

Dynamic View Expansion for Enhanced Navigation in Natural Orifice Transluminal Endoscopic Surgery

Mirna Lerotic, Adrian J. Chung, James Clark, Salman Valibeik,
and Guang-Zhong Yang

Institute of Biomedical Engineering, Imperial College, London SW7 2AZ, UK
{mirna.lerotic, a.chung, j.clark, salman.valibeik,
g.z.yang}@imperial.ac.uk

Abstract. Natural Orifice Transluminal Endoscopic Surgery (NOTES) is an emerging surgical technique with increasing global interest. It has recently transcended the boundaries of clinical experiments towards initial clinical evaluation. Although profound benefits to the patient have been demonstrated, NOTES requires highly skilled endoscopists for it to be performed safely and successfully. This predominantly reflects the skill required to navigate a flexible endoscope through a spatially complex environment. This paper presents a method to extend the visual field of the surgeon without compromising on the safety of the patient. The proposed dynamic view expansion uses a novel parallax correction scheme to provide enhanced visual cues that aid the navigation and orientation during NOTES surgery in periphery, while leaving the focal view undisturbed. The method was validated using a natural orifice simulated surgical environment and demonstrated on *in vivo* porcine data.

1 Introduction

Since the advent of Minimally Invasive Surgery (MIS) [1], surgery has strived towards developing ever less invasive procedures without compromising on treatment outcomes. The recent development of Natural Orifice Transluminal Endoscopic Surgery (NOTES) had taken a giant step towards this goal [2]. In NOTES, the endoscope is guided within the lumen to a point which is closest to the target operative site. The luminal wall is then breached and the endoscope is inserted into either the peritoneal or thoracic cavity depending on the type of procedure performed. The endoscope is navigated to the operative site within the cavity from where the operation can be carried out typically with tools passed through biopsy channels within the endoscope. Once the operation is completed, the endoscope and specimen are removed and the luminal breach closed behind.

Despite significant clinical interests in NOTES, its adoption in surgery relies heavily on the instrument design and improvements in the ergonomic and navigation factors. The problems with the technique revolve predominantly around the equipment, which often must be adapted towards, rather than specifically designed for NOTES [3]. However, despite the future improvement in the functionality of these instruments, the technique itself has a number of fundamental problems which need to be

refined. Irrespective of instrument design, the procedure will still rely upon the navigation of a flexible endoscope through a spatial environment which lacks the familiar luminal constraints for guidance.

In NOTES, since the instrument is no longer constrained in a single lumen, navigational cues are essential in orientating the instrument in relation to the overall patient anatomy. Loss of awareness of the endoscope's rotation and position highlights significant safety concerns as vision can easily be impaired (see Fig. 1) and random movements undertaken to reinstate visual cues can cause inadvertent damage to adjacent structures [4]. In current NOTES practice, the endoscope is continually required to be retroflexed to approximate and adjust the position and forward direction of the instrument, ensuring not only that the endoscopist is navigating the space within the remit of the operation, but also that the instrument is not in the vicinity of any structure which has the potential of being damaged during its transposition. Correcting the visual orientation can be achieved by twisting the external portion of the endoscope and although small corrections in this way are of little concern, the endoscopist must be aware that rotation is also occurring in the unseen trailing portion of the instrument. If a loop has developed in this portion of the endoscope, as is frequently seen in intraluminal studies and is more likely to occur when no constraints are in place on the instrument, it can potentially turn in a wide arc and cause serious damage to structures in its path without the endoscopist even being aware of the potential danger [5].

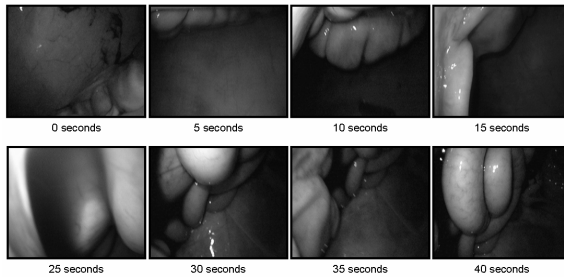


Fig. 1. An image sequence extracted from a NOTES procedure on a porcine model, demonstrating the disorientation that can occur during exploration of the peritoneal cavity using a flexible endoscope

Methods to aid the positioning, as well as the visio-spatial orientation will provide a major step forward in the wider acceptance of the NOTES technique. Visual navigation and orientation cues designed into the NOTES equipment will reduce procedure time, improve safety and shorten the technique's learning curve. Previous research into NOTES navigation involved a combination of EM tracking and endoscopic ultrasound to provide augmented reality sufficient for navigation purposes [6]. Real-time endoscopic mosaicing for improved visualization during MIS for locally planar (retinal) and cylindrical (endometrial) imaging has also been described [7]. The purpose of this paper is to present a new method for dynamic view expansion during NOTES navigation. It is expected that by further expanding the field-of-view of the current optics, an improved awareness of the endoscope's trajectory and orientation can be

achieved, thus aiding the confidence of the operator and avoiding unnecessary or potentially dangerous maneuvers. The method is validated and demonstrated with a simulated NOTES environment of the transluminal cavity and *in vivo* porcine studies.

2 Method

The proposed concept of dynamic view expansion is closely related to existing research on image mosaicing by which an image with an extended field-of-view (FOV) is reconstructed from a set of narrow FOV images. The major difference is that for NOTES we can no longer consider the projected surface to be planar. For accurate image mosaicing, it is necessary to know the relative poses of the camera for each image frame. Estimation of the camera motion can be performed from the video sequences in case of a static scene. If there are deforming regions in the scene, which is common in NOTES procedures, this estimation will be highly inaccurate, and subsequent mosaicing is likely to fail.

2.1 View Expansion Using Optical Flow

Instead of camera motion, the dynamic view expansion is conducted using optical flow extracted from the video sequence. To this end, the method by Black *et al.* [8] was used due to its robustness to varying environment factors. The optical flow beyond the image boundaries was estimated by fitting a homography to the known optical flow data. In essence, homography gives point-to-point mapping between two overlapping images taken from different camera positions [9]. A 2D homography \mathbf{H} was used to fit to optical flow data for each frame. The estimated homography \mathbf{H} was then used to generate estimated optical flow beyond the image boundaries.

It should be noted however, that a homography can only transform images under planar perspective mapping, and therefore it will not accurately map non-planar objects, nor collections of out-of-plane features. Such geometry exhibits parallax under translational motion of the camera. If the objects were mapped and mosaiced using just the homography, the resulting image will appear “flattened” since everything is treated as planar. However, due to parallax, the change of object position due to camera movement will depend on object depth in the scene. To cater for this effect, a new correction factor was introduced. The parallax correction factor $s(x, y)$ is given by

$$s(x, y) = 1 + M \cdot N \cdot \frac{\left| \sqrt{u(x, y)^2 + v(x, y)^2} - \sqrt{u'(x, y)^2 + v'(x, y)^2} \right|}{\sum_{x, y=1..M, N} \sqrt{u'(x, y)^2 + v'(x, y)^2}} \quad (1)$$

where $u(x, y)$ and $v(x, y)$ denote the extracted optical flow in the x and y directions respectively, and $u'(x, y)$ and $v'(x, y)$ represent the estimated optical flow for view expansion by applying \mathbf{H} (see Fig. 2). As the actual motion of local image features tended to deviate from the fitted homography by varying amounts, the parallax correction $s(x, y)$ was used to scale the motion of local features according to their individual depths, thus avoiding the *flattening* effect. This map of correction factors

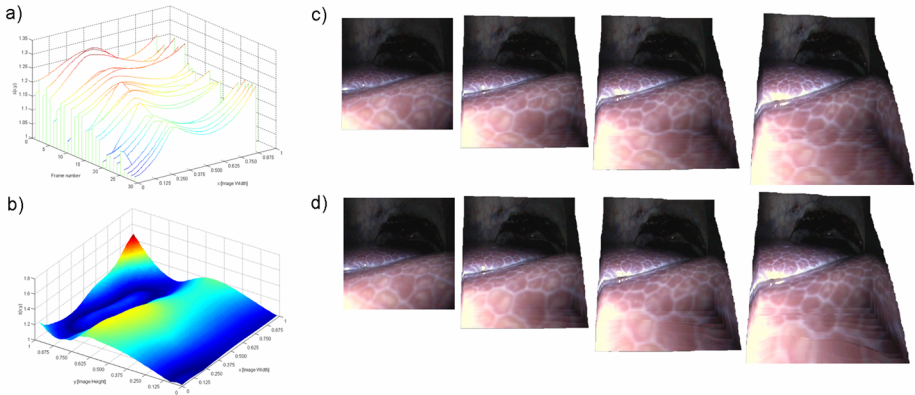


Fig. 2. Image mosaicing created with and without the parallax correction. The change in correction factor is illustrated in (a) and an example of the parallax scale for one frame is shown in (b). Without parallax correction, the reconstructed scene appears flattened (c). After correction, a correct perspective view is recreated as shown in (d). Seamless stitching was not used for the images in (c-d) to accentuate the parallax effect.

s was also dynamically expanded using the same image warp that was applied to the mosaic so that the parallax effect can be reproduced in the expanded view. If depth map information is available, it is possible to use image-based modeling to cater for parallax effects [10].

2.2 Seamless Image Blending

Due to the unique optical configuration of the endoscope, the light source moves in synchrony with the camera. This causes surface brightness to change as the endoscope moves, thus creating visible seams in the composite view. To reduce this artifact, an image composition method was adopted in which the brightness of the mosaic was adjusted locally to match the perimeter of the current video frame using a Poisson image editing operation [11]. In this method, an arbitrarily shaped source image can be mapped onto a typically larger destination image by reformulating it as a partial differential equation, the solution of which is the final mosaic. The seam between the source and destination image is hidden by enforcing Dirichlet boundary conditions [11] on the perimeter of the source image which is constrained to equal the destination image intensities. The second order differential equation is discretised down to pixel resolution thus producing a large sparse positive definite system of linear equations which are solved iteratively for each of the color channels independently.

Prior to merging, the mosaic pixel values were adjusted locally to match the current video frame whose pixel values were left unchanged. The larger mosaic is being composited onto the smaller video frame, hence a hole was created in the middle of the mosaic 2 pixels smaller than the video frame, thus leaving a 1 pixel wide overlap at the perimeter $\delta\Omega$. As a result, the center of the new mosaic was overwritten with the current video frame. The mosaic is updated by solving the Poisson equation:

$$|N_i|f'_i - \sum_{j \in N_i} f'_j = \sum_{j \in \delta\Omega \cap N_i} g_j + |N_i|f - \sum_{j \in N_i - \delta\Omega} f_j \quad (2)$$

where N_i is the set of pixels adjacent to pixel i , f_i are the pixel values of the mosaic prior to updating, f'_i are the unknown pixel values of the updated mosaic that are to be determined, and g_i are the pixel values of the current video frame. Dirichlet boundary conditions were thus defined on the shared perimeter loop $\delta\Omega$. The set of equations formed a sparse linear system which was solved iteratively using a conjugate gradient method with successive over-relaxation as the pre-conditioner.

Once completed, the extended view adds peripheral vision to the camera, which improves visual orientation and navigation. While enabling extension of the FOV, it should be noted that the expanded view may contain image artifacts, therefore they should only be used to aid navigation as they do not constitute full “true” representation of the scene.

3 Results

In order to demonstrate the navigational challenges encountered in NOTES, a Natural Orifice Simulated Surgical Environment (NOsSE) as shown in Fig. 3(e) was constructed. This NOsSE phantom incorporates a box MIS trainer with recreated anatomical structures. Using a standard flexible endoscope, a clinician carried out repeated simulated tasks on the model that represented NOTES procedures. The result of applying the above view expansion scheme for the NOsSE phantom for a short navigation path is shown in Fig. 3. As the endoscope moved through the peritoneal cavity, as shown in Fig. 3(a-d), a homography was fitted to the extracted optical flow

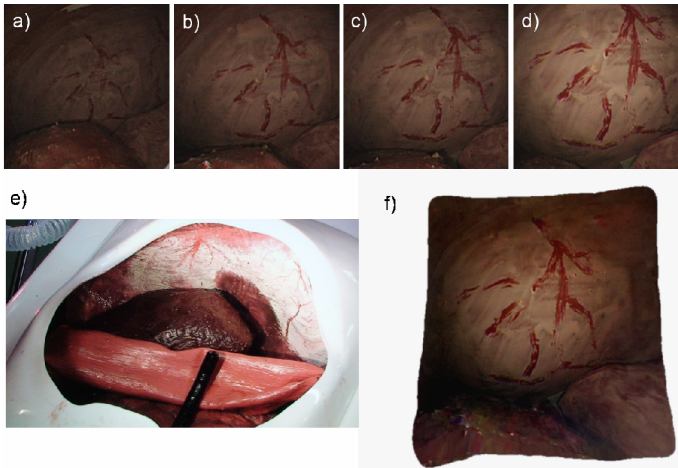


Fig. 3. Natural Orifice Simulated Surgical Environment (NOsSE) model (e) used in this paper is deformable, CT compatible, representative of gas inflated peritoneal cavity. Video sequence from NOsSE phantom (a-d), was used to create a mosaic of the scene (f). Pixels in the expanded periphery are of varying age, some having been mapped through several image warps.

for each video frame. It was then used to modify the preceding frames so that they could be blended into a mosaic that conforms to the new camera view. The current video frame is then merged with the mosaic using Poisson image editing, resulting in a seamless blend, as shown in Fig. 3(f). The current frame was left unaltered in order to preserve integrity of the displayed information.

In order to quantitatively assess the visual fidelity of the proposed dynamic view expansion scheme, the method was validated by testing the expanded view conformance with epipolar geometrical constraints. The fundamental matrix \mathbf{F} that relates point correspondences between two of the video frames was evaluated using manually selected landmarks. The geometric error E , defined as the mean geometric distance between points in the first image that were used to estimate \mathbf{F} and their corresponding epipolar lines in the second image, was within 2.3% of the image size. The expanded view was validated using validation points that moved out of the visible region in the original video frame, but remained visible in the dynamically expanded view. Geometric errors measured for each of the validation points, presented in Table 1, are within the accuracy of the evaluated fundamental matrix, and confirm that the mosaic conforms to epipolar geometric constraint.

Table 1. Geometric error was used to test the conformation to epipolar geometry constraints for the NOsSE phantom (the values are normalised to image width and height, respectively)

	1	2	3	4	5
x / Image Width	0.693	0.533	0.523	0.550	0.775
y / Image Height	0.213	0.227	0.202	0.213	0.712
Geometric error	0.0057	0.0118	0.0047	0.0084	0.0083

To demonstrate the value of the proposed technique for *in vivo* procedures, a NOTES experiment was performed on a porcine model. The peritoneal cavity was accessed trans-vaginally. The endoscope footage was post-processed using the dynamic view expansion, with the parallax correction factor and the results are shown in Fig. 4. Three video sequences were selected from the procedure, Fig. 4(a-d), and used to generate corresponding expanded views, Fig. 4(e). Using the described technique for parallax correction (see Fig. 2) the expanded views are presented as viewed from the perspective of the camera position of the last frame. The last, real-time frame remains unchanged and is merged with the extended view using seamless stitching.

Table 2. Geometric error is used to verify that the motion of feature points in the expanded view conform to epipolar geometry constraints for the *in vivo* porcine data (the values are normalised to image width and height, respectively)

	1	2	3	4	5
x / Image Width	0.507	0.371	0.783	0.716	0.580
y / Image Height	0.830	0.778	0.782	0.848	0.797
Geometric error	0.0050	0.0088	0.0052	0.0019	0.0021

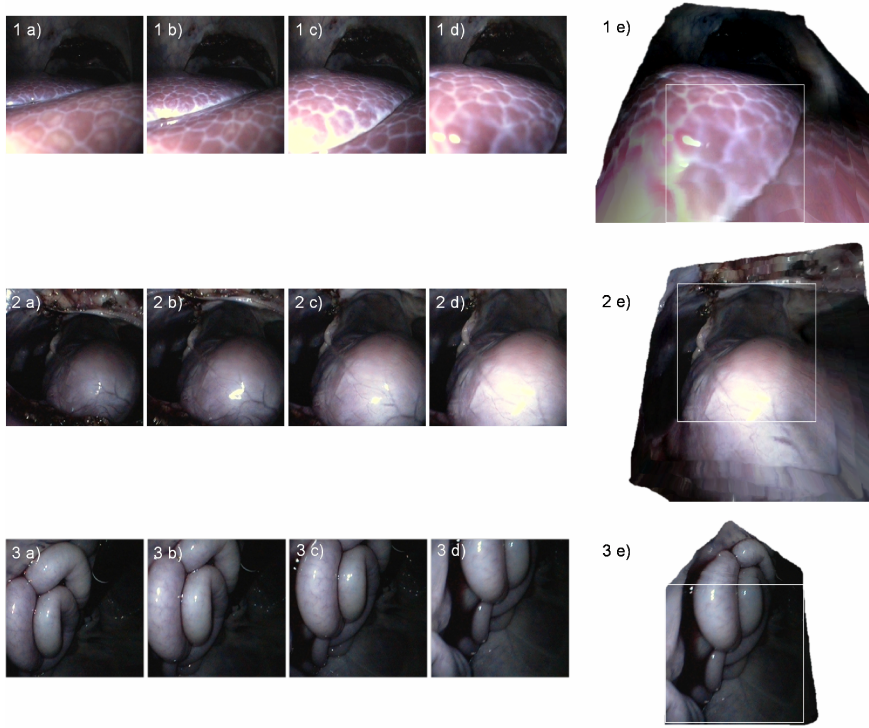


Fig. 4. Selected video sequences from a porcine procedure for evaluating the proposed dynamic view expansion scheme

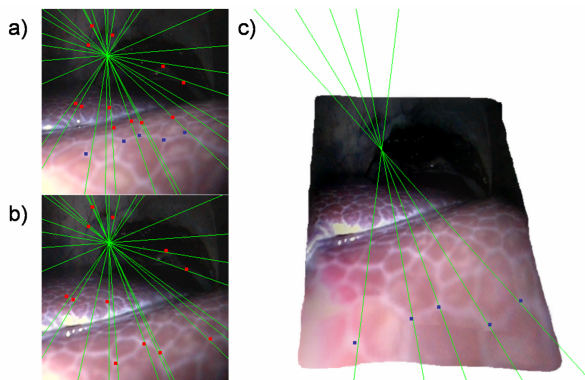


Fig. 5. Error analysis for the proposed view expansion scheme. The fundamental matrix was evaluated from 15 markers manually selected between two images in (a-b). The markers visible in the first frame, shown as blue dots in (a), which had migrated to the mosaiced region (c) in the second frame, were chosen for validation. The distance of the validation markers to the corresponding epipolar lines (shown in green) was measured.

Validation results are given in Table 2 and further illustrated in Fig. 5 for visual assessment. As before, the fundamental matrix was evaluated using selected points (as illustrated in Fig. 5(a-b)), with the geometric error falling within 0.8% of the image size. The geometric errors, given in Table 2, are illustrated in Fig. 5(c) and confirm that the validation points lie within the accuracy of the epipolar lines.

4 Conclusions

In this paper, we have proposed the dynamic view expansion scheme for extending the effective visual field in NOTES to encompass the past positioning of the endoscope and thus facilitate both navigation and spatial awareness. The technique uses a novel parallax correction scheme to seamlessly blend the expanded view with real-time video images. With this technique, it is expected that fewer gross maneuvers will be necessary to correct for both direction and orientation of the instrument, thus leading to improved safety and accuracy. The technique provides visual cues that aid the navigation and orientation during NOTES in periphery, leaving the focal view undisturbed. Initial validation of the method on a NOsSE phantom and an *in vivo* porcine experiment demonstrates the potential practical value of the technique.

Future work includes implementation on a GPU to facilitate real time performance. Additionally, combination of the dynamic view expansion technique methods with methods that track long term camera moments, such as the simultaneous localization and mapping or structure from motion, could improve the method over longer sequences.

References

1. Muhe, E.: Long-term follow-up after laparoscopic cholecystectomy. *Endoscopy* 24(9), 754–758 (1992)
2. Marescaux, J., Dallemagne, B., Perretta, S., Wattiez, A., Mutter, D., Coumaros, D.: Surgery without scars: report of transluminal cholecystectomy in a human being. *Arch. Surg.* 142(9), 823–827 (2007)
3. Bardaro, S.J., Swanstrom, L.: Development of advanced endoscopes for Natural Orifice Transluminal Endoscopic Surgery (NOTES). *Minimally Invasive Therapy and Allied Technologies* 15(6), 378–383 (2006)
4. Holden, J.G., Flach, J.M., Donchin, Y.: Perceptual-motor coordination in an endoscopic surgery simulation *Surgical Endoscopy* 13(2), 127–132 (2004)
5. Cao, C.G.L., Milgram, P.: Disorientation in Minimal Access Surgery: A Case Study. In: *Proceedings of the IEA/HFES 2000 Congress, San Diego*, pp. 169–172 (2000)
6. Estepar, R.S.J., Stylopoulos, N., Ellis, R., Samset, E., Westin, C.-F., Thompson, C., et al.: Towards scarless surgery: An endoscopic ultrasound navigation system for transgastric access procedures. *Computer Aided Surgery* 12(6), 311–324 (2007)
7. Seshamani, S., Lau, W., Hager, G.: Real-Time Endoscopic Mosaicking. In: Larsen, R., Nielsen, M., Sporning, J. (eds.) *MICCAI 2006. LNCS, vol. 4190*, pp. 355–363. Springer, Heidelberg (2006)

8. Black, M.J., Anandan, P.: A framework for the robust estimation of optical flow. In: Proceedings Fourth International Conference on Computer Vision, pp. 231–236 (1993)
9. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision, 2nd edn. Cambridge University Press, Cambridge (2004)
10. ElHelw, M.A., Lo, B.P.L., Chung, A.J., Darzi, A., Yang, G.-Z.: Photorealistic Rendering of Large Tissue Deformation for Surgical Simulation. In: Barillot, C., Haynor, D.R., Hellier, P. (eds.) MICCAI 2004. LNCS, vol. 3216, pp. 355–362. Springer, Heidelberg (2004)
11. Perez, P., Gangnet, M., Blake, A.: Poisson image editing. *ACM Trans. Graph* 22(3), 313–318 (2003)