Development and Evaluation of an Augmented-Reality Training System for Planning Brain Tumour Resection Interventions

Kamyar Abhari[1]
kabhari@robarts.ca
John S.H. Baxter[1], Ali R.Khan[1], Elvis S. Chen[1], Chris Wedlake[1], Terry Peters[1], Sandrine de Ribaupierre[2], Roy Eagleson[1]

[1] Robarts Research Institute, Western University, London, Canada
[2] Clinical Neurological Sciences, Western University, London, Canada

Abstract

By regarding medical image understanding as a form of visualization involving a Human-Computer Interface, the onus is on system designers to consider the specific capacities and constraints of the human perceptual system in the design of visualization systems. From the perspective of Cognitive Science, the design and evaluation of these systems must be informed by the results of Basic Science in the domains of Perception, Scene Understanding, and Perceptual-Motor Control. In this paper, we present the results of such an approach to the design of an Augmented-Reality based visualization system for Neurosurgical Planning and Neuro-Anatomical training. We hypothesize that the proposed AR system facilitates training of novice residents to plan tumour/hematoma resection interventions. To test our hypothesis, a number of experiments were conducted where subjects were asked to perform relevant spatial judgment tasks using three conventional visualization approaches as well as our proposed AR system. Our preliminary results indicate that, compared to traditional methods, the proposed AR system a) greatly improves the user performance in tasks involving 3D spatial reasoning about the tumour relative to the anatomical context, b) reduces error associated with mental transformation, and c) supports generic spatial reasoning skills, over this range of sensory-motor tasks.

1 Introduction and Clinical Motivation

The primary goal of a visualization system is to represent data in such a way that relevant information is made explicit, facilitating comprehension. Learning 3D anatomical structures and spatial relationships sometimes requires the trainee to visualize anatomical structures from within, adding a level of complexity for the trainee as well as challenges for the design of visualization systems. One example is tumour resection interventions. Approximately 23,000 Americans and 2,800 Canadians were diagnosed with primary brain tumours in 2012, resulting in 13,700 and 1,800 deaths respectively [1] [2]. Compared to alternatives, surgical resection is the most recommended option [3] to treat brain tumours. Pre-operative planning involves identifying optimal surgical paths and entry points based on a number of criteria in order to minimize post-surgery complications, and teaching these skills is a challenge. One must take into account functional areas, white matter tracts, and major vessels while planning a trajectory. Additionally, determining the shortest path from the skull to the tumour and aligning the surgical trajectory with the longest axis of the tumour, can further reduce damage. Formulating the optimal path and entry point demands the perception of and complex reasoning about spatial relationships between the tumour and other key structures. Conventional visualization approaches for planning these interventions involve examining 2D orthogonal slices of pre-operative MR images, but the inherent limitations of these 2D views and the reliance on more complex spatial reasoning can slow the process of planning and make it prone to error [4]. While it is clear that users can be trained to perform well on non-intuitive tasks, they are more prone to error added cognitive demands. Although this is not controversial, novel visualization methods are not generally evaluated against task-relevant user performance metrics, and in the following section we present a methodology for doing so.

2 Materials and Methods

The AR system is comprised of a head phantom and off-the-shelf AR eyewear (Vuzix 920AR, Vuzix corporation, Rochester, NY), both of which are tracked using an optical tracking system (Polaris, NDI, Canada), (Figure 1). The Vuzix eyewear is equipped with twin cameras and displays to record and display stereo images. Tracking the head phantom and the Vuzix goggles allows for the correct fusion of virtual and physical spaces. Our visualization approach involves making use of a tracked stylus as a manipulandum.

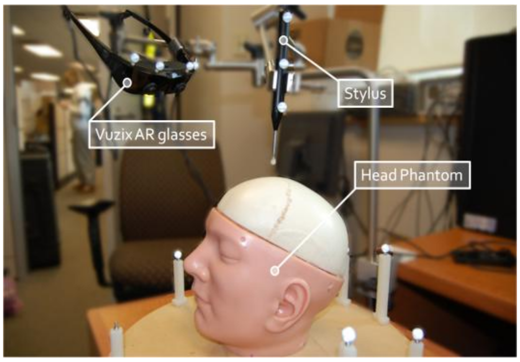

Figure 1: Our AR system includes Vuzix eyewear, a head phantom, and a tracked stylus

Using a window-and-context paradigm [5], the tracked stylus controls a focus window in which a visual scene is visualized (Figure 2 (left)). Voxels in this window are rendered using direct volume rendering (DVR) and two-dimensional transfer functions (2D TF), a standard approach for reducing ambiguity and conveying subtle surface properties to the user [6]. Furthermore, early ray termination is explicitly triggered by an additional transfer function, which prevents rendering the volumetric data outside the focus window. This results in a keyhole-like aperture into the volume. The size of the aperture can be adjusted using a multifunction USB-knob (Griffin Tech., TN, US). Similar to virtual windows [7], the aperture reduces depth misperception to some extent, and draws attention to a region-of-interest while preventing cognitive overload [8]. Cel-shading, a non-photorealistic shading technique, was used to enhance perception of object boundaries, improving the use of occlusion as a perceptual cue. These boundaries are detected by sudden changes in the depth of penetration of the virtual rays, allowing areas of larger change to be shaded more heavily. It has been shown that enhancing contours using cel-shading improves the perception of continuity and depth [9]. Distance shading is also employed, providing additional cues to relative depth. To further increase the usability of the system, the opacity of each individual tissue type (e.g. tumour, cortex), and the strength of the shading techniques are easily adjustable in the provided interface.

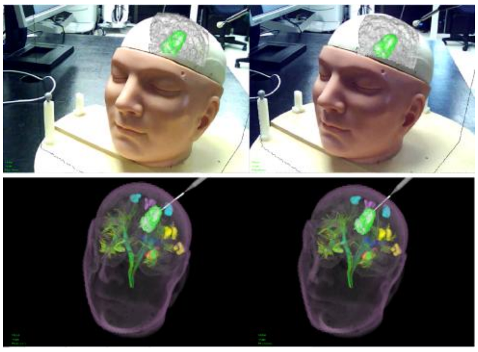

Figure 2: Stereo images of the proposed AR (left) and VR system (right)

While the AR mode provides a better understanding of the spatial relationship between the virtual data and the physical context, visualization of too much information through the key-hole aperture can lead to visual clutter. Therefore, we have incorporated a Virtual Reality (VR) mode in which virtual data can be visualized in its complete form by halting the early ray termination process (Figure 2 (right)). Additionally, specific DTI tracts and eloquent areas of the brain as well as a virtual representation of the stylus and its trajectory are also included. The location of these areas is crucial for planning tumour resection interventions as avoiding them mitigates the risk of post-surgical complications. Users can toggle the visibility of these tracts/regions within either hemisphere. A foot pedal was provided to participants allowing them to switch back and forth between these two modes, in order to benefit from both AR and VR.

2.1 Experimental Methodology and Objective Metrics of Performance

Phase 1 & 2: The objective of our evaluation process was to compare the user performance associated with conventional approaches against our AR system. Conventional techniques consist of two-dimensional slice-by-slice (2D), crossed-plane (XP), and three-dimensional volume rendering (3D) (Figure 3). Ten novices (8 Male, 2 Female, all graduate students), participated in our validation study and were asked to perform a series of relevant spatial reasoning tasks while exploring data via available techniques. Experiments were conducted in two different phases. In the first phase, each task was defined to isolate one of the planning principles mentioned earlier in the introduction: 1) finding the maximal distance from the target to a nearby critical structure, and thereby avoiding it; 2) finding the shortest path to the target from the surface of skull; and 3) determining the longest axis of the target. Based on these criteria, subjects were asked to use of the head phantom and stylus to indicate the optimal entry point/surgical path. Stimuli consisted of the head phantom's CT images which were modified to meet the requisites of each specific task. Simulated structures derived from patient anatomical data were used to increase the clinical validity, while target structures were simulated to increase our control over the experimental design.

In the second phase, subjects were asked to perform the same tasks as the first phase, but in this stage, the ground truth entry points and surgical paths were provided as synthetic lines.

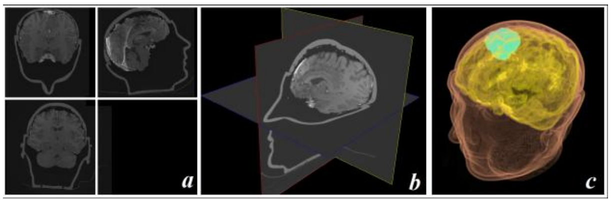

Figure 2: Conventional visualization methods: a) slice-by-slice (2D), b) crossed-plane (XP), and c) three-dimensional direct volume rendering (3D)

Our empirical methodology was designed; 1) to study the effect of different visualization approaches on user performance, and 2) to determine whether providing visual assistance can diminish the difference between available visualization techniques. Rotational error is measured as the deviation (in degrees) between the chosen and optimal paths. Translational error is measured as the Euclidean distance (in mm) between the optimal point of entry and that selected by participants (Figure 4).

Phase 3: Although the first two phases may demonstrate the efficacy of the proposed system in assisting subjects in performing simple spatial tasks, the efficacy of the AR environment within a clinical context is still questionable. Thus, we extended our evaluation to include clinically relevant data and expert neurosurgeons and neurosurgery residents. In order to increase the statistical power, a large set of patient-specific images was required. Tumour data (randomly selected from a set of previously segmented tumours[1]) was added systematically to different regions of an MRI dataset. Similar to the previous phases, each subject was asked to perform two tasks: finding the point on the skull with shortest-distance to tumour, and the tumour's longest axis estimation.

3 Results and Discussion

Phase 1 and 2 involved 12 trials (3 tasks x 4 visualization methods) where participants (n = 10, no prior training) were presented with a randomized collection of synthetic data. Phase 3 involved 64 trials[2] (32 trials per task) in which the patient MR data was randomly selected from the database and displayed in the 4 different modalities described earlier.To minimize the effects of learning and fatigue, the stimuli, the visualization mode, and the task were all randomized. The user overall performance was calculated by averaging the rotational and translational error over each visualization technique.

Phase 1: A multivariate ANOVA test indicated that the mode of visualization was indeed significant in the first phase (rotational error: $p<0.05$, translational error: $p<0.05$). This level of significance was derived from the Šidák correction which would

[1] DTI challenge workshop, MICCAI 2010-11
[2] 48 trials for one of our experts

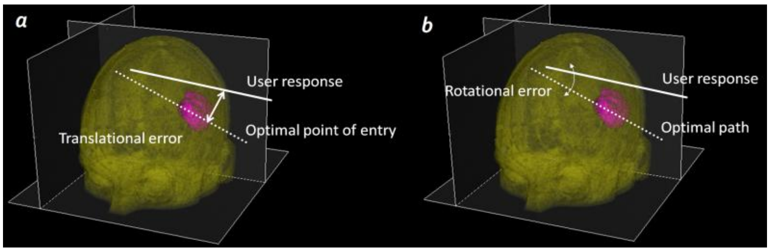

Figure 4: Rotational/translational errors as metrics to measure users' performance

lead to a combined level of significance of 1%. No interaction effect between visualization method and task was observed. Post-hoc analysis using Tukey HSD test indicated that the difference between 2D/XP visualization environments and 3D/AR were statistically significant. However, the difference between the 3D and AR environments was not statistically significant. Therefore, it can be concluded that when there is no visual assistance, a) the method of visualization significantly affects the user performance error, and b) the impact of method of visualization is not affected by the task performed (and vice versa). This indicates the generic usability of the visualization regardless of task. Significant improvement of performance in AR/3D demonstrates that 3D perception of the target location/orientation in 2D/XP can be facilitated with appropriate visualization methods.

Phase 2: Unlike the previous stage, significant interaction among factors was observed, i.e. the magnitude of difference between environments depends in part upon the task performed. A test of significance for each revealed that most interactions occur in 2D, XP, and 3D environments. This indicates lower usability scores for these environments. Additionally, the increased variation in rotational and translational error in 2D and XP illustrates the veridicality of 3D and particularly AR environments. Significant improvement of performance in AR while providing visual assistance illustrates that the AR mode of presentation reduces the mental transformation load in the 2D/XP/3D modes.

Phase 3: Speed-accuracy trade-off was taken into account by calculating the index of performance ($I_p = 1/_{\text{time}} \log_2\left(2\mu/_{\text{Effectiveerror}}\right)$) in accordance with Fitts' methodology [10]. Our preliminary results show that for the longest axis task, I_p was significantly higher ($p < 0.05$) in AR compared to 2D and XP (μ_{AR}= .093, μ_{3D}=.066, μ_{2D}=.045, μ_{XP}=.035), and for the shortest axis, it was significantly higher in AR and 3D compared to XP (μ_{AR}= .085, μ_{3D}=.073, μ_{2D}=.055, μ_{XP}=.040). All subjects with a neurosurgery background performed better in all visualization modalities with the exception of AR for identifying the longest axis where novices performed better than residents.

4 Conclusion

Locating structures such as tumours and tracts and perceiving the spatial relationships between them is necessary for successful neurosurgical planning, which is heavily influenced by the visualization and interaction in the planning environment. In this study, we investigated whether performing these spatial tasks could be facilitated by visualizing

MR images through DVR in an AR environment. We conducted a number of experiments where subjects performed relevant spatial judgment tasks under four different visualization approaches. Our preliminary results indicate that, AR environments could improve spatial reasoning with respect to clinically-relevant tasks, for trainees. Phase 3 data illustrates the potential of using AR in neurosurgical training. However, more number of subjects is required to increase the power of such study. Nevertheless, the proposed work has the potential to improve the quality of tumour resection planning purely from the perspective of perceptual enhancement, despite the fact that it does not trump skilled observational training. When using AR and 3D modes, the performance of the novice group is enhanced. Performance in the AR environment is more independent from the specific spatial tasks indicating that this presentation of visual information has enhanced utility for generic tasks.

Acknowledgment

This project was supported by the Canadian Institutes for Health Research (Grant MOP 74626), the National Science and Engineering Research Council of Canada (Grants R314GA01 and A2680A02), the Ontario Research and Development Challenge Fund, the Canadian Foundation for Innovation and Ontario Innovation Trust, and GRAND-NCE. Graduate student funding for K. Abhari was provided scholarships from the National Science and Engineering Research Council of Canada (CAMI).

References

[1] American Cancer Society, Cancer Facts and Figures, 2012.

[2] Canadian Cancer Statistics, Public Health Agency of Canada, 2012.

[3] National Cancer Institute (NCI) booklet: NIH Publication No. 09-1558

[4] M. Hegarty, M. Keehner, C. Cohen, D.R. Montello, Y. Lippa, G.L. Allen: the role of spatial cognition in medicine: Applications for selecting and training professionals, Applied spatial cognition: From research to cognitive technology, Lawrence Erlbaum Associates Publishers pp. 285-315, (2007)

[5] Prastawa, M., Bullitt, E., and Gerig. G.: Simulation of Brain Tumors in MR Images for Evaluation of Segmentation Efficacy. Medical Image Analysis (MedIA). v.13(2), 297-311, 2009

[6] Drebin, R.A., Carpenter, L., Hanrahan, P.: Volume rendering. SIGGRAPH Comput. Graph. vol. 22(4), pp. 65--74. (1988)

[7] Abhari, K., Baxter, J.S.H, de Ribaupierre, S., Peters, T., and Eagleson, R.: Perceptual Improvement of Volume-Rendered MR Angiography Data using a Contour enhancement Technique. International Society for Optics and Photonics (SPIE), 8318, 831809, USA. (2012)

[8] Baxter, J.S.H., Peters, T.M., and Chen, E.C.S.: A unified framework for voxel classification and triangulation. Proc. SPIE 7964, 796436. (2011)

[9] Interrante, V., Fuchs, H., and Pizer, S.M.: Conveying the 3D Shape of Smoothly Curving Transparent Surfaces via Texture. IEEE Trans. Visualization and Computer Graphics, vol. 3(2), pp. 98--117. (1997)

[10] Fitts, P.M., The information capacity of the human motor system in controlling the amplitude of movement. J Exp Psychol, 47(6): p. 381—91. (1954)