Tissue Surface Reconstruction Aided by Local Normal Information Using a Self-calibrated Endoscopic Structured Light System

Jianyu Lin[1,2], Neil T. Clancy[1,3], Danail Stoyanov[4,5], and Daniel S. Elson[1,3]

[1] Hamlyn Centre for Robotic Surgery, Imperial College London, UK
[2] Department of Computing, Imperial College London, UK
[3] Department of Surgery and Cancer, Imperial College London, UK
[4] Centre for Medical Image Computing, University College London, UK
[5] Department of Computer Science, University College London, UK

Abstract. The tissue surface shape provides important information for both tissue pathology detection and augmented reality. Previously a miniaturised structured light (SL) illumination probe (1.9 mm diameter) has been developed to generate sparsely reconstructed tissue surfaces in minimally invasive surgery (MIS). The probe is inserted through the biopsy channel of a standard endoscope and projects a pattern of spots with unique spectra onto the target tissue. The tissue surface can be recovered by light pattern decoding and using parallax. This paper introduces further algorithmic developments and analytical work to allow free-hand manipulation of the SL probe, to improve the light pattern decoding result and to increase the reconstruction accuracy. Firstly the "normalized cut" algorithm was applied to segment the light pattern. Then an iterative procedure was investigated to update both the pattern decoding and the relative position between the camera and the probe simultaneously. Based on planar homography computation, the orientations of local areas where the spots are located in 3D space were estimated. The acquired surface normal information was incorporated with the sparse spot correspondences to constrain the fitting of a thin-plate spline during surface reconstruction. This SL system was tested in phantom, *ex vivo*, and *in vivo* experiments, and the potential of applying this system in surgical environments was demonstrated.

Keywords: Structured light, Endoscopy, Light pattern decoding, Self-calibration, Normal, Reconstruction.

1 Introduction

The measurement of tissue surface shape is important in medical applications since it may indicate tissue pathology and facilitate navigation in MIS. For instance, colonic polyp detection is aided by their morphological appearance [1]. Furthermore, intra-operative information, which can be acquired by optical techniques for surface detection, may be combined with pre-operative information provided by imaging modalities such as MRI and CT, facilitating augmented reality for surgical guidance [2, 3].

N. Navab et al. (Eds.): MICCAI 2015, Part I, LNCS 9349, pp. 405–412, 2015.
DOI: 10.1007/978-3-319-24553-9_50

SL has demonstrated potential in challenging surgical environments due to its non-reliance on salient features [3-5]. In SL a setup is designed to generate a pattern of light and project it through a probe onto the target object. According to the location of the projected light pattern recorded by a single camera, the object surface is estimated. Similar to stereoscopy a geometrical calibration should be applied to establish the spatial relationship between the projector and camera prior to 3D reconstruction.

A unique system has been developed [6] using a 4W supercontinuum laser (420-750 nm) dispersed by an SF-11 prism and focused onto a linear array of fibres. Those fibres, each carrying light of a unique narrowband spectrum, are converted into a random 1.9 mm diameter circular distribution at the distal tip, which is then imaged onto the tissue using a gradient refractive index (GRIN) lens. The system has many advantages, including the use of a narrow diameter flexible probe compatible with instrument ports, unique spectra for each fibre and high pattern brightness due to the use of a coherent laser source that can be focused into the 50 micron core diameters of the fibre array. The system schematic and the projected pattern are shown in Fig. 1(a). Despite its advantages SL systems have limitations, for instance, the calibration needs to be applied every time the flexible probe pose changes. Furthermore, SL systems do not generally provide a reconstruction result denser than the inter-spot spacing.

A system specific software platform has been established previously for surface recovery by conventional calibration and pattern decoding methods [7]. New algorithmic developments and analytical work is proposed to make full use of SL, emphasising two aspects: on-the-fly self-calibration which enables a free-hand manipulation of the probe; incorporation of local normal information in surface reconstruction. Both of these are general techniques that can also be applied to other SL systems. Self-calibration can be used whenever the relative positions of projector and camera are changed, and local normal incorporation can be used whenever light patterns with specific known shapes or sub-patterns are used.

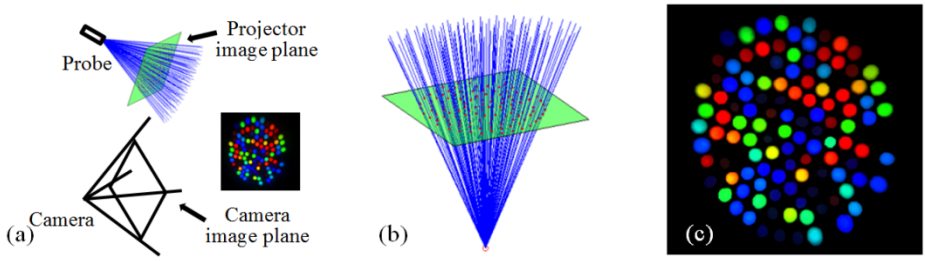

Fig. 1. (a). SL system schematic. (b). The projector model with probe centre (red circle), rays (blue lines), image plane (green), and features (red dots). (c). Virtual projector image plane.

2 Methods

Pre-calibration. Pre-calibration to estimate the intrinsic parameters of the SL system is applied prior to the experiment (Fig. 1 (b)) [7]. However, a virtual projector image plane, which provides a pixel-level description of the ellipse shapes, is also useful as it

allows 3D plane orientation estimation and denser reconstruction. A white planar target is used to generate the projector image plane. The probe projects light onto the target and the SL image is captured by the camera. Induced by a plane at a general position in 3D space, a homography matrix estimated according to the spot centroid correspondences links the projector and camera image planes. The pixels of each spot in the camera image plane are mapped to the projector image plane using this homography and, after interpolation, a virtual projector image plane is generated (Fig. 1 (c)).

Pre-processing. Colour is the most important feature for light pattern decoding using this SL system. A colour vector, defined by the normalised magnitude of each colour component (red, blue and green), is used to represent pixel colour. The colour difference between two pixels can be quantified by the angle between their colour vectors. Due to strong light-tissue interaction pre-processing of SL images is required prior to light pattern decoding. This begins with specular highlight removal and includes high intensity and low saturation area detection, morphological transformation, and interpolation. Colour vectors for all pixels are calculated after Gaussian smoothing.

Light Pattern Decoding. SL images acquired during surgery have a number of challenges. For instance inter-spot distances are small and, due to lateral diffusion at longer wavelengths, may overlap, causing "contamination". In this work we use "normalized cut", which formulates the segmentation problem as spectral clustering [8]. Prior to segmentation, a mask is firstly generated based on regional maxima detection and morphological transformation to cover the spots (foreground). Then spectral clustering is applied to the pixels inside this mask to cluster them into different spots.

Spots on the SL image should be identified to find the correspondences between the projector image plane and camera image plane. In this work, by comparing the projector image with the captured SL image, unique labels are assigned to the spots. Firstly, neighbours for all the spots in both the projector and SL images are detected based on Delaunay triangulation. Then their colour vectors are computed and stacked in a clockwise sequence allowing some unique correspondences to be built. Next, the spot matching procedure is propagated to their neighbourhood from these initial matches until no more correspondences can be found and the projected spots are labelled.

On-the-fly Calibration. Calibration is required whenever the relative position between the probe and the camera changes. We propose on-the-fly self-calibration to estimate the relative positions, enabling free-hand manipulation during surgery.

In this SL system, the image plane of the projector can be regarded as the image plane of the second camera in a normal passive stereo system. The coordinates for all spot centroids X_p on the image plane are calculated in the pre-calibration procedure. On the camera image plane the image coordinates of corresponding spots X_c, after distortion correction and normalisation by the camera matrix, are also available after pre-calibration. Assuming that X_p and X_c correspond to X in 3D space the projective matrices of the projector P_p and the camera P_c, which satisfy $X_p = P_p X$ and $X_c = P_c X$, can be defined as $P_p = [R|t]$, $P_c = [I|0]$. Here R stands for the rotation matrix of the projector, and $t = -Rc$, where c is the position of the probe. The fundamental matrix F indicating the correspondences between these two image planes can be defined as

$$X_p{}^T F X_c = 0. \tag{1}$$

Together with the constraint that $\det F = 0$, F can be derived using the Maximum Likelihood (ML) estimation which minimizes the reprojection error [9]. In practice during the experiment, RANSAC [10] is applied to exclude wrong correspondences (outliers), and only those reliable correspondences are used to estimate F. The essential matrix E, which describes the relative position, is defined as $E = [t]_\times R$. E can be estimated from F based on the SVD and the geometrical interpretation of four possible solutions. In this way, both the probe orientation R and position c can be acquired, and the 3D position of spot centres can be estimated up to scale.

Feature Matching Refinement Using the Epipolar Constraint. Occasionally, due to strong optical scattering or absorption, the spot identification method fails to provide a perfect feature matching result. Therefore, the following iterative procedure is applied to refine both the epipolar geometry and spot identification: In each iteration, the fundamental matrix F is estimated with RANSAC outlier rejection and F optimisation. We then search for more correspondences using F as a constraint. A tolerance is set as the stopping criterion. Then F and correspondences are both updated.

Incorporation of Local Normal Information in 3D Reconstruction. In two view geometry, points on one plane can be related to corresponding points on another plane by a homography uniquely induced by a plane at a general position in 3D space [9]. Given the projective matrices P_p and P_c, and a plane defined by $N^T X = 0$ with $N = (n^T, 1)^T$ in 3D space, a point X on the plane is projected as X_p and X_c on the projector and camera planes, respectively. Then $X_p = HX_c$, where

$$H = R - t * n^T. \tag{2}$$

Normally, H can be estimated according to at least four 2D to 2D point correspondences. If H is known and R and t are calibrated beforehand, the normal of the plane n can be estimated. Based on this, we propose a method to estimate the orientation of the object surface at the reconstructed spot centres. Given that the individual spots are very small in 3D space, we assume that all the pixels inside spot S_i are located on one plane defined by $N_i{}^T X = 0$, with $N_i = (n_i{}^T, 1)^T$, then the image coordinates of those pixels on two image planes can be linked by a homography H_i. However, since pixelwise correspondences cannot be built between the two image planes in our SL system, another method is adopted to estimate H_i. It is noticeable that the shapes of spots on the projector image plane, which represents the intersection between the cone-like rays and the image plane, are all ellipses. It can also be observed that the shapes of most spots on the camera image plane are similar to ellipses, showing the feasibility of the assumption, since a conic can be transformed into another conic by a given homography. An ellipse on a 2D plane can be represented as

$$Ax^2 + 2Bxy + Dy^2 + 2Ex + 2Fy + G = 0. \tag{3}$$

Eq. (3) can be written as $X^T C X = 0$, where $X = [x, y, 1]^T$, and

$$C = \begin{bmatrix} A & B & E \\ B & D & F \\ E & F & G \end{bmatrix}. \tag{4}$$

In our SL system, if a pair of corresponding conics C_p and C_c on two image planes are known, they can be related by a homography H [9]:

$$C_c \sim H^T C_p H. \tag{5}$$

Combining eq. (2) and eq. (5),

$$C_{ci} \sim (R - t * n^T)^T C_{pi} (R - t * n^T), \tag{6}$$

Since eq. (6) defines an overdetermined system, n^T can be estimated through optimisation as long as the conics are non-degenerate. Based on the assumption that each pair of projected spots are located on one single plane in 3D space, corresponding spots in the two image planes of the SL system can be seen as pairs of corresponding conics C_p and C_c. Therefore, we start with ellipse fitting for all the spots to acquire the conic coefficient matrices, using the method proposed by Fitzgibbon et al. [11]. Based on the fitting cost, some spots are excluded since either they are not accurately delineated or they are not located in single planes, hence do not affect the reconstruction. Because all the fitted ellipses are non-degenerate, the local normal n where the spot is located can always be estimated by optimising eq. (6). The optimisation procedure is applied with the constraint that the estimated plane should always go through the 3D location of the spot centre, which is acquired in the previous self-calibration step. The initial guess of a local normal can be estimated from the normal of plane determined by the three surrounding spot locations in 3D space. After optimisation, the estimated normals are filtered using the cost of the optimised objective function. Given the 3D locations of spot centres, the object surface is reconstructed using a thin-plate spline (TPS) [12]. A simple but effective method is used to constrain the surface orientation: for one point with estimated orientation, six new neighbouring points close to it on the estimated plane are added. All the points including both the original and the newly added ones are used to fit the TPS model. In this way, those newly added neighbouring points constrain the local surface direction when TPS fitting takes place.

3 Experiments and Results

In order to evaluate the feasibility of the software platform, experiments on different objects have been carried out. Previous studies with this system have shown that, due to the narrowband spectra of the spots, pattern decoding is robust to changes in background hue and albedo [6]. Additional experiments have been carried out to measure the hue difference for the projected patterns on paper with different colours, finding an RMS difference of approx. 0.01. The accuracy of reconstruction and normal estimation depend on pattern decoding results, which has been demonstrated in previous work on tissue of varying albedos and hues *ex vivo* and *in vivo* [7]. The following

section is divided into two parts: evaluation of self-calibrated reconstruction accuracy through comparison with corresponding manually calibrated results; observation of changes in 3D reconstruction brought by incorporating local normal information.

Reconstruction Results Using Self-calibration. The accuracy of reconstruction with manual calibration has been demonstrated in previous work (0.65 mm error at ~100 mm working distance) [7] and it was used to evaluate the self-calibration result. A silicon heart phantom was imaged as well as *ex vivo* porcine liver, kidney, heart, and *in vivo* porcine large bowel. In total 49 SL images were used and the result was evaluated by registering surfaces from manual and self-calibration. Since the surfaces are reconstructed up to scale, the scaling factor along line-of-sight rays is the only variable used in registration. The average and maximum distances between the two surfaces in percentages are used as indicators of reconstruction error (Fig. 2) (Table 1).

Table 1. Reconstruction errors using the self-calibration technique on benchmark data

Object	Mean distance	Median of mean distance	Max. distance	Median of Max. distance
Heart phantom	0.32% ± 0.13%	0.28%	1.03% ± 0.51%	0.88%
Liver	2.36% ± 5.26%	0.80%	3.84% ± 2.54%	2.68%
Kidney	0.40% ± 0.37%	0.20%	1.76% ± 1.70%	0.88%
Heart	0.81% ± 0.68%	0.44%	2.93% ± 2.39%	1.43%
Large Bowel	3.58% ± 4.38%	2.07%	16.70% ± 23.29%	7.55%

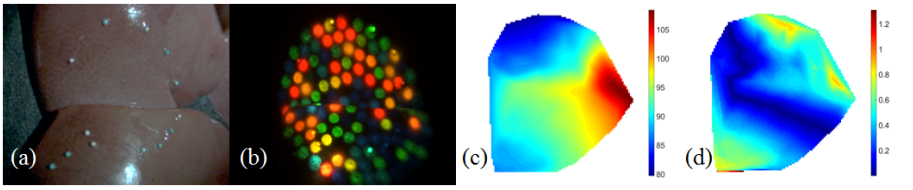

Fig. 2. (a). Liver under white light. (b). Liver under SL. (c). Reconstructed surface using self-calibration where colour indicates the surface depth (in mm). (d). Relative error map on the reconstructed surface using self-calibration where colour indicates the reconstruction error (%).

The mean and median values of both the average and maximum distance show promise for tissue reconstruction using the self-calibration technique. But it is noticeable that according to the maximum distance, despite generally good performance for the case of porcine large bowel, large errors still occur in some areas. This phenomenon was caused by the poor light pattern imaging due to strong scattering of light at long wavelengths and absorption of light at short wavelengths. But the performance does show the potential of the self-calibration technique in surgical environments.

Incorporation of Local Normal Information. The estimated local normals provide additional surface detail. The TPS interpolation is able to provide a reconstruction with high quality for objects with smooth surfaces. However, adding normal information helps to improve the surface detail when the object surface is more complex.

We tested the normal estimation algorithm on a rectangular box, a cylinder, and a liver phantom. Fig. 3 (b), (d), (f) demonstrate the reconstruction of an area around one edge of a rectangular box, a cylinder, and the liver phantom, respectively. The local normals (red arrows) are well estimated using the proposed method.

Since local normals do not benefit the reconstruction much on relatively smooth surfaces, a small object consisting of two steps was used to intuitively show the improvement of surface reconstruction after incorporating normal information. Light patterns were projected onto the object, and Fig. 3 (g) and (h) show the reconstructed surfaces before and after addition of local normal information. The step shape is only accurately defined when incorporating normal information, demonstrating that this method could improve the reconstruction when a sparse light pattern is projected on a non-smooth object surface. Images were acquired using 33 ms exposure times and it took ~50 s to process a single frame using Matlab (Intel i7-3770, 8 GB RAM).

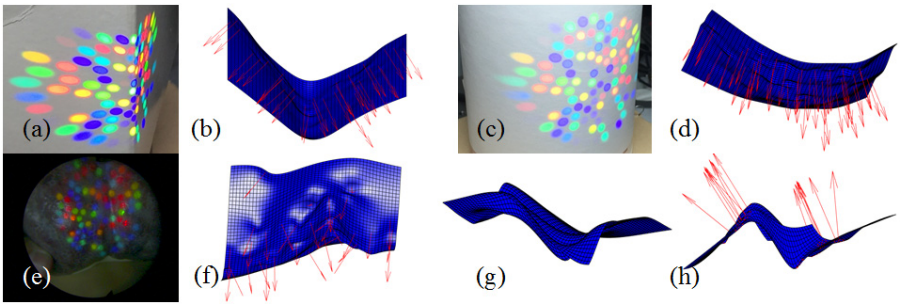

Fig. 3. (a). Rectangular box edge under white light and SL. (b). Reconstructed rectangular box edge (blue) aided by normals (red). (c). cylinder under white light and SL. (d). Reconstructed cylinder (blue) aided by normals (red). (e). Liver phantom under white light and SL. (f). Reconstructed liver phantom (blue) aided by normals (red). (g). Reconstructed step surface (blue) without normal information. (h). Reconstructed step surface (blue) aided by normals (red).

4 Discussion and Conclusion

In this paper we have developed an SL system that is capable of estimating 3D depth as well as inferring information about the surface normal at each SL point. We propose an on-the-fly self-calibration technique along with an iterative algorithm to update both the epipolar geometry and feature matching to enhance SL inference of 3D structural information. Comparison between the resulting reconstruction and that from manual calibration exposes the feasibility and robustness of this technique in object surface reconstruction. But further studies are needed to localise depth in metric space. This method is appropriate for fibre-based SL systems that rely on point projection. Based on the assumption that some projected spots are located on one plane in 3D space, local normal information is estimated according to the homography between two corresponding ellipses. Our experiments show that the improvement of incorporating normal information is notable and that this technique can also be adapted in other SL systems using light patterns with specific known shapes. Future

work will focus on further hardware improvements to reduce ambiguities in pattern decoding and development of implementations where real-time robust and dense tissue surface reconstruction is possible through algorithm design and parallelisation architectures. This would allow dynamic tissue imaging on a per-frame basis.

Acknowledgements. This work is funded by ERC 242991 and an Imperial College Confidence in Concept award. NC is supported by an Imperial College Junior Research Fellowship. The authors thank Northwick Park Institute for Medical Research for surgical arrangements.

References

1. Schwartz, J.J., Lichtenstein, G.R.: Magnification endoscopy, chromoendoscopy and other novel techniques in evaluation of patients with IBD. Techniques in Gastrointestinal Endoscopy 6, 182–188 (2004)
2. King, A.P., Edwards, P.J., Maurer, C.R., Cunha, D.A.D., Gaston, R.P., Clarkson, M., Hill, D.L.G., Hawkes, D.J., Fenlon, M.R., Strong, A.J., Cox, T.C.S., Gleeson, M.J.: Stereo Augmented Reality in the Surgical Microscope. Presence: Teleoper. Virtual Environ. 9, 360–368 (2000)
3. Maier-Hein, L., Mountney, P., Bartoli, A., Elhawary, H., Elson, D., Groch, A., Kolb, A., Rodrigues, M., Sorger, J., Speidel, S., Stoyanov, D.: Optical techniques for 3D surface reconstruction in computer-assisted laparoscopic surgery. Medical Image Analysis 17, 974–996 (2013)
4. Maier-Hein, L., Groch, A., Bartoli, A., Bodenstedt, S., Boissonnat, G., Chang, P.L., Clancy, N.T., Elson, D.S., Haase, S., Heim, E., Hornegger, J., Jannin, P., Kenngott, H., Kilgus, T., Muller-Stich, B., Oladokun, D., Rohl, S., dos Santos, T.R., Schlemmer, H.P., Seitel, A., Speidel, S., Wagner, M., Stoyanov, D.: Comparative Validation of Single-Shot Optical Techniques for Laparoscopic 3-D Surface Reconstruction. IEEE Transactions on Medical Imaging 33, 1913–1930 (2014)
5. Schmalz, C., Forster, F., Schick, A., Angelopoulou, E.: An endoscopic 3D scanner based on structured light. Medical Image Analysis 16, 1063–1072 (2012)
6. Clancy, N.T., Stoyanov, D., Maier-Hein, L., Groch, A., Yang, G.-Z., Elson, D.S.: Spectrally encoded fiber-based structured lighting probe for intraoperative 3D imaging. Biomedical Optics Express 2, 3119–3128 (2011)
7. Lin, J., Clancy, N.T., Elson, D.S.: An endoscopic structured light system using multispectral detection. In: International Journal of Computer Assisted Radiology and Surgery (2015) ISSN:1861-6410
8. Shi, J., Malik, J.: Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 22, 888–905 (2000)
9. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press (2004). ISBN: 0521540518
10. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM 24, 381–395 (1981)
11. Fitzgibbon, A., Pilu, M., Fisher, R.B.: Direct least square fitting of ellipses. IEEE Transactions on Pattern Analysis and Machine Intelligence 21, 476–480 (1999)
12. Franke, R.: Smooth interpolation of scattered data by local thin plate splines. Computers & Mathematics with Applications 8, 273–281 (1982)