

Optical imaging for medical diagnosis based on active stereo vision and motion tracking

Tao T. Wu and Jianan Y. Qu*

Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology,
Clear Water Bay, Kowloon, Hong Kong, P. R. China
eequ@ust.hk

Abstract: This study aims to develop a novel imaging technique to improve the accuracy of the colposcopic diagnosis of cervical cancer. An optical imaging system based on active stereo vision is built to measure the 3-D surface topology of cervix and track the motion of patient. The information of motion tracking are used to register the time-sequenced images of cervix recorded over the period of examination. The imaging system is evaluated by tracking the movements of cervix models. The results show that the error of 2-D image registration is 0.8 pixels, equivalent to the motion tracking error of 0.05 mm in the field-of-view. The imaging technique holds the promise to enable quantitative mapping of the acetowhitening kinetics over cervical surface for more accurate diagnosis of cervical cancer.

©2007 Optical Society of America

OCIS codes: (170.3890) Medical optics instrumentation; (170.0110) Imaging systems; (150.6910) Three-dimensional sensing.

References and links

1. B. W. Pogue, H. B. Kaufman, A. Zelenchuk, W. Harper, G. C. Burke, E. E. Burke, and D. M. Harper, "Analysis of acetic acid-induced whitening of high-grade squamous intraepithelial lesions," *J. Biomed. Opt.* **6**, 397-403 (2001).
2. C. Balas, "A novel optical imaging method for the early detection, quantitative grading, and mapping of cancerous and precancerous lesions of cervix," *IEEE Trans. Biomed. Eng.* **48**, 96-104 (2001).
3. T. T. Wu, J. Y. Qu, T. H. Cheung, S. F. Yim, and Y. F. Wong, "Study of dynamic process of acetic acid induced-whitening in epithelial tissues at cellular level," *Opt. Express* **13**, 4963-4973 (2005).
4. Y. F. Wang and J. K. Aggarwal, "An overview of geometric modeling using active sensing," *IEEE Control Syst. Mag.* **8**, 5-13 (1988).
5. J. Batlle, E. Mouaddib, and J. Salvi, "Recent progress in coded structured light as a technique to solve the correspondence problem: a survey," *Pattern Recogn.* **31**, 963-982 (1998).
6. K. Hasegawa, K. Noda, and Y. Sato, "Electronic endoscope system for shape measurement," in *Proceedings of the 16th International Conference on Pattern Recognition* (IEEE Computer Society, 2002), pp. 761-764.
7. T. Wu, J. Qu, T. -H. Cheung, K. Lo, and M. -Y. Yu, "Preliminary study of detecting neoplastic growths in vivo with real time calibrated autofluorescence imaging," *Opt. Express* **11**, 291-298 (2003).
8. R. Y. Tsai, "A versatile camera calibration technique for highaccuracy 3D machine vision metrology using off-the-shelf TV camera and lenses," *IEEE J. Rob. Autom.* **3**, 323-344 (1987).
9. D. Q. Huynh, R. A. Owens, and P. E. Hartmann, "Calibrating a structured light stripe system: a novel approach," *Int. J. Comput. Vis.* **33**, 73-86 (1999).
10. P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Trans. Pattern Anal. Mach. Intell.* **14**, 239-256 (1992).
11. Y. Chen and G. Medioni, "Object modeling by registration of multiple range images," *Image Vis. Comput.* **10**, 145-155 (1992).
12. Z. Zhang, "Iterative point matching for registration of free-form curves and surfaces," *Int. J. Comput. Vis.* **13**, 119-152 (1994).
13. S. Rusinkiewicz and M. Levoy, "Efficient variants of the ICP algorithm," in *Proceedings of 3rd International Conference on 3-D Digital Imaging and Modeling (3DIM'01)* (IEEE Computer Society, 2001), pp. 145-152.
14. J. L. Bentley, "Multidimensional binary search trees used for associative searching," *Commun. ACM* **18**, 509-517 (1975).

1. Introduction

Colposcopic diagnosis of cervical cancer is based on acetowhitening phenomenon, the transient whitening of tissue induced by acetic acid. The temporal kinetics of acetowhitening process produces the diagnostic information to detect and grade the precancerous cervical lesions known as Cervical Intraepithelial Neoplasia (CIN). It has been shown that the quantitative measurement of acetowhitening kinetics at cervix *in vivo* can potentially improve the diagnostic accuracy of colposcopy [1-3]. However, the challenge is that the patient can not remain completely stationary during the colposcopy procedure which normally takes several minutes. Slight motion of the patient will cause the loss of the correspondence between the time-sequenced images of cervix recorded during the examination. Without image registration, the measurement of acetowhitening kinetics using the time-sequenced images will generate false diagnostic information because of patient's motion. Therefore, accurate registration of the time-sequenced images during colposcopy procedure is crucial to accurately measure the kinetics of acetowhitening in the area of interest.

Since the cervix surface is smooth and lacks of obvious features such as edges and lines, conventional methods based on 2-D image processing may not provide reliable information for motion tracking and image registration. The active stereo vision is a kind of mature 3-D optical imaging technology that is particularly effective for the surface reconstruction of the objects with relatively plain spatial characteristics such as cervix [4-6]. A typical active 3-D imaging system projects the structured light on an object to generate feature points. With a simple calibration of imaging and projection channels of the imaging system, 3-D coordinates of the feature points can be measured for reconstructing the surface topology of the imaged object. In this work, we build an imaging system based on the configuration of standard colposcope. A projection channel is added to produce the structured light of grid pattern on the imaged object surface. The imaging channel takes image of the pattern from the object to reconstruct the 3-D surface. The 3-D surfaces measured at different time are used to track the motion of the object. The information of motion tracking are then used for the registration of time-sequenced images recorded right after the application of acetic acid and continuously throughout the acetowhitening process. The ratio of registered image to the image recorded right after the application of acetic acid provides a map of acetic acid induced changes over the imaged cervix surface. Therefore, the imaging method can be used for quantitative imaging of the temporal kinetics of the acetowhitening process.

2. Materials and methods

2.1 Imaging system based on active stereo vision

The schematic of the 3-D imaging system used in this study is shown in Fig. 1(a). The system consists of an imaging and projection channel. The imaging channel is designed with the similar configuration of a standard colposcope, a low power microscope. A three-CCD color camera (Model No. DXC-390P, Sony Co., Japan) and a fast frame grabber (Genesis, Matrox Inc., Canada) are used to capture images with the resolution of 768×576 pixels in the imaging channel. Two polarizers, P1 and P2, with cross-polarization configuration are placed in front of camera and white-light illumination source (Fiber-Lite, MH-100 Metal Halide illuminator, Dolan-Jenner, USA), respectively. This arrangement effectively eliminates the specular reflection from the surface of the examined object [2, 7]. The cross-polarized reflection image of a human cervix model (Model No. PP01915, Nasco, USA) is shown in Fig. 1(b). The full width and height of the image field-of-view is about 44 mm and 33 mm, respectively. The size and shape of the model are similar to the real human cervix. The landmarks (red dots) on the model surface are used to evaluate the performance of the imaging system in 3-D reconstruction of cervix surface, motion tracking and image registration.

In the projection channel a collimated light source from a 100 mw laser diode at wavelength 660 nm (Model Lasiris, StockerYale Inc., Canada) and a holographic grating are used to generate the structured light. The grating (Model SLH-533(0.09)-5, StockerYale Inc., Canada) converts the collimated beam into 33 stripes with 0.09° separation angle between adjacent stripes. A 50/50 beam splitter divides the set of stripes into two sets. The orientation of one set of stripes was rotated by 90° using a pair of mirrors. A 33×33 grid pattern was formed by the combination of two set of stripes that are perpendicular to each other. The angle between the imaging and projection channel is 8° , about the same as that between the imaging channel and white-light illumination of a standard colposcope. The image of the grid pattern projected on a cervix model is shown in Fig. 1(c). The spacing of the lines of the grid projected at the cervix model is about 0.82 mm.

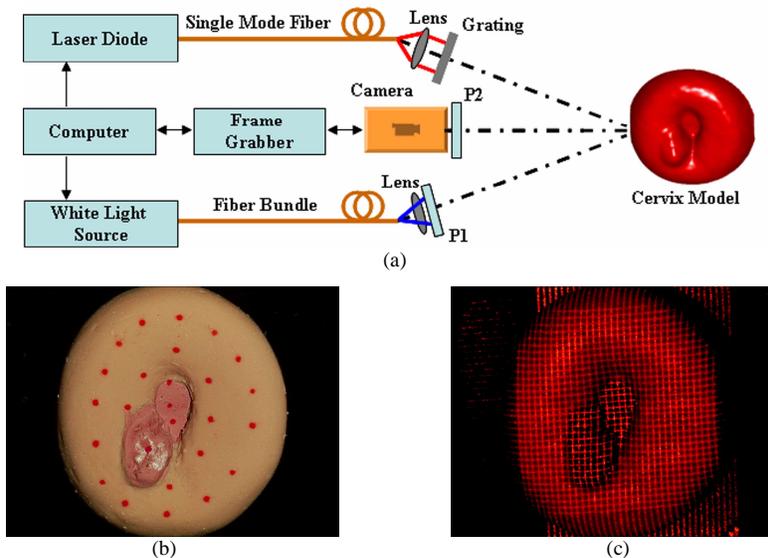


Fig. 1. (a) Schematic of the 3-D imaging system. P1 and P2 are the polarizers with their polarization orientations perpendicular to each other. (b) Image of a cervix model with white-light illumination. (c) Image of the 33×33 grid pattern projected on the model.

2.2 System calibration and 3-D reconstruction

The calibration procedures establish the relationship between the coordinates of the imaging system and a real-world coordinates. First, a well-known Tsai's pinhole camera model was used to calibrate the imaging channel using a non-plane checker board with known 3-D feature points [8]. Then, the calibration parameters were used to calculate the 3-D coordinates of the grid pattern projected on the same checker board. Finally, the 3-D grid will be used for the calibration of the projection channel.

The Tsai's calibration model consists of a rigid body transformation from a 3-D world coordinate system to a 2-D image coordinate system. It follows a projection of pinhole camera geometry from a 3-D real-world coordinate system to an ideal image coordinate in the plane of a CCD camera sensor. The ideal image coordinate is then transformed to an actual image coordinate to characterize the distortion of the imaging system. Finally, the actual image coordinate is related to an image pixel based the knowledge of the CCD camera and the frame grabber. This technique is commonly used in the field of camera calibration, and the details of Tsai's calibration method are presented in [8].

In the calibration of the projection channel the grid pattern is projected on the same checker board used in the calibration of imaging channel. The checker board consists of two orthogonal surfaces with known plane equations. Because each laser stripe can be modeled as a fan plane emitted from the laser center, the grid pattern is considered as the intersections of the fan planes of all laser stripes with the checker board. The 3-D coordinates along the

intersection lines on the checker board can be calculated from the image of grid pattern using the camera calibration model and the plane equations of checker board [9]. Thus, the plane equations of all stripes can be determined using the 3-D coordinates of the intersection lines.

To perform the 3-D surface measurement, the correspondence of feature points on the imaged object must be established between the imaging and the projection channels. The 3-D coordinates of the feature points then can be calculated using a simple triangulation of imaging and projection channels. When the 33×33 grid pattern is projected on the object, the spatial arrangement of the stripes is known. The correspondence of the stripes between the imaging and projection channels can be obtained by simply sorting out the stripes in the image according to their spatial arrangement. For the laser stripes projected on the object surface, the points along each stripe are used as the feature points. After the calculation of the fan plane equation of each stripe and the ray equation of each feature point, the reconstruction problem can be solved based on the triangulation of the ray and plane [9]. The reconstructed grid projected on a cervix model is shown in Fig. 2(a). The corresponding continuous 3-D surface generated by interpolation is shown in Fig. 2(b).

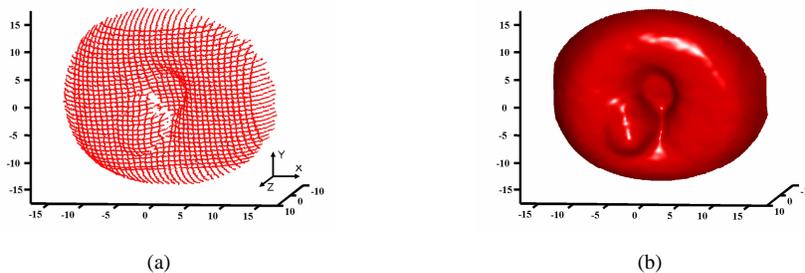


Fig. 2. (a). Reconstructed 3-D grid projected on the surface of a cervix model. (b) Reconstructed continuous surface by using interpolation. The unit of scale is mm.

2.3 Motion tracking and image registration

The iterative closest point (ICP) algorithm is a commonly used method of motion tracking [10-13]. In the motion tracking based on 3-D measurement, the most computationally expensive step in ICP algorithm is to find the paired closest points in a target and reference surface. For accurate closest point calculation, a continuous reference surface of the cervix model at initial position is constructed. The regions in the reference surface not sampled by the grid pattern are filled up with triangle meshes using the interpolation. We adopt the k-d tree method to accelerate the search of the paired closest points [12, 14]. It has been shown that the average complexity of a closest point query can be significantly reduced by the use of the k-d tree method.

Briefly, the k-d tree technique facilitates a fast lookup of the closest vertex of a triangle mesh in the reference surface to a given point, P_i , in the target surface located at a position different from the reference one. The output of the k-d tree algorithm is the vertex, X_i , which is closest to the point, P_i . Given X_i , the closest point Y_i in the reference surface should lie within, or on the border of one of the triangle meshes sharing the vertex X_i . In order to find Y_i , it is necessary to project P_i into the planes defined by these triangle meshes. Y_i is found as the projected point which is closest to P_i . To exclude the outliers due to noise and incomplete sampling, we use a dynamic threshold-based outlier detection [12]. After removing the outliers, we calculate the rigid transformation that minimizes a least-squared distance between the paired closest points. The process is iterated until convergence is reached.

After the completion of 3-D motion tracking, the registration of 2-D images can be achieved by using the 3-D transformation provided by the tracking. Firstly, the coordinates of laser grids in the image of the target surface, $P(X_d, Y_d)$, and in the 3-D real-world coordinates, $P(X_w, Y_w, Z_w)$, are known. We apply the calculated 3-D transformation on $P(X_w, Y_w, Z_w)$ to obtain the 3-D coordinates, $P'(X_w, Y_w, Z_w)$, which is supposed to be the closest corresponding points in the reference surface. Secondly, we project the $P'(X_w, Y_w, Z_w)$ onto the 2-D image

plane to obtain the coordinates, $P'(X_d, Y_d)$, in the image of the reference surface using the calibrated camera model. Thirdly, the corresponding points, $P(X_d, Y_d)$ and $P'(X_d, Y_d)$, can be used to calculate the 2-D transformation and to register target image to reference image [15].

2.4 Performance evaluation

Human cervix model is used to evaluate the accuracy of motion tracking and image registration. The human cervix model is mounted on a multi-dimensional translation and rotation stage that can create precise motion of translation and rotation along the X , Y , and Z axis of the stage, respectively. The Z axis is parallel to the optical axis of the imaging channel and the origin of the stage coordinates is about 280 mm away from the imaging system, typical working distance of a standard colposcope.

As shown in Fig. 1(b), a total of 26 landmarks evenly distributed on the surface of cervix model are used for the performance evaluation. A quaternion method is used to determine the actual motion [10]. It provides a standard for the evaluation of motion tracking results produced by the ICP algorithm. In the quaternion method, the 3-D coordinates of the landmarks on reference and target surfaces are first measured by the 3-D imaging system. The transformation between the target and reference surfaces can then be computed by using the two sets of 3-D coordinates of landmarks based on the quaternion calculation. After applying the transformation to the target surface, the 3-D Euclidean distances and 2-D pixel distances between the corresponding landmarks are calculated to obtain the tracking accuracy of the quaternion method. The accuracy of quaternion method is mainly determined by the measurement error of the 3-D imaging system. The tracking accuracy of ICP algorithm is calculated in the same fashion by using the transformation obtained from 3-D imaging and ICP algorithm. The performance evaluation is based on the comparison in tracking accuracy between ICP and quaternion methods. It should be emphasized that quaternion method relies on the landmarks on object surface and ICP method does not require any landmark.

3. Results and discussions

We evaluated the performance of the 3-D imaging system, motion tracking algorithm and image registration method by creating a series of translations along and rotations around the X - Y - Z -axis of the stage coordinates. The motion ranges of the translation and rotation were ± 10 mm and ± 10 degrees, respectively. These are considered to be large movements of patient during colposcopy procedure because it has been reported that the maximal average patient motion is 4.5 mm [1]. We calculated the 3-D distances in mm and 2-D distances in pixel between the landmarks on the reference surface and the target surface after applying quaternion and ICP transformation, respectively. The performance evaluations were based on the comparison between the mean distances over the motion ranges calculated by using the quaternion and ICP tracking algorithms, respectively. The results are summarized in Table 1.

As can be seen, the mean 3-D distance over the ranges of translational and rotational motions is less than 0.14 mm for quaternion method and less than 0.16 mm for the ICP method, respectively. With the image registration the mean 2-D distance is less than 0.53 pixels for quaternion method and 0.84 pixels for ICP method. The accuracy of ICP algorithm is comparable to the quaternion method of the accuracy only limited by the error of 3-D surface reconstruction.

The primary goal of the imaging technique is to track the motion of patient and produce accurate registration of the time-sequenced images during colposcopy. A movie shown in Fig. 3 demonstrates the accurate 2-D image registration over the motion range based on the 3-D imaging and ICP algorithm. The image on right is the superposition of the reference and target images of the cervix model. The image on left is the superposition of the reference and registered target images. The gravity centers of the landmarks in reference image and target images are labeled by red crosses and green crosses, respectively. The mean distance between the landmarks in the reference and registered images is less than one pixel compared with the distance over a few hundred pixels for maximal motion without using image registration. The results indicate that the imaging system based on active stereo vision and motion tracking is

capable of producing accurate 2-D image registration for large motion of human cervix. The accuracy of image registration is less than 1 pixel, equivalent to 0.06 mm in X-Y plane that is adequate for the measurement of temporal kinetics of acetowhitening process from a CIN lesion normally with the size of a few mm.

Table 1. Comparison of quaternion and ICP tracking results

| Motions in stage coordinates | Mean 3-D distances (mm) | | Mean 2-D distances (pixel) | |
|------------------------------|-------------------------|-------------|----------------------------|-------------|
| | Quaternion | ICP | Quaternion | ICP |
| X – Translation | 0.10 ± 0.08 | 0.13 ± 0.09 | 0.49 ± 0.53 | 0.73 ± 0.57 |
| Y – Translation | 0.14 ± 0.10 | 0.16 ± 0.11 | 0.42 ± 0.52 | 0.84 ± 0.57 |
| Z – Translation | 0.13 ± 0.09 | 0.15 ± 0.10 | 0.35 ± 0.50 | 0.78 ± 0.63 |
| X – Rotation | 0.10 ± 0.07 | 0.11 ± 0.07 | 0.53 ± 0.54 | 0.72 ± 0.58 |
| Y – Rotation | 0.11 ± 0.08 | 0.13 ± 0.09 | 0.42 ± 0.46 | 0.81 ± 0.53 |
| Z – Rotation | 0.11 ± 0.10 | 0.14 ± 0.12 | 0.31 ± 0.49 | 0.79 ± 0.60 |

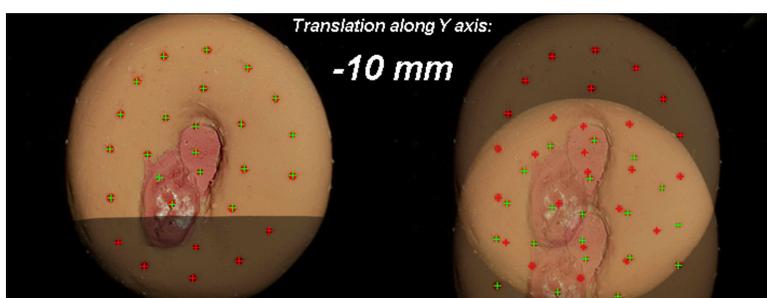


Fig. 3. (3.98 MB) Movie of tracking results (13.71 MB version).

It should be noted that this study is still at preliminary stage. The program takes about 100 seconds to complete the registration of one image. There are a few approaches that can accelerate the process significantly. Firstly, a Kalman filter can be applied to estimate the motion based on a reasonable assumption of slow motion of patient during the examination. This may tremendously reduce the time to search for the paired closest points in ICP algorithm. Secondly, the sampling density of the 33×33 grid may be too high for the human cervix with smooth surface in the region of ectocervix. The sampling density can be reduced to speed up the 3-D reconstruction and ICP algorithms. Finally, a C++ based program will be developed to run on a real-time image processor with multiple CPUs on board, such as GenesisPlus vision processor (Genesis, Matrox Inc., Canada).

4. Conclusion

We developed an active stereo vision system for 3-D imaging of surface topology, motion tracking and image registration. The results show that with the accurate 3-D reconstruction of cervix surface, the ICP algorithm can track the movements of cervix over a large range of motion. The transformation provided by the motion tracking can be used to register the images of cervix recorded at different positions. The accuracy of the image registration is adequate to measure the acetowhitening kinetics of the CIN lesions normally with size of a few mm. In the future work, a colposcope equipped with the 3-D imaging features and image registration algorithm will be built for the study of imaging of the acetowhitening kinetics at clinical level.

Acknowledgments

The authors gratefully acknowledge financial support by the Hong Kong Research Grants Council through the grant HKUST6025/02M and HKUST6408/05M.