pq-space Based Non-Photorealistic Rendering for Augmented Reality

Mirna Lerotic, Adrian J. Chung, George Mylonas, and Guang-Zhong Yang

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

Abstract. The increasing use of robotic assisted minimally invasive surgery (MIS) provides an ideal environment for using Augmented Reality (AR) for performing image guided surgery. Seamless synthesis of AR depends on a number of factors relating to the way in which virtual objects appear and visually interact with a real environment. Traditional overlaid AR approaches generally suffer from a loss of depth perception. This paper presents a new AR method for robotic assisted MIS, which uses a novel *pq*-space based non-photorealistic rendering technique for providing see-through vision of the embedded virtual object whilst maintaining salient anatomical details of the exposed anatomical surface. Experimental results with both phantom and *in vivo* lung lobectomy data demonstrate the visual realism achieved for the proposed method and its accuracy in providing high fidelity AR depth perception.

1 Introduction

Augmented reality (AR) is becoming a valuable tool in surgical procedures [1-3]. Providing real-time registered preoperative data during a surgical task removes the need to refer to off-line images and aids the registration of these to the real tissue. The visualization of the objects of interest becomes accessible through the see-through vision that AR provides. In recent years, medical robots are increasingly being used in Minimally Invasive Surgery (MIS). With robotic assisted MIS, dexterity is enhanced by microprocessor controlled mechanical wrists, allowing motion scaling for reducing gross hand movements and the performance of micro-scale tasks that are otherwise not possible. The unique operational setting of the surgical robot provides an ideal platform for enhancing the visual field with pre-operative/intra-operative images or computer generated graphics. The effectiveness and clinical benefit of AR has been well recognized in neuro and orthopedic surgeries. Its application to cardiothoracic or gastrointestinal surgery, however, remains limited as the complexity of tissue deformation imposes significant challenges to the AR display.

Seamless synthesis of AR depends on a number of factors relating to the way in which virtual objects appear and visually interact with a real scene. One of the major problems in AR is the correct handling of occlusion. Although the handling of partial occlusion of the virtual and real environment can be achieved by accurate 3D reconstruction of the surgical scene, particularly with the advent of recent techniques for real-time 3D tissue deformation recovery [4], most surgical AR applications involve

N. Ayache, S. Ourselin, A. Maeder (Eds.): MICCAI 2007, Part II, LNCS 4792, pp. 102–109, 2007.

the superimposition of anatomical structures behind the exposed tissue surface. This, for example, is important for coronary bypass for which improved anatomical and functional visualization permits more accurate intra-operative navigation and vessel excision. In prostatectomy, 3D visualization of the surrounding anatomy can result in improved neurovascular bundle preservation and enhanced continence and potency rates. Whilst providing a good in plane reference in stereo vision environments, traditionally overlaid AR suffers from inaccurate depth perception. Even if the object is rendered at the correct depth, the brain perceives the object as floating above the surface [5, 6]. For objects to be perceived as embedded in the tissue, our brains expect some degree of occlusion. To address the problem of depth perception in AR, a number of rendering techniques and display strategies have developed to allow for accurate perception of 3D depth of the virtual structures with respect to the exposed tissue surface. A recent study has also compared the perceptual fidelity of different visualization techniques designed to overcome misleading depth perception cues [7].

The purpose of this paper is to investigate a new AR method for robotic assisted MIS. The method is based on *pq*-space based Non-Photorealistic Rendering (NPR) for providing a see through vision of the embedded virtual object whilst maintaining salient anatomical details of the exposed anatomical surface. Detailed user experiments demonstrate the effectiveness of the technique in making object appear embedded in the tissue with accurate depth perception due to accentuated ridges that "occlude" the object. This inversion of realism due to NPR has shown to correctly bring the perceptual focus to the object whilst maintaining the relative depth to the real environment.

2 Method

The key task of the proposed technique is to render the exposed anatomical surface as a translucent layer while keeping sufficient details to aid navigation and depth cueing. To this end, surface geometry based on *pq*-space representation is first derived, where *p* and *q* represent the slope of the surface along the *x*, *y* axes, respectively. They can be solved with photometric stereo by introducing multiple lighting conditions. For deforming tissue, however, the problem is ill posed and the introduction of multiple light sources *in vivo* is not feasible. Nevertheless, the problem can be simplified for cases where both camera and a light source are near to the surface [8, 9], such as bronchoscopes and endoscopes. In such cases, the value of image intensity at coordinates x, y for a near point light source is given by

$$E(x,y) = \frac{s_0 \rho(x,y) \cos\theta}{r^2} \tag{1}$$

where s_0 is the light source intensity constant, $\rho(x,y)$ is the albedo or reflection coefficient, r is the distance between the light source and the surface point (x,y,z), and θ is the angle between the incident light ray and the normal to the surface $\hat{n}$. In gradient space, the normal vector to the surface is equal to

$$\hat{n} = \frac{\left(p,\, q,\, -1\right)}{\sqrt{1 + p^2 + q^2}} \tag{2}$$

where p and q represent surface slopes in directions x and y respectively.

For a smooth Lambertian surface in the scene, image irradiance given by equation Eq. 1 can be reduced to

$$E(x,y) = s_0 \rho_{average} \frac{\left(1 - p_0 x - q_0 y\right)^3}{Z_0^2 \left(1 - p_0 x_0 - q_0 y_0\right)^2 \left(1 + p_0^2 + q_0^2\right)^{1/2} \left(1 + x^2 + y^2\right)^{3/2}} \tag{3}$$

and it defines the relationship between the image irradiance $E(x,y)$ at the point (x,y) and scene radiance at the corresponding surface point $(x_0 Z_0, y_0 Z_0, Z_0)$ with surface normal $(p_0, q_0, -1)$, where $\rho_{average}$ denotes the average albedo in a small neighborhood of the surface. By utilizing partial derivatives of the irradiance in equation Eq. 3, the following set of linear equations can be derived [9]

$$\begin{aligned} \frac{1}{E}\frac{\partial E}{\partial x} &= -3\left(\frac{p_0}{\left(1 - p_0 x - q_0 y\right)} + \frac{x}{\left(1 + x^2 + y^2\right)}\right) \\ \frac{1}{E}\frac{\partial E}{\partial y} &= -3\left(\frac{q_0}{\left(1 - p_0 x - q_0 y\right)} + \frac{y}{\left(1 + x^2 + y^2\right)}\right) \end{aligned} \tag{4}$$

They can be solved for surface gradients (p_0, q_0) at each point (x,y) of the image. This technique generally provides normal estimation errors smaller than 0.1 radians [10].

In this study, *pq* values of the surface provide 3D details of the exposed anatomical structure and are used to accentuate salient features while making smoothly varying background semi-transparent. To create the desired visual clues, surfaces of the scene that are parallel to the viewing plane are rendered as transparent whilst the sloped structures are accentuated. A measure of the surface slope is generated from the *pq*-values for each image point (x,y) by

$$S(x,y) = \log(abs(p_0) + abs(q_0)), \tag{5}$$

where high values of $S(x,y)$ correspond to large gradients. The smooth background image B is created by applying a wide 6×6 Gaussian filter ($\sigma = 7$) on S. The two textures are then combined by using a mask $mask(x,y)$ defined by a Catmull-Rom spline.

3 Experimental Setup and Results

In order to assess the accuracy of the proposed method in providing AR stereo-depth perception, a user experiment was designed to compare the traditional AR overlay with the new NPR AR rendering. To provide 3D stereo vision, a da Vinci robotic

surgical console (Intuitive Surgical, Inc., Mountain View, USA) was used. Ten subjects with experiences in surgical AR were recruited for the study. After a short training period, the subjects were asked to locate prescribed virtual targets with an Omni™ Phantom® device (The SensAble Technologies, Woburn, MA, USA). After positioning the tool on the target, subjects recorded the position of the Omni™ Phantom® device by pressing a button.

In this study, two experiments were conducted. The first experiment required the users to locate eight small virtual spheres positioned at different depths of a real thoracic model viewed from the stereo laparoscopic camera of the da Vinci system. Both the standard AR overlay and the proposed NPR AR rendering were provided. In the second experiment, the subjects were asked to define the tumor margin with the Omni™ Phantom® device by using a recorded da Vinci lung lobectomy procedure. Lung lobectomy is the most common surgery performed to treat lung cancer, and Video-Assisted Thoracoscopic Lobectomy (VATS) has significant advantages over conventional surgery due to reduced blood loss and shorter recovery time [11]. In all experiments, the subjects had to rely on their perceived depth of the virtual object in relation to the real anatomical structure to navigate the instrument. No tactile feedback was given to the subjects about the relative position the instrument tip to the target.

Fig. 1 illustrates three examples of the rendering results by using the proposed *pq*-space based NPR scheme for the thoracic model and robotic assisted lung lobectomy procedure used for the two user experiments. It is evident that salient surface details are preserved whilst the overall surface now appear as semi-transparent. Under stereo vision, the overall 3D structure of the exposed anatomical site appeared extremely realistic.

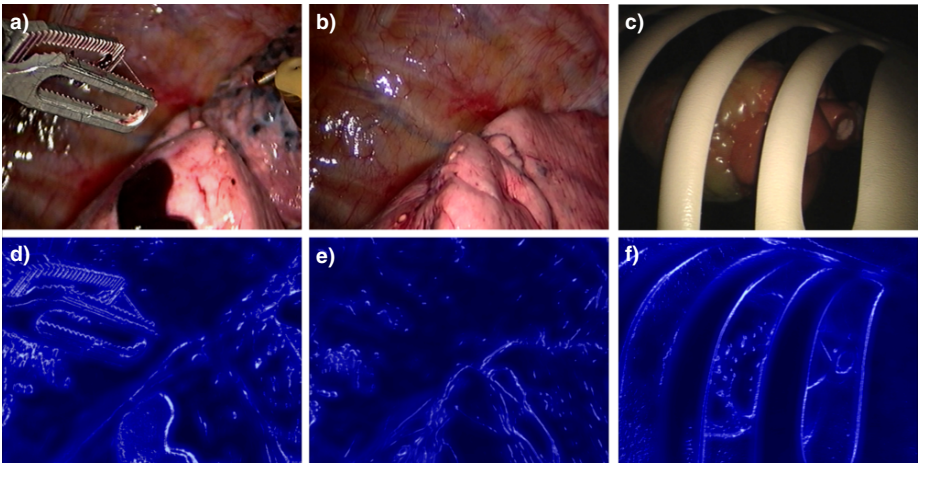

Fig. 1. Example AR rendering results rendered using the proposed *pq*-space based NPR method for robotic assisted lung lobectomy (a-b) and the thoracic phantom model used for the user experiments (c) and (f) are shown on (d-e) respectively

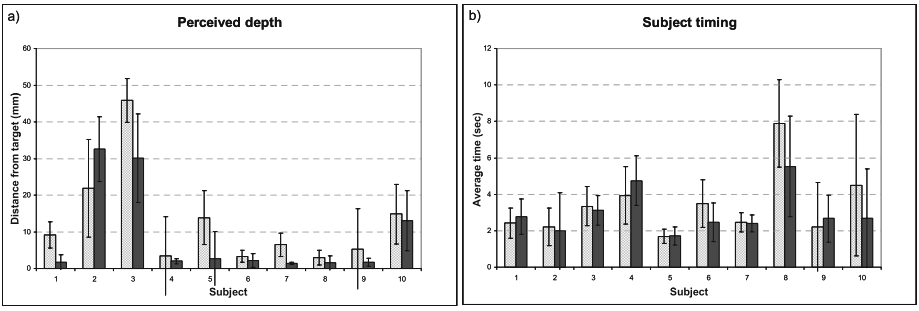

Fig. 2. Error analysis of the perceived depth for the ten subjects studied by using the traditional AR (gray bars) and new NPR AR (black bars) methods, as well as the corresponding time required to locate the targets for the thoracic phantom model. In both graphs, median values and 95% confidence intervals (CI) are given.

Fig. 2 summarizes the results of the first experiment for the ten subjects studied. In Fig. 2(a), the perceived versus the actual depths of the object for the standard and the proposed NPR AR environments are provided. It is evident that there is a marked improvement in the accuracy of the depth perception when NPR AR is used. Non-parametric Wilcoxon signed rank test has given a p-value <0.001 (the data was not normally distributed as determined by Kolmogorov-Smirnov test), confirming that the error in perceived depth is significantly smaller in NPR AR as compared to that of traditional AR. The overall time taken by the subjects in reaching to the virtual spheres for the first experiment is shown in Fig. 2(b), which shows a marginal improvement in performance speed. The Wilcoxon signed-rank test gives a p-value of 0.16, which highlights the fact that the significant impact of the proposed NPR AR scheme is in depth perception accuracy rather than the speed of instrument navigation.

In Fig. 3, the left and right stereo views of the lung lobectomy procedure are provided, where the original da Vinci images are shown in Figs. 3(a-b), the traditional AR overlay in Figs. 3(c-d), pq-space NPR AR rendering in Figs. 3(e-f), and finally *pq*-space NPR AR fused with the original laparoscopic view in Figs. 3(g-h). It is apparent that in the traditional AR view, the object appears as if floating above the surface due to the effect of mixed occlusion. The *pq*-space based NPR AR rendering provided an effective "see-through" vision of the exposed anatomical surface and under stereo viewing, the virtual object appeared correctly embedded into the lung parenchyma. The fused views in Figs. 3(g-h) provided a strikingly realistic viewing experience with which the achieved visual realism and depth correspondence are of high quality.

The errors in perceived depths of this experiment for the ten subjects studied are summarized in Table 1. As in the first experiment, the subjects were able to judge the depth more accurately by using the proposed NPR AR rendering when compared to the traditional AR overlay. Similar results were observed for the time it took for each subject to reach the target. Tool trajectories of a representative subject are shown in Fig. 4 to demonstrate the improved depth perception and instrument targeting by using the proposed NPR AR rendering technique.

After the experiment, the users were asked to provide a subjective measure of the quality of the two AR environments based on a Likert scale (1-poor, 5-excellent).

They were also asked to score the overall visual impression and depth perception of the object. The mean visual score for the traditional AR was 2.5, as compared to 4.4 for the proposed *pq*-space based NPR AR method. The mean depth perception score was 2.6 for the traditional AR, as compared to 4.7 for the proposed technique.

Another AR technique based on a "virtual window" [7] uses occlusion as a depth clue to achieve improved depth perception.

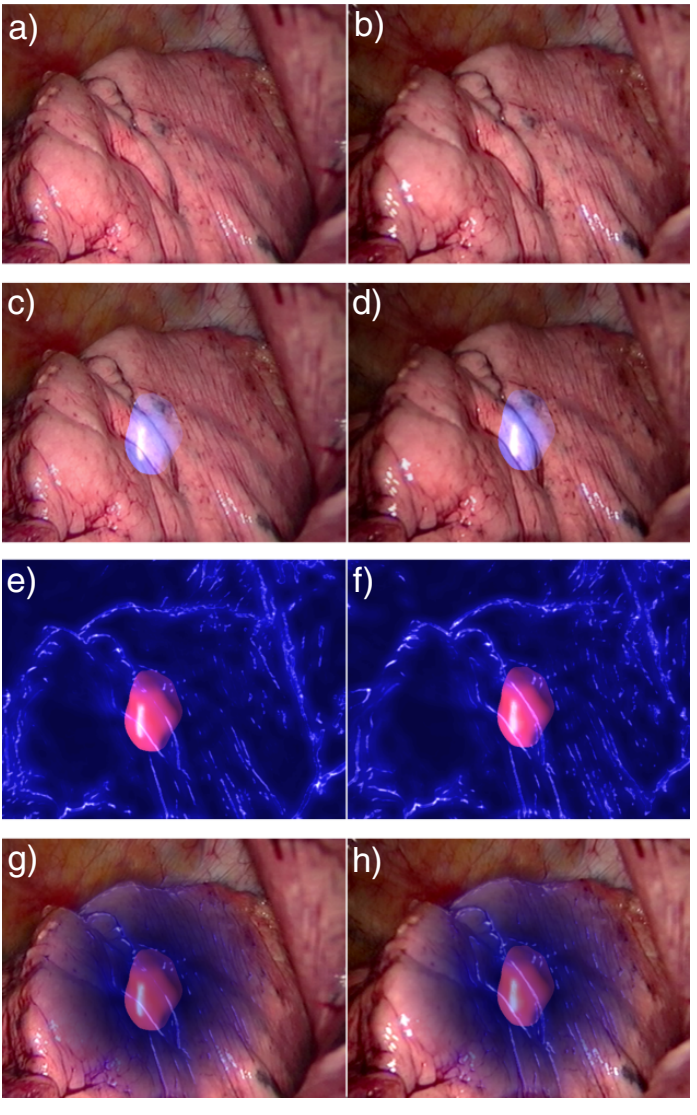

Fig. 3. (a-b) Stereo views of the original robotic assisted lung lobectomy, (c-d) traditional AR overlay, (e-f) *pq*-space NPR AR rendering, and (g-h) fused NPR AR with the original video. The virtual object is rendered at the same depth in all cases.

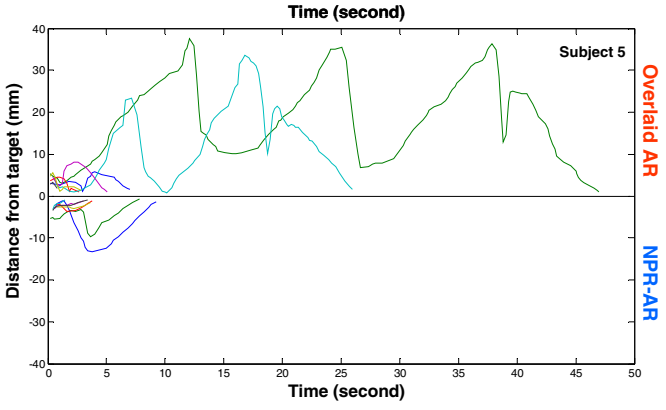

Fig. 4. Tool trajectories of a representative subject of the study showing improved depth perception and instrument targeting by using the proposed NPR AR rendering scheme. For comparison, overlaid AR is shown on the positive and NPR AR on negative axis. More oscillation and variation can be seen when the subject was provided with the traditional overlaid AR overlay.

Table 1. Comparison of median errors in depth perception [in mm] and median task completion times [in seconds] for the ten subjects studied by using the overlaid AR and the new NPR AR methods. The 95% Confidence Interval of each measure is shown in parentheses.

Subject	1	2	3	4	5	6	7	8	9	10
Overlaid AR depth error	2.3 (1)	19.3 (5)	29.7 (4)	8.4 (6)	4.1 (5)	10.6 (3)	5.4 (2)	4.9 (4)	10.1 (2)	10.2 (3)
NPR AR depth error	1.6 (1)	19.1 (5)	27.6 (5)	6.3 (8)	1.7 (2)	7.7 (4)	3.4 (2)	5.1 (3)	2.6 (1)	7.1 (2)
Overlaid AR task timing	3.4 (4)	3.4 (10)	6.0 (4)	5.8 (5)	5.2 (12)	9.4 (5)	4.4 (3)	9.0 (11)	13.6 (6)	13.2 (5)
NPR AR task timing	4.2 (3)	4.2 (4)	3.6 (4)	2.4 (4)	4.2 (2)	4.4 (4)	4.6 (8)	5.0 (5)	2.8 (3)	9.6 (5)

4 Conclusion

In this paper, we have demonstrated the benefits of *pq*-space based non-photorealistic rendering for augmented reality. Experiments with both phantom and *in vivo* lung lobectomy data were used to evaluate the accuracy of the proposed method for AR depth perception. One of the most promising advances in surgical technology in recent years is the introduction of robotic assisted MIS which allows the performance of procedures that are otherwise prohibited by the confines of the operating environment. The use of robotic assisted MIS provides an ideal environment for integrating pre-operative data of the patient for performing image guided surgery. In this regard, the proposed technique is expected to have a significant role in the future deployment of AR to robotic assisted MIS as it is perceptually accurate and computationally efficient due to the linear form of the *pq*-vector derivation and light computational/graphics loading on the rendering engine.

Acknowledgments

The authors would like to thank Danail Stoyanov for the help in stereo camera calibration.

References

[1] Bajura, M., Fuchs, H., Ohbuchi, R.: Merging virtual objects with the real world: seeing ultrasound imagery within the patient 19th annual conference on Computer graphics and interactive techniques, pp. 203-210 (1992)

[2] Edwards, P.J., King, A.P., Maurer, J.C.R., De Cunha, D.A., Hawkes, D.J., Hill, D.L.G., et al.: Design and evaluation of a system for microscope-assisted guided interventions (MAGI). Medical Imaging, IEEE Transactions 19(11), 1082–1093 (2000)

[3] Sauer, F., Khamene, A., Vogt, S.: An Augmented Reality Navigation System with a Single-Camera Tracker: System Design and Needle Biopsy Phantom Trial. In: Dohi, T., Kikinis, R. (eds.) MICCAI 2002. LNCS, vol. 2488, pp. 116–124. Springer, Heidelberg (2002)

[4] Stoyanov, D., Darzi, A., Yang, G.-Z.: A Practical Approach Towards Accurate Dense 3D Depth Recovery for Robotic Laparoscopic Surgery. Computer Aided Surgery 10(4), 199–208 (2005)

[5] Johnson, L.G., Edwards, P., Hawkes, D.: Surface transparency makes stereo overlays unpredictable: the implications for augmented reality. Studies in Health Technology and Informatics 94, 131–136 (2003)

[6] Swan II, J.E., Jones, A., Kolstad, E., Livingston, M.A., Smallman, H.S.: Egocentric Depth Judgments in Optical, See-Through Augmented Reality. IEEE Transactions on Visualization and Computer Graphics 13(3), 429–442 (2007)

[7] Sielhorst, T., Bichlmeier, C., Heining, S.M., Navab, N.: Depth Perception - A Major Issue in Medical AR: Evaluation Study by Twenty Surgeons. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4190, pp. 364–372. Springer, Heidelberg (2006)

[8] Okatani, T., Deguchi, K.: Shape reconstruction from an endoscope image by shape from shading technique for a point light source at the projection center. Computer vision and image understanding 66(2), 119–131 (1997)

[9] Rashid, H.U., Burger, P.: Differential algorithm for the determination of shape from shading using a point light source. Image and Vision Computing 10(2), 119–127 (1992)

[10] Deligianni, F., Chung, A., Yang, G.-Z.: pq-Space Based 2D/3D Registration for Endoscope Tracking. In: Ellis, R.E., Peters, T.M. (eds.) MICCAI 2003. LNCS, vol. 2878, pp. 308–311. Springer, Heidelberg (2003)

[11] Shigemura, N., Akashi, A., Funaki, S., Nakagiri, T., Inoue, M., Sawabata, N., et al.: Long-term outcomes after a variety of video-assisted thoracoscopic lobectomy approaches for clinical stage IA lung cancer: A multi-institutional study. J. Thorac. Cardiovasc. Surg. 132(3), 507–512 (2006)