Gaze-Contingent Motor Channelling and Haptic Constraints for Minimally Invasive Robotic Surgery

George P. Mylonas, Ka-Wai Kwok, Ara Darzi[*], and Guang-Zhong Yang

Royal Society/Wolfson Medical Image Computing Laboratory,
Institute of Biomedical Engineering,
[*] Department of Surgical Oncology and Technology,
Imperial College London, London, United Kingdom
{george.mylonas,k.kwok07,a.darzi,g.z.yang}@imperial.ac.uk

Abstract. The use of master-slave surgical robots for Minimally Invasive Surgery (MIS) has created a physical separation between the surgeon and the patient. Reconnecting the essential visuomotor sensory feedback is important for the safe practice of robotic assisted MIS procedures. This paper introduces a novel gaze contingent framework with real-time haptic feedback by transforming visual sensory information into physical constraints that can interact with the motor sensory channel. We demonstrate how motor tracking of deforming tissue can be made more effective and accurate through the concept of gaze-contingent motor channelling. The method also uses 3D eye gaze to dynamically prescribe and update safety boundaries during robotic assisted MIS without prior knowledge of the soft-tissue morphology. Initial validation results on both simulated and robotic assisted phantom procedures demonstrate the potential clinical value of the technique.

Keywords: Gaze-Contingent control, human robot interaction, haptic constraints, visuomotor fusion, eye tracking, robotic surgery, virtual fixtures.

1 Introduction

It is well recognised that the success of Minimally Invasive Surgery (MIS) is coupled with an increasing demand on surgeons' manual dexterity and visuomotor control due to the complexity of instrument manipulations. Tissue deformation combined with restricted workspace and visibility of an already cluttered environment imposes significant challenges that raise critical issues related to surgical precision and safety. With the availability of robotic assisted MIS, existing research has explored the use of motion stabilisation to simplify the execution of delicate surgical tasks such as small vessel anastomosis [1]. Current work also suggests that global stabilisation of the soft tissue may not be necessary and local stabilization of the foveal field-of-view (FOV) is sufficient [2]. This paves the way for designing adaptive motion stabilisation techniques that are clinically practical. Despite these advances in robotic assisted MIS, the lack of haptic feedback and the physical separation between the surgeon and the patient still impose a potential safety risk. Work published by Rosenberg and others [3][4] introduced the concept of virtual fixtures for enhancing the performance and

safety of MIS procedures by overlaying the operator's motor ability with alternative sensory pathways. However, most of these techniques require prior knowledge of the tissue morphology using pre-operatively 3D imaging data.

The purpose of this paper is to introduce a gaze contingent scheme for realizing real-time haptic constraints *in situ*. The concept transubstantiates visual sensory information into physical entities that can interact with the motor sensory channel. In this paper, we demonstrate how motor tracking of deforming tissue can be made more effective and accurate through the concept of *Gaze-Contingent Motor Channelling* (GCMC). One important part of the proposed GCMC framework is the use of 3D eye gaze to dynamically prescribe and update safety boundaries during robotic assisted MIS without prior knowledge of the soft-tissue morphology. The work presented here further extends our existing experience in real-time eye tracking [5] and to our knowledge this is the first study that successfully bridges the visual and motor modalities by shifting the computational burden towards perceptually enabled channels.
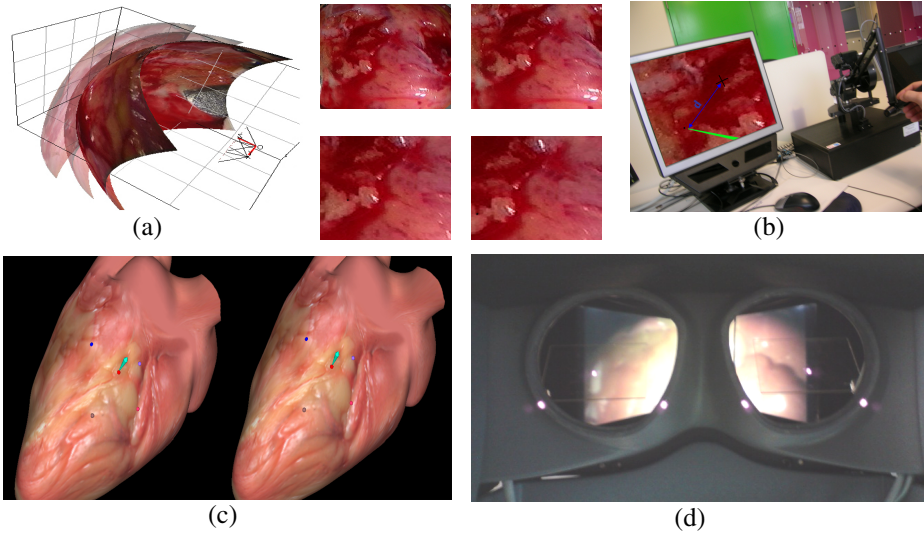
2   Methods

In this paper, the concept of GCMC is studied and demonstrated in both 2D planar and full 3D manipulations. For 2D haptic interactions, a synthetic surgical scene with a deforming soft-tissue surface was used. During simulation, the virtual endoscope camera observes a deformable mesh, which is texture-mapped using footage from an *in vivo* robotic assisted totally endoscopic coronary artery bypass grafting procedure. The mesh is repeatedly deformed by a function combining respiratory and cardiac induced tissue deformation of varying frequencies (Fig.1a). For simplicity, we chose the respiratory component to be a rigid translation and the cardiac component to be represented by a mixture of Gaussians with respect to a selected deformation centre. The two deformations are combined with linear weighting [2]. Finally, the 3D mesh is projected in 2D and a video sequence of the artificially deformed tissue is generated from the extracted image series and displayed on a monitor for subsequent 2D experiments (Fig. 1a, right).

For the 3D experiments, a 3D heart model was reconstructed from MR images of a silicon phantom heart (Chamberlain Group, MA). Surface texture was also acquired and rendered to provide realistic visualisation (Fig. 1c). The model can be deformed between the prescribed key frames and in this paper the principal modes of deformation are limited to the short and long axes of the heart. The two modes of deformation are harmonically combined to represent realistic cardiac motion.

Eye tracking for the 2D experiments is performed using a Tobii x50 stand-alone eye tracker. This is an infrared video-based remote eye tracking system used to record fixation points on a screen at 50Hz with an accuracy of 0.5 degrees and drift <1 degree across the work plane. The system allows for a certain amount of head movement within a working volume ($30 \times 16 \times 20$ cm^3) [6]. Prior to eye tracking, subject specific calibration is necessary and the gaze data is median filtered with a window size of 140ms.

For studying GCMC in 3D, a binocular eye tracker was developed and integrated to a daVinci surgical console (Fig 1d). The eye tracker allows seamless localization of the surgeon's 3D fixation point relative to the operating FOV without obstructing the

surgeon's direct vision. The eye tracker consists of a pair of near infrared sensitive cameras, an array of externally switchable sub-miniature IR emitting diodes at 940nm and a pair of dichroic beam splitters with their cut-off wavelength set above 750nm. This configuration provides appropriate illumination of the surgeon's eyes and their images are captured at a rate of 50 fps. The outputs from the two cameras are then digitized at a resolution of 768×576 and processed in real-time. By calculating ocular vergence, the 3D depth of the fixation point can be determined [7][8].

(a)   (b)

(c)   (d)

Fig. 1. Experimental setup for user studies on the effectiveness of the proposed GCMC framework. (a) The deformable mesh used for the 2D planar experiment where a moving target is tracked by a combined use of eye tracking and a 6-DoF haptic device. The four snapshots show the target on the bottom left of the image. (b) The experimental setup for the 2D GCMC experiments where the configuration of the haptic manipulator and the remote eye tracker is shown. (c) A stereo pair of the deforming 3D cardiac model used for the 3D GCMC experiments along with the ablation targets and the surgical probe. (d) The binocular eye tracker integrated with the daVinci robot stereoscopic console.
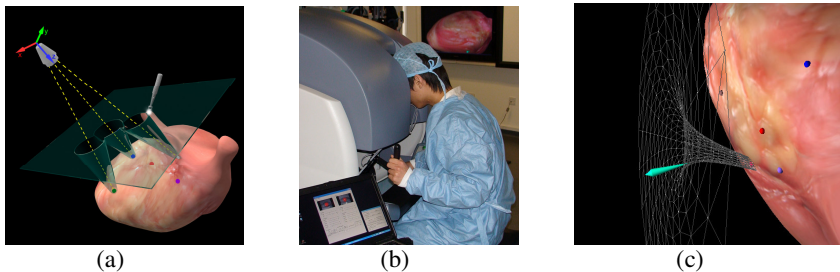
To provide tissue interaction, a haptic device is used to control a virtual surgical tool (proxy) within the FOV. For these experiments, the Phantom Premium 1.5/6DoF Haptic Device from SensAble Technologies (MA) was used. The device provides a workspace with a translational dimension of 381(W)×267(H)×191(D) mm^3 and a rotational dimension of (297, 260, 335) degrees (yaw, pitch, roll). The nominal position resolution in terms of translation is 860dpi/0.03mm and in terms of rotation 0.0023 degrees (yaw and pitch) and 0.0080 degrees (roll). The device is able to provide force feedback and position sensing in 6 DoF.

The GCMC haptic interactions within the simulated 2D surgical scene are applied based on the relative separation between the user's fixation point on the screen and

the 2D position of the manipulated surgical instrument (both on the screen frame-of-reference as shown in Fig 1b). For the 3D surgical scene, the forces exerted by the fixation point to the haptic manipulator are based on the Euclidean distance between the 3D fixation and the position of the manipulated surgical instrument as illustrated in Fig. 1c. In both 2D and 3D cases, the force direction is towards the fixation point. Hand-motion scaling can also be introduced and different force profiles are implemented to assess the optimal force profile to be used in practical applications.

2.1 Experimental Setup

To assess improved accuracy in 2D targeting by the use of the GCMC framework, a moving target is introduced to the deforming 2D tissue. The target accurately follows the tissue deformation and during a full cycle it spans an area corresponding to 8.5^{o} of horizontal and 5.5^{o} of vertical visual angle from a viewing distance of 60 cm. The coordinates of the target are registered and known on a per video-frame basis. The composite frequency of oscillation of the target is constant at 0.23Hz. For assessing the effectiveness of GCMC in manually tracking the target on the deforming tissue, 10 subjects were recruited for this study. They were presented with the video of the tissue and all the subjects were given as much time as they needed to familiarize themselves with the Phantom manipulator and the basic concept of GCMC. Their task was to track the moving target as accurately as possible with the tip of the virtual instrument. During the experiment, the force profile and the motion scaling of the haptic device were changed according to the order as they appear in Table I. A high value of hand-motion scaling results in small hand movements being translated into large displacements of the virtual tool. A unit scaling factor provides a 1:1 mapping between space and screen coordinate system. For tasks 1 and 2, no force was exerted on the phantom manipulator. For the remaining tasks, force feedback was enabled and exerted as a function of the 2D distance between the fixation point and the tool tip.
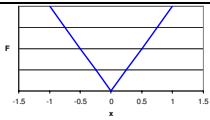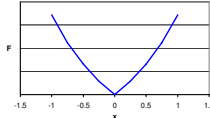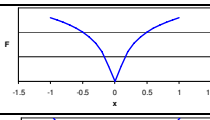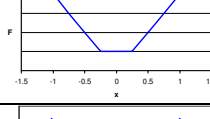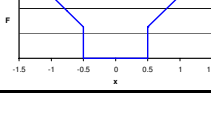

(a)   (b)   (c)

Fig. 2. (a) The conical pathways carved out from the safety boundary which is positioned based on the the fixation during the experiment. (b) The experimental setup for the 3D GCMC experiments. (c) An illustrarion of the deformable conic pathways that can be used for prescribing dynamic active constraints.

For the second part of the 2D GCMC study, the optimal force profile identified in the previous experiment step was adopted. It was used to assess the method during non visually-servoed motor-tracking and a number of force graduations. The same

video, setup and task were used as with the first experiment with the same 10 subjects. Initially, visual servoing was enabled using a high spring constant. Gradually, the spring stiffness was reduced to zero. To minimise the learning effect, after a number of tissue deformation cycles over a force profile, the video was paused and the subjects were asked to look away from the screen and leave hold of the haptic manipulator. After 20sec, they were cued with a timed audible signal that the experiment was about to resume. Another 10sec were available for them to grasp the haptic manipulator and start fixating on the target to be tracked. During both experiments, time, fixations, target and tool-tip coordinates were recorded for post processing. Analysis was based on measuring the error of the tool-tip position and the actual position of the tracked target.

Table 1. The different force and hand-motion scaling profiles in the order as they have been introduced during the 2D experiment. The horizontal axis x represents the distance between the fixation point and the tool-tip and the vertical axis (F) represents the normalized force exerted by the haptic device. Parameter k is the spring constant. The same principle applies to other axes.

Task	Force Feedback Profile	Description	Formula
1,2		High (1) and 1:1 (2) hand-motion scaling	$F = 0$
3, 4		High (3) and 1:1 (4) hand-motion scaling, linear spring	$F = k \cdot x$
5		1:1 hand-motion scaling, exponential force profile	$F = k \cdot \left(e^{|x|} - 1\right)$
6		1:1 hand-motion scaling gain, logarithmic force profile	$F = k \cdot \log\left(|x| + 1\right)$
7		1:1 hand-motion scaling gain, linear spring and constant force well	$F = k \cdot x$ for $-t > x > t$ else $F = const$
8		1:1 hand-motion scaling gain, linear spring and zero-force well	$F = k \cdot x$ for $-t > x > t$ else $F = 0$

   To examine how the proposed GCMC could be used to provide haptic constraints in robotic assisted MIS procedures, 6 subjects were asked to observe the deforming 3D synthetic cardiac tissue through the daVinci stereoscopic console (Fig 2b). Prior to each trial, all subjects underwent user specific binocular calibration. During the experiment, eye tracking provided the 3D fixation point, which was also fed to the
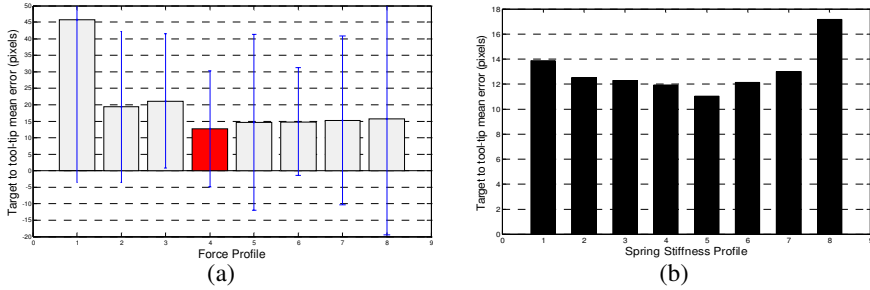
6DOF haptic manipulator in real-time. When GCMC was enabled, the user felt forces exerted from the fixation point to the virtual instrument in 3D. As before, all subjects were given time to familiarize themselves with the setup. The task prescribed was to approach a preset fiducial with the virtual instrument and ablate it by pressing the manipulator button. A total of 10 fiducial markers were displayed in sequence but in random order. Throughout the experiment, time, instrument trajectory and action events were recorded and the experiments were repeated with and without GCMC. When GCMC was enabled, a planar safety boundary was also provided to prevent the instrument from penetrating into the tissue and cause inadvertent damage. The boundary was parallel to the x-y plane of the virtual camera and its position on the z axis (depth) is determined by the depth of the fixation that was closer to the camera origin minus a small shift towards the origin (Fig. 2a). The only way for the virtual instrument to penetrate behind the safety boundary and touch the tissue was through narrow conical channels that were created based on the position of each fiducial. The apex of a cone was inserted under the tissue surface at a small amount so that the virtual instrument was relatively free to move and track the deforming tissue. In this experiment, the vector connecting the apex and the base centre coincided with the vector connecting the camera origin and the fiducial. Only one cone was enabled at a time and this selection was based on the user's fixation point and a preset threshold.

3   Results

Fig. 3a depicts the results for the 2D experiments for all the subjects studied. It is evident that the error reduces dramatically when GCMC is enabled. When a high hand-motion scaling factor is used (columns 1 and 3), the improvement is over 50%. Columns 2 and 4 show the non-servoed and servoed errors respectively for the 1:1 hand scaling and the associated accuracy improvement (30%). From the same graph, it can also be observed that the optimal force profile is the linear spring model shown in column 4. For this reason, it was chosen for all the subsequent tasks. Perceptually, this profile also gives the best smoothness and haptic consistency subjectively.

Fig. 3b shows the error analysis of the linear spring force model with different spring constants. Moving from left to right, the first column represents a stiff spring exerting high force and the last column represents a spring with zero stiffness. Column 7 has a very low stiffness spring that is barely noticeable through the haptic device but still remains effective for reducing the error over the non-servoed task.

For the 3D experiments carried out on the daVinci robot, we have compared the amount of tissue penetration with/without the use of GCMC. Tissue penetration occurs every time the virtual tool penetrates the surface of the tissue. Table II demonstrates the results for all the subjects studied. It is evident that there is a clear reduction in terms of the amount of penetration when GCMC is enabled with an improvement of 22.33% on the number of times the surgical tool has penetrated the tissue and 29.08% on the average penetration duration. Also the accuracy in manually targeting the ablation target has improved by 37.72%.

Fig. 3. a) Error and performance analysis for all the subjects studied by using the described GCMC concept when different force profiles are used. The first two bars correspond to free motor control with decreasing gain. The last six bars represent errors with visual servoing enabled over different force profiles according to Table I. The error bars show the standard deviation b) The effect of spring stiffness on the accuracy of targeting when force profile 4 in (a) is adopted. Force decreases from left to right and the last column is the resulting error with forces disabled (*i.e.*, free motion without haptic constraints).

Table 2. Results of the 3D experiments comparing the amount of estimated tissue penetration and the accuracy of manually targeting the ablated target, with and without the use of GCMC control. The results are averaged over all the participating subjects.

Method	Number of times	Total duration (sec)	Max duration (sec)	Average duration (sec)	Total length (mm)	Max length (mm)	Average length (mm)	Error in targeting (mm)
Free	17.17	33.25	4.52	1.62	233.14	48.04	12.69	19.75
GCMC	13.33	16.47	2.59	1.15	101.29	23.92	7.05	12.30
% Improv.	22.33	50.46	42.55	29.08	56.55	50.22	44.47	37.72

4   Discussion and Conclusions

In this paper, we have demonstrated the use of GCMC for improved targeting and hand-eye coordination. We have also shown the flexibility of the proposed method in setting up dynamically updated haptic constraints to avoid accidental damage to the soft tissue. The results derived have shown the effectiveness of the proposed concept in improving safety and task performance during robotic assisted MIS without prior knowledge of the 3D tissue morphology. In this study, post-experiment survey revealed that the use of GCMC enhanced the users' conceptualization of the task by increasing their confidence levels and making the task to be more natural and easier.

This study re-enforces findings in previous studies that for pointing and grasping tasks, the eye often precedes the hand motion [9][10]. It has been shown that motor tasks are visually guided and that eye-hand latency varies for different task contexts and across different subjects. The proposed GCMC framework is potentially useful in reducing such latencies by indirectly linking the motor and the visual channels for

delicate micro-surgical tasks. Our study also builds on previous findings in the use of virtual fixtures for improving task performance and accuracy. One significant feature that differentiates the current approach from previously published work is the use of the surgeon's 3D fixation point as a dynamic virtual fixture and makes use of real-time *in situ, in vivo* structural information directly acquired from ocular vergence. For clinical applications, this is important as it avoids the difficult task of pre- and intra-operative registration, which for cardiothoracic and abdominal procedures remains a major research challenge. It should be noted that this paper does not address instability issues arising when the user's gaze is not detected or directed towards a location other than the desired target. In practice, aberrant eye movements can be detected by using saccade velocity and fixation duration. It is also possible to use an auxiliary button/switch to engage the instrument. Under the current framework, the hand is not overpowered by the haptic manipulator throughout the experiments and the subjects are always in control of the surgical tool. Future work will concentrate on implementing deformable pathways and boundaries that will allow dynamic haptic constraints for more complex anatomical structures.

References

1. Wimmer-Greinecker, G., Deschka, H., Aybek, T., Mierdl, S., Moritz, A., Dogan, S.: Current status of robotically assisted coronary revascularization. Am. J. Surg. 188(4A Suppl.), 76S-82S (2004)
2. Mylonas, G.P., Stoyanov, D., Darzi, D., Yang, G.-Z.: Visual assessment of Gaze-Contingent Motion Stabilization for Minimally Invasive Robotic Surgery: A study of perceptual quality. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part II. LNCS, vol. 4792, pp. 660–667. Springer, Heidelberg (2007)
3. Rosenberg, L.B.: Virtual fixtures: Perceptual Tools for Telerobotic Manipulation. In: Proc. IEEE Virtual Reality Annual Intl. Symposium, pp. 76–82 (1993)
4. Park, S., Howe, R.D., Torchiana, D.F.: Virtual Fixtures for Robotic Cardiac Surgery. In: Niessen, W.J., Viergever, M.A. (eds.) MICCAI 2001. LNCS, vol. 2208, pp. 1419–1420. Springer, Heidelberg (2001)
5. Yang, G.-Z., Dempere-Marco, L., Hu, X.P., Rowe, A.: Visual search: psychophysical models and practical applications. Image and Vision Computing 20, 291–305 (2002)
6. Tobii technology. User Manual (2003), `http://www.tobii.se`
7. Mylonas, G.P., Stoyanov, D., Deligianni, F., Darzi, A., Yang, G.-Z.: Gaze-Contingent Soft Tissue Deformation Tracking for Minimally Invasive Robotic Surgery. In: Duncan, J.S., Gerig, G. (eds.) MICCAI 2005. LNCS, vol. 3749, pp. 843–850. Springer, Heidelberg (2005)
8. Mylonas, G.P., Darzi, A., Yang, G.-Z.: Gaze Contingent Depth Recovery and Motion Stabilisation for Minimally Invasive Robotic Surgery. In: Yang, G.-Z., Jiang, T. (eds.) MIAR 2004. LNCS, vol. 3150, pp. 311–319. Springer, Heidelberg (2004)
9. Pelz, J., Hayhoe, M., Loeber, R., et al.: The coordination of eye, head, and hand movements in a natural task. Exp. Brain Res. 139, 266–277 (2001)
10. Leong, J., Atallah, L., Mylonas, G.P., Leff, D., Emery, R., Darzi, A., Yang, G.-Z.: Investigation of Partial Directed Coherence for Hand-Eye Coordination in Laparoscopic training. In: Dohi, T., Sakuma, I., Liao, H. (eds.) MIAR 2008. LNCS, vol. 5128, pp. 270–278 (2008)