6-2.0mm respectively[10]. 2D optical imaging based registration has been used for registering video image with CT data of human face[14]. The reported accuracy of registration is 1.45-1.59mm. The difficulty in 2D video image based registration is that the set of features must be exposed to the intraoperative imaging (or laser scanning) and matched accurately between the CT/MR dataset and the patient. In the case of laryngoplasty, the bone fixed fiducial markers would make potentially damage to the thin laryngeal cartilage. 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. 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. Surface based registration is a research area in image guided surgery, where the exposed anatomical structure is laser scanned and registered to the preoperative CT or MRI volume. Although, the surface based registration is mainly performed at large anatomical structures like skull, knees and hip joint, it provides a possibility to apply surface based registration method to the image guided laryngoplasty. The distinctive features in larynx and laryngeal cartilage make it possible to perform surface based registration for image guidance.

Structured light based surface reconstruction requires light projection device (LCD projector) and one or more cameras. The structured light based surface reconstruction system can be classified into three categories: time-multiplexing, spatial neighborhood and direct coding. Time-multiplexing is a way to encode the pixel information in the temporal domain. Several binary coded light stripes are projected onto the measuring surface and further decoded to calculate the 3D coordinate using triangulation[15]. One problem with the binary code is that the pixel intensity at stripe boundary is highly sensitive to the noise and the sensor resolution. Inokuchi et al.[16] improved the coding scheme with gray code to make the code word robust to the noise. Caspi et al.[17] reduced the number of images by using colored gray codes. Hattori and Sato[18] refined the original hierarchical stripe based technique by introducing sub-pixel offsets to the final stripe pattern to get better resolution. Recently, Gühring[19] combined the gray code light pattern and line shifting to reconstruct highly accurate 3D surface model. The sub-pixel accuracy peak detection is the key component of accurate 3D reconstruction, which is also used in traditional laser scanning. In spatial neighborhood based coding, the code word of projected light pattern is obtained by considering the neighborhood pixels around it. However, due to the local smoothness constraint, the code word could not be recovered robustly. As a result, the spatial resolution of reconstructed surface is much lower than the time multiplexing. Boyer and Kak[20] projected a color stripe pattern with a unique color intensity value. The advantage of this method is capable of obtaining shape from moving objects. The disadvantage of this approach is the complex decoding algorithms and the uncertainty of the code word. Vuylsteke et al.[21] proposed a De Bruijn sequence based neighborhood coding. The code word is recovered by traversing through the Hamilton circuits of De Bruijn graph. The De Bruijn sequence resolves the ambiguity problem in neighborhood coding. Later, Hall-Holt and Rusinkiewicz[22] described a coding scheme with time-varying stripe patterns. The most common problem in the neighborhood coding is the robust decoding of the code word. Zhang et al.[23] proposed a multi-pass dynamic programming to recover the code word from colored De Bruijn sequence. The major advantage in neighborhood coding is the possibility of reconstructing dynamic scenes. However, the decoding of the code word is a global optimization process and the 3D reconstruction result is less accurate than time multiplexing. Direct coding is to create a pattern in which every pixel can be directly labeled by the image pixel value. Carrihill and Hummel[24] introduced an intensity ratio depth sensor: the ratio between full-lit image and the linearly varying illumination pattern is used to index every pixel along horizontal scanline. Since the ratio based direct coding technique is highly susceptible to sensor noise, Chazan and Kiryati[25] combined this method with hierarchical stripes to reduce noise susceptibility. Miyasaka et al.[26] calculated more accurate intensity ratio depth map by using an LCD projector and a 3CCD camera. Direct coding techniques are useful for achieving large spatial resolution with few images. Spatial neighborhood and direct coding 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. For our experimental framework, since the primary goal is to reconstruct accurate 3D surface for registration, we have combined gray code pattern and multi-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. Besl and McKay introduced ICP (Iterative Closest Point) algorithm to geometrically align two similar geometric models[27]. A new geometric transform matrix is calculated by minimizing the MSE (Mean Square Error) between the closet point pairs. Horn[28] described a closed form solution for the quaternion calculation from matched closet point pairs. Chen and Medioni[29] proposed a similar iterative scheme using a different criterion. Instead of searching for the closest distance between point pairs, they searched for the point to surface distance using surface normal vector. This method is restricted to the 3D models with surface description. Rusinkiewicz and Levoy[30] compared various ICP algorithms based on sub-sampling scheme from the geometric data, closest point searching method, rejection of outliers and error minimization method. To reduce point samples from geometric data, Turk and Levoy[31] uniformly sub-sampled the source data, while Masuda et al. randomly sub-sampled the point data with different sample of points at each iteration step[32]. Rusinkiewicz et al. proposed a point sub-sampling schema that maximizes the normal vector distribution of the selected point samples[30]. To remove the outliers from the matched point sets, Masuda et al reject the point pairs whose distance is larger than certain times of the standard deviation[32]. Turk and Levoy[31] removed the point pairs which contains a point from mesh boundary. KD-tree[27] and approximated KD-tree[33] is used to accelerate the closest point searching process. For the error metric and minimization, most people minimized the sum of squared distance for the matched point pairs. Masuda and Yokoya[32] minimized the least median of square (LMS) distance instead of summed square distance. The LMS estimator minimizes the median of the squared residuals, and it is more robust to the mismatched point pairs. To avoid local minima, Simon et al. started with several perturbations in the initial conditions and select the best result[34]. We used the closed form solution from Horn[28] 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

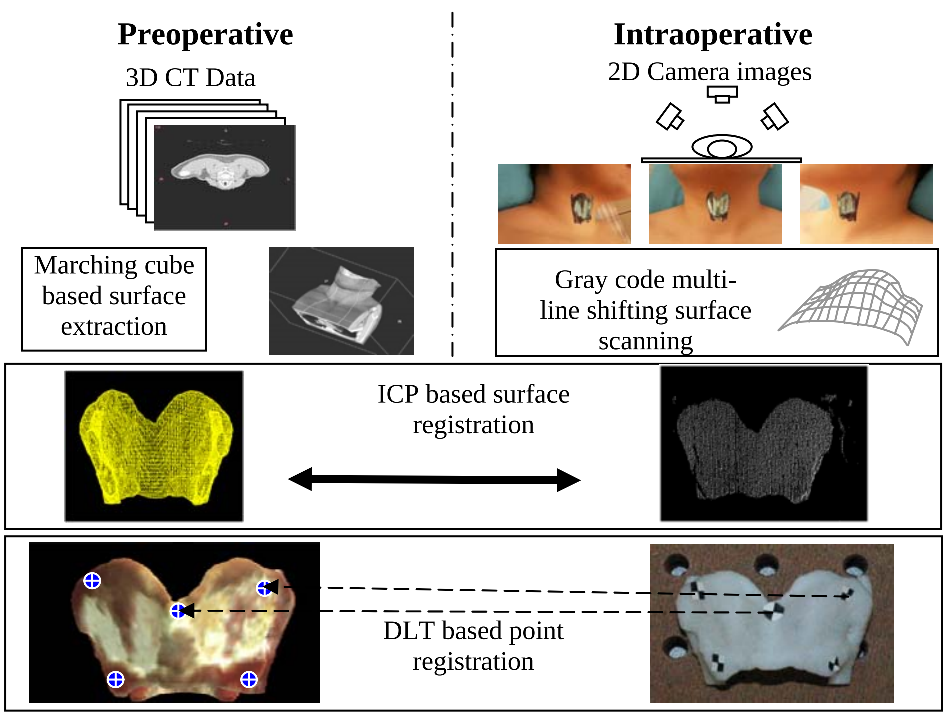

Fig. 1. Work flow of surface registration

The work flow of our surface registration process is shown in figure 1. There are five major steps: 1) In the patient's preoperative CT image, the laryngeal cartilage surface is extracted using marching cube iso-surface extraction algorithm.

2) Stereo camera pairs are calibrated using planar chessboard pattern image taken from different positions and orientations. 3) Intraoperatively, gray code pattern and multi-line shifting based surface reconstruction method is used to reconstruct the exposed laryngeal cartilage surface, which is defined as intraoperative surface. 4) The preoperative surface from step 1 is registered to the intraoperative surface from step 3 by matching the geometric shapes. A shape priori based ICP algorithm is designed to quickly and accurately match the two surfaces. After shape matching, the fiducial marker positions are registered to the corresponding positions in CT data. 5) During the surgery, the fiducial markers are tracked by the real-time stereo vision system and provide the rigid transformation between 3D CT data and patient's space.

3.1 Medialization Laryngoplasty

The medialization laryngoplasty (Fig. 2) is a surgical procedure to restore the voice of patients with vocal fold paresis or paralysis. The etiologies of this condition are typically idiopathic or iatrogenic; examples of the latter include injury to the vagus nerve during thyroidectomy or skull base surgery. The medialization laryngoplasty procedure is the thyroplasty procedure, which is aimed at medializing the membraneous aspects of the vocal fold. In thyroplasty, a surgical opening is created in the lamina of the thyroid cartilage for the purpose of implanting a permanent structural support that pushes the injured vocal fold medially so as to approximate the normal vocal fold of the opposite side during phonation. The medialization laryngoplasty is a collective term for surgical manipulations of the laryngeal framework, including manipulations of the thyroid, cricoid, and/or the arytenoid cartilages designed to restore normal voice in individuals with vocal cord paralysis or paresis. The surgical implant is typically comprised of a rigid material, such as Silastic®, that can be modified in configuration during the surgery. 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 uninjured vocal cord. The anatomy of the thyroid cartilage varies in size and shape among individuals as does the precise position of the vocal folds. The vocal folds are located deep under the thyroid cartilage surface. Therefore, the anatomical landmarks on the external laryngeal cartilage surface are insufficient to accurately locate the underlying vocal folds. In surgical practice, the location is approximated by pre- and intraoperative study of the patient's CT scans, visual inspection of the thyroid cartilage, and the geometry of the vocal folds seen through a trans-nasal fiber optic image.

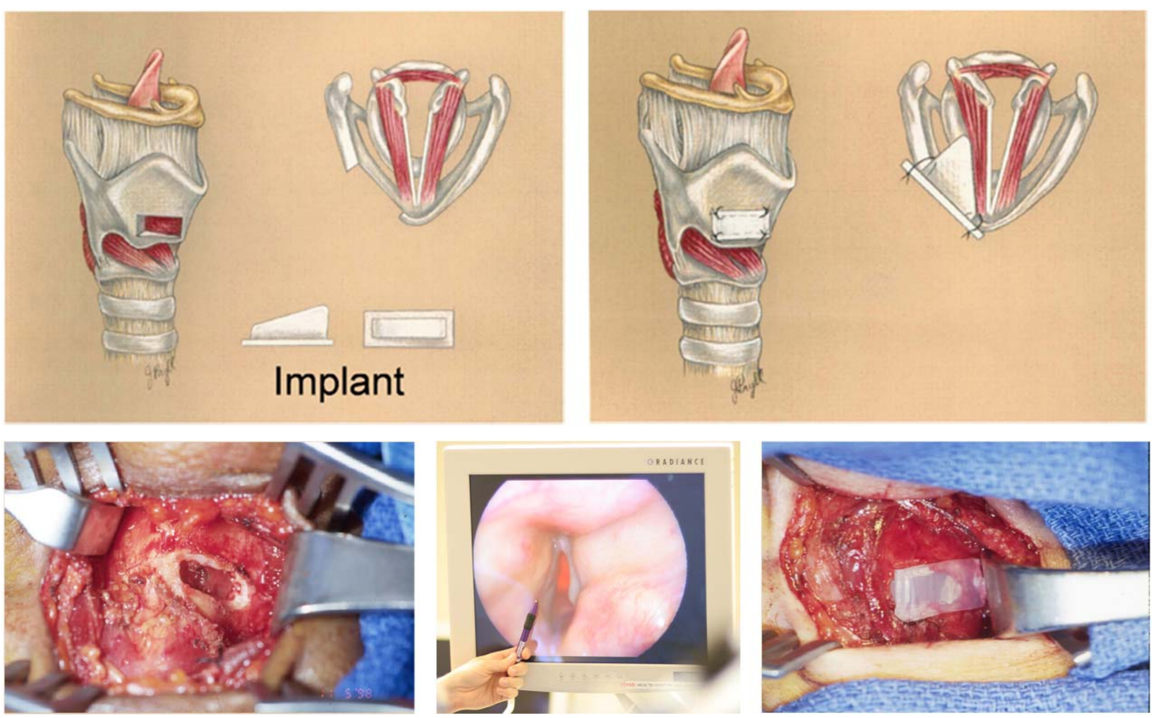

Fig. 2. Medialization laryngoplasty

3.2 Surface Extraction from CT Data and Phantom Model Construction

For the preliminary experiment, the Visible Human CT data set from NIH National Library of Medicine is used for the experiment. The thyroid cartilage surface is extracted using marching cube algorithm[35] with the iso-density value of 1070. he extracted triangular mesh is rendered in wire frame, flat shading and texture mapping (Fig 3. Left). To experiment the surface reconstruction and registration, a phantom model with corresponding CT data set is required. It is difficult to find an actual scale phantom model of laryngeal cartilage. A reverse engineering approach is used to build a phantom model from the Visible Human CT data set. The extracted 3D surface model is converted to a solid CAD model and sent to 3D prototyping device (Stratasys FDM 3000). The prototyper is capable of constructing a 3D phantom model from the CAD input with the accuracy of 0.1mm. (Fig 3. Right).

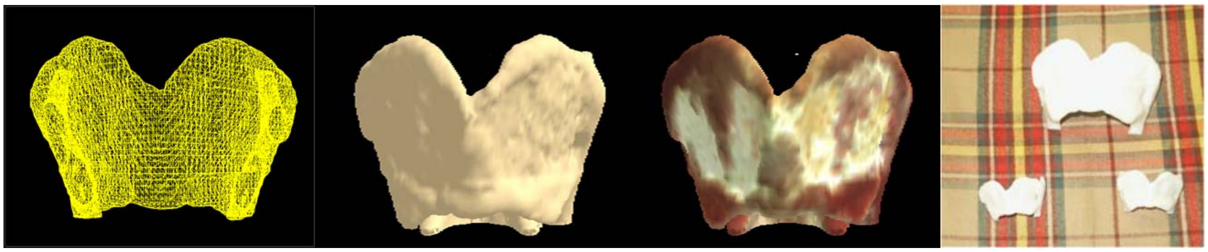

Fig. 3. Left: Iso-surface extraction from the CT data, Right: Phantom model

3.3 Structured Light based Intraoperative 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[36] (Fig. 4).

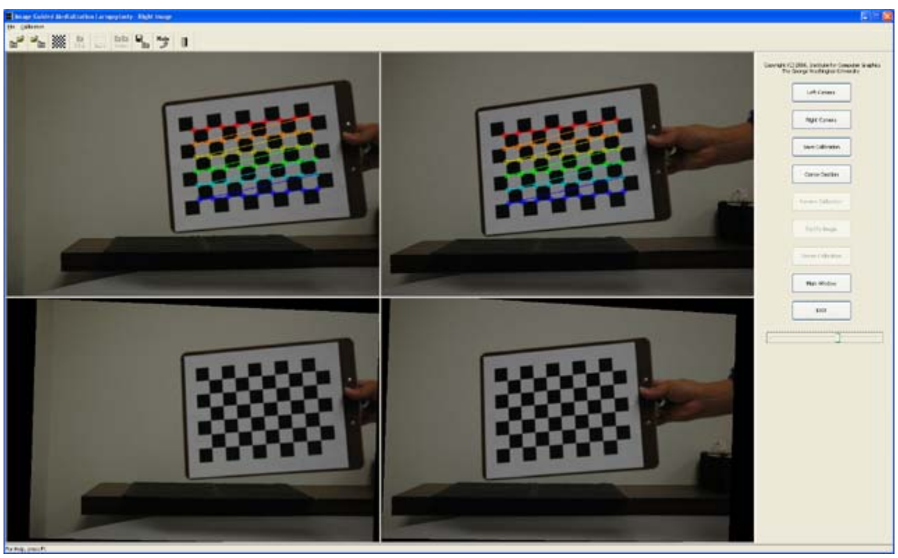

Fig. 4. Camera calibration and rectified images

$$M_1 = \begin{bmatrix} f & 0 & C_{x1} \\ 0 & f & C_y \\ 0 & 0 & 1 \end{bmatrix} \quad M_2 = \begin{bmatrix} f & 0 & C_{x2} \\ 0 & f & C_y \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = M_1^{-1} \begin{bmatrix} x1 \\ y \\ 1 \end{bmatrix} = M_2^{-1} \begin{bmatrix} x2 + Dx \\ y \\ 1 \end{bmatrix} \qquad (1)$$

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. Suppose, *P1*, *P2* is the camera position vector and the *R1*, *R2* is the camera rotation matrix represented by quaternion. In equation 1, the *M1*, *M2* is the pinhole camera projection matrix. If we find the pixel correspondence in left and right images, we can calculate the real 3D position of the pixel in camera coordinate system by solving the linear equations (Equation 1).

From equation 1, 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. We have applied Blais and Rioux sub-pixel peak detection method to reconstruct the surface of thyroid cartilage phantom model. Similar with laser range scanning, we swept the surface of an object with single high intensity line pattern. In order to increase the speed of surface scanning and reduce the number of images used for reconstruction, we have combined gray code pattern and multi-line shifting method. The benefit of this approach is that the number of images required for computation is dramatically reduced, while the accuracy of surface reconstruction is maintained at the same level of single line shifting method. The key challenge in gray code multi-line shifting method is to resolve the ambiguity and mislabeling problem originated from image segmentation, shadows and occlusions. We have tackled these problems with multi-pass dynamic programming to maximize the sequence matching cost. In the first pass, since we knew the ground truth projection pattern, we matched the recovered pattern sequence from both left and right images to the ground truth pattern sequence. This will resolve the mislabeling problem in a single multi-line shifting image. In the second pass, we have sorted all the sub-pixel peak position from multi-line shifting images and applied dynamic programming. The second pass is to maximize the sequence matching cost in both spatial and temporal domain (Fig. 5).

Fig. 5. Gray code and multi-line shifting for surface reconstruction

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.

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 stage, the goal of ICP algorithm is to find the optimal rotation and translation that minimize equation 2. We used unit quaternion *Q(q0,q1,q2q3)* to represent the rotation matrix.

$$E(R,T)=\sum_{i\in M}\sum_{j\in D}\left\|(m_i-(R*d_j+T)\right\| \qquad R=\begin{bmatrix} q_0^2+q_1^2-q_2^2-q_3^2 & 2(q_1q_2-q_0q_3) & 2(q_1q_3+q_0q_2) \\ 2(q_1q_2+q_0q_3) & q_0^2+q_2^2-q_1^2-q_3^2 & 2(q_2q_3-q_0q_1) \\ 2(q_1q_3-q_0q_2) & 2(q_2q_3+q_0q_1) & q_0^2+q_3^2-q_1^2-q_2^2 \end{bmatrix} \qquad (2)$$

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. The new quaternion vector *Q* is the eigenvector of largest positive eigen value of *N*.

$$N=\begin{bmatrix} \mathrm{Sxx+Syy+Szz} & \mathrm{Syz-Szy} & \mathrm{Szx-Sxz} & \mathrm{Sxy-Syx} \\ \mathrm{Syz-Szy} & \mathrm{Sxx-Syy-Szz} & \mathrm{Sxy+Syx} & \mathrm{Szx+Sxz} \\ \mathrm{Szx-Sxz} & \mathrm{Sxy+Syx} & \mathrm{Syy-Sxx-Szz} & \mathrm{Syz+Szy} \\ \mathrm{Sxy-Syx} & \mathrm{Szx+Sxz} & \mathrm{Syz+Szy} & \mathrm{Szz-Syy-Sxx} \end{bmatrix} \quad S_{xx}=\sum_{i\in M}\sum_{j\in D}(m'_{ix}*d'_{jx})\cdot \qquad (3)$$

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 $[\mathrm{X_{M1}},\mathrm{Y_{M1}},\mathrm{Z_{M1}}]$ $[\mathrm{X_{M2}},\mathrm{Y_{M2}},\mathrm{Z_{M2}}]$ is closest matching point pair from CT and structured light based reconstruction, the new translation vector is calculated using equation 4.

$$\begin{bmatrix}Tx & Ty & Tz\end{bmatrix}=[\sum_n(X_{M1}-X_{M2})/n,\sum_n(Y_{M1}-Y_{M2})/n,\sum_n(Z_{M1}-Z_{M2})/n] \qquad (4)$$

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 (Fig. 6). Point to plane distance 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 SVD based closed form solution is used to approximate the plane equation. The plane equation is the vector associated with smallest singular value (Equation 5). Geometric description based on initial shape approximation will provide close initial pose estimation for the ICP method.

$$Dist=\sum_{v\in M}\frac{(ax+by+cz+d)^2}{a^2+b^2+c^2};\ \ \mathrm{SVD:D}=\frac{1}{N}\sum_{v\in M}\begin{bmatrix}x\\y\\z\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}^T \qquad (5)$$

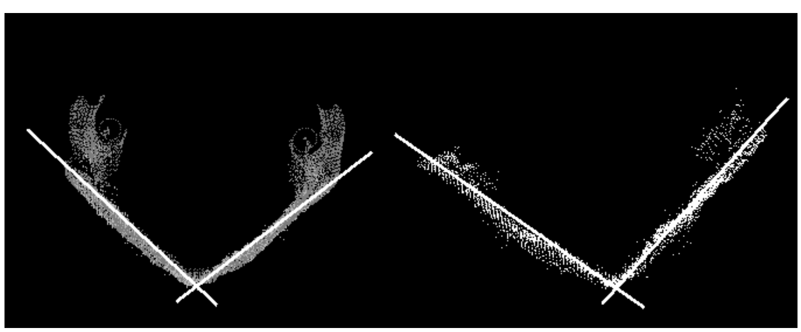

Fig.6. Approximation of the larynx with two crossing planes

4. EXPERIMENT AND RESULT

We used Intel Xeon 3.2GHz Workstation with 4GB memory for our experiment. We have implemented a prototype system with Visual Studio 2005 and MFC. For the camera calibration and image processing, we have utilized the Intel OpenCV library. The prototype system can directly import the DICOM format CT data and extract iso-surface (Fig 7).

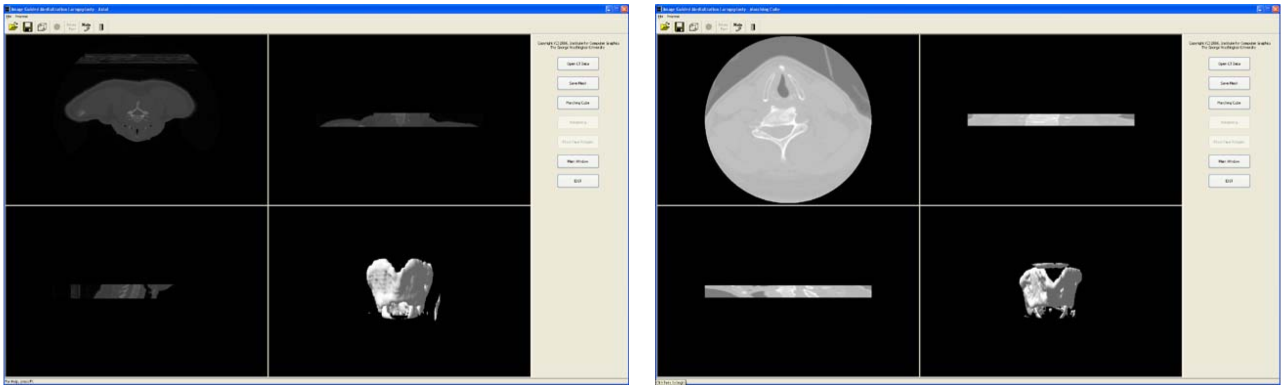

Fig. 7. Iso-surface extraction result

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 8. To mimic the real situation, color dotted model and animal bone are used for the experiment (Fig 9).

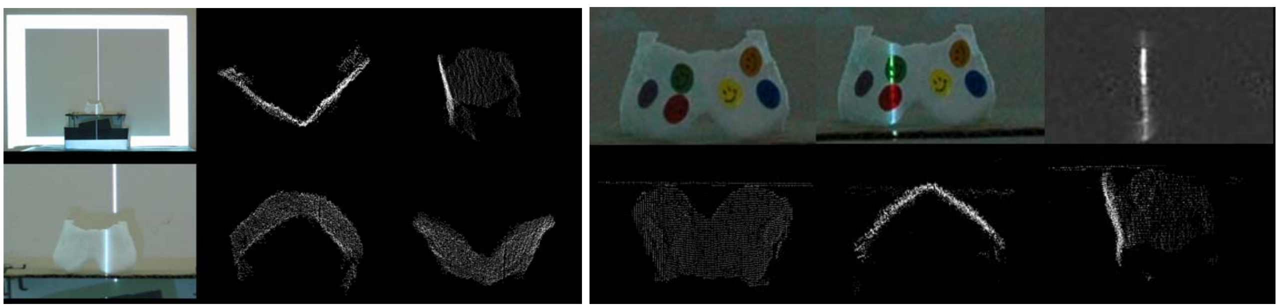

Fig. 8. Surface reconstruction result for phantom model

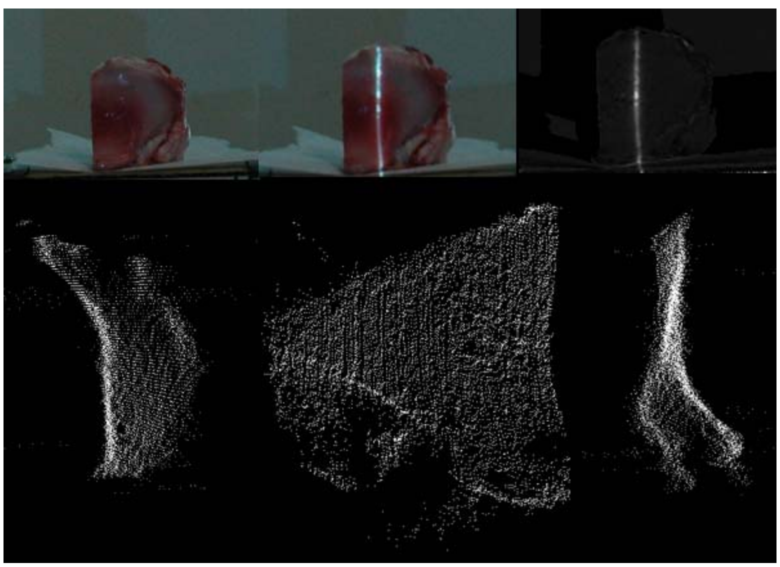

Fig. 9. Animal bone surface reconstruction

To reduce the number of images required for surface reconstruction, we have implemented multi-pass dynamic programming to resolve the ambiguity issue in gray code multi-line shifting method. The surface reconstruction results are shown on figure 10.

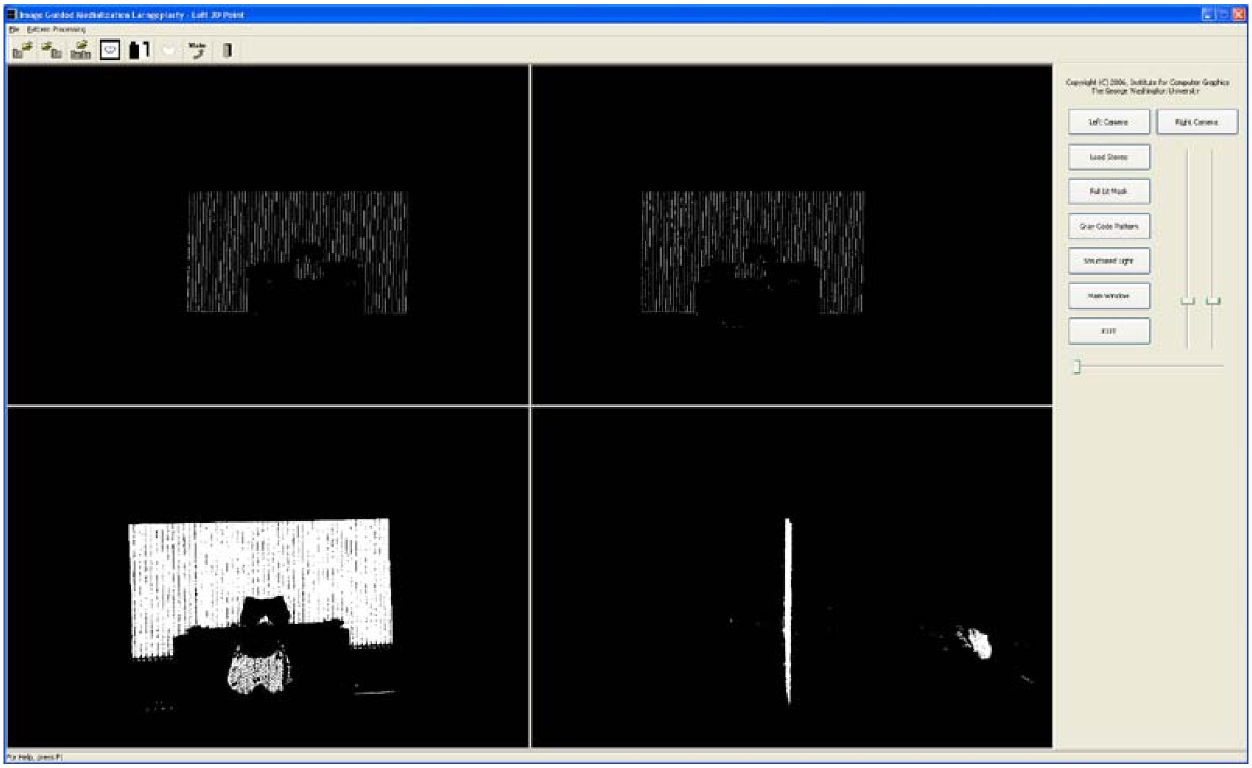

Fig. 10. Multi-pass dynamic programming for surface reconstruction

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 11.

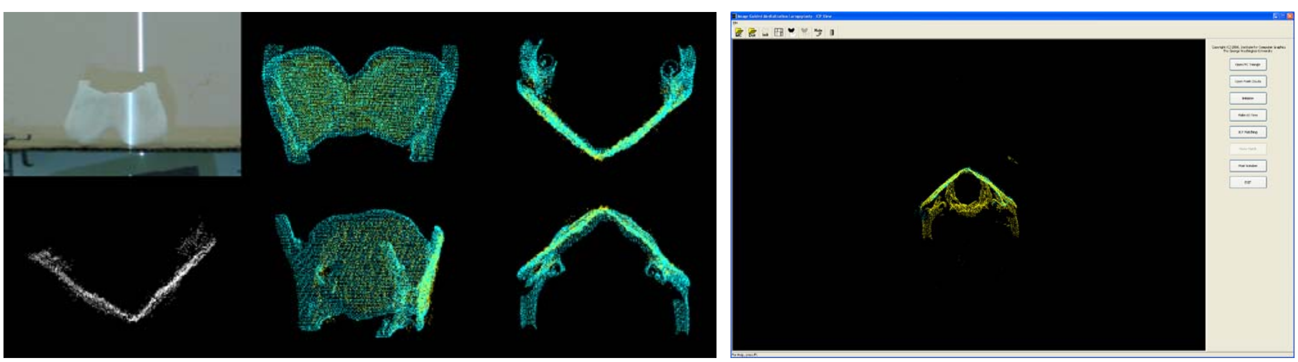

Fig. 11. Shape priori based ICP matching result and interface

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 gray code multi-line shifting for the 3D reconstruction. To resolve the ambiguity in multiple line-shifting, we used the dynamic programming method to maximize the sequence matching cost. To mimic the real situation, color dotted phantom model and animal bone is used for experiment. 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. N. Isshiki, H. Morita, H. Okamura, et al. “Thyroplasty as a new phonosurgical technique”, Acta Otolarynol 78:451-7, 1974.
2. J.A. Koufman. “Laryngoplasty for vocal cord medialization: an alternative to Teflon”, Laryngoscope 96:726-31, 1986.
3. S. Bielamowicz. “Aging voice or vocal fold paresis: what can I do about it?,” Arch Otolaryngology Head Neck Surg. 2004 Sep;130(9):1114-6.
4. T.D Anderson, J.R Spiegel, and R.T Sataloff. “Thyroplasty revisions: frequency and predictive factors.” J Voice;17: 442-448, 2003.
5. P.J.Kelly “Computer-assisted stereotaxis: new approaches for the management of intracranial intra-axial tumors" Neurology.1986;36:535.
6. K. P. Gall and L. J. Verhey “Computer-assisted positioning of radiotherapy patients using implanted radioopaque fiducials”. Medical physics. 1993;20(4):1152:159.
7. 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 1997; 16:447--462.
8. Darabi K, Grunert P, Perneczky A: Accuracy of intraoperative navigation using skin markers, in: CARS 1997. Berlin, Springer-Verlag, 1997, pp 920–924.
9. D. L. G. Hill, D. J. Hawkes, J.E. Crossman, M. J. Gleeson, T.C.S. Cox, E.C. Bracey, A. J. Strong, and P. Graves (1991).” Registration of MR and CT images for skull base surgery using pointlike anatomical features”. British journal of radiology, 64(767), 1030:035.
10. 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, 2003:22(8) 973-985
11. A. Gueziec, P. Kazanzides, B. Williamson, and R. H. Taylor, “Anatomy-based registration of CT-scan and intraoperative X-ray images for guiding a surgical robot,” IEEE Trans. Med. Imag., vol. 17, pp. 715-728, Oct 1998.
12. . H. Livyatan, Z. Yaniv, L. Joskowicz: Gradient-Based 2D/3D Rigid Registration of Fluoroscopic X-ray to CT. IEEE Trans. Med. Imaging 22(11): 1395-1406 (2003).
13. Wells, W., Viola, P., et al.: Multi-Modal Volume Registration by Maximization of Mutual Information. Medical Image Analysis 1 (1996) 35—51
14. M. J. Clarkson, D. Rueckert, D. L. G. Hill, D. J. Hawkes: Using Photo-Consistency to Register 2D Optical Images of the Human Face to a 3D Surface Model. IEEE Trans. Pattern Anal. Mach. Intell. 23(11): 1266-1280 (2001)
15. J. L. Posdamer, M. D. Altschuler, Surface measurement by space-encoded projected beam systems, Computer Graphics and Image Processing 18 (1) (1982) 1–17.
16. S. Inokuchi, K. Sato, F. Matsuda, Range imaging system for 3-D object recognition, in: Proceedings of the International Conference on Pattern Recognition, 1984, pp. 806–808.
17. D. Caspi, N. Kiryati, and J. Shamir. Range imaging with adaptive color structured light. IEEE Trans. on Pattern Analysis and Machine Intelligence, 20(5), 1998.
18. K. Hattori and Y. Sato. Accurate range finder with laser pattern shifting. In Int. Conf. on Pattern Recognition, pages 849–853, 1996.

19. 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
20. K. L. Boyer, A. C. Kak, Color-encoded structured light for rapid active ranging, IEEE Transactions on Pattern Analysis and Machine Intelligence 9 (1) (1987) 14–28.
21. P. Vuylsteke, A. Oosterlinck, Range image acquisition with a single binary-encoded light pattern, IEEE Transactions on Pattern Analysis and Machine Intelligence 12 (2) (1990) 148–163.
22. O. Hall-Holt and S. Rusinkiewicz. Stripe boundary codes for real-time structured-light range scanning of moving objects. In Int. Conf. on Computer Vision, pages 359–366, 2001.
23. Li Zhang, Brian Curless, and Steven M. Seitz. Spacetime Stereo: Shape Recovery for Dynamic Scenes. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Madison, WI, June, 2003, pp. 367-374.
24. B. Carrihill and R. Hummel. Experiments with the intensity ratio depth sensor. Computer Vision, Graphics, and Image Processing, 32:337–358, 1985.
25. G. Chazan and N. Kiryati. Pyramidal intensity ratio depth sensor. Technical Report 121, Center for Communication and Information Technologies, Department of Electrical Engineering, Technion, Haifa, Israel, Oct 1995.
26. T. Miyasaka, K. Kuroda, M. Hirose, K. Araki, High speed 3-D measurement system using incoherent light source for human performance analysis, in: Proceedings of the 19th Congress of The International Society for Photogrammetry and Remote Sensing, The Netherlands, Amsterdam, 2000, pp. 65–69.
27. Besl P. J., Mckay N. D.: A method for registration of 3d shapes. IEEE PAMI. 14, 2 (1992), 239-256.
28. B.K.P. Horn. Closed-form solution of absolute orientation using unit quaternions. Journal of the Optical Society of America, 4(4):629--642, 1987
29. Chen Y., Medioni G.: Object modelling by registration of multiple range images. Image Vision Comput. 10, 3 (1992). 2
30. Szymon Rusinkiewicz and Marc Levoy. Efficient Variants of the ICP Algorithm. Third International Conference on 3D Digital Imaging and Modeling (3DIM). June 2001
31. Turk, G. and Levoy, M. "Zippered Polygon Meshes from Range Images," Proc. SIGGRAPH, 1994.
32. Masuda, T., Sakaue, K., and Yokoya, N. "Registration and Integration of Multiple Range Images for 3-D Model Construction," Proc. CVPR, 1996.
33. Greenspan, M.; Yurick, M.; Approximate k-d tree search for efficient ICP 3-D Digital Imaging and Modeling, 2003. 3DIM 2003. Proceedings. 2003 Page(s):442 – 448
34. D. Simon "Fast and Accurate Shape-Based Registration", Ph. D. Dissertation, Carnegie Mellon University, 1996.
35. W. E. Lorensen and H. E. Cline. Marching cubes: a high resolution 3D surface construction algorithm. In M. C. Stone, editor, Computer Graphics (SIGGRAPH '87 Proceedings), volume 21 (4), pages 163--170, July 1987
36. Zhengyou Zhang. Flexible camera calibration by viewing a plane from unknown orientations. In 7th IEEE International Conference on Computer Vision, pages 666--673, 1999