Medicine Meets Virtual Reality
J.D. Westwood et al. (Eds.)
IOS Press, 1999

Virtual Hand Laboratory Meets Endoscopic Surgery

C.L. MacKenzie[1], Ph.D., E.D. Graham[2], Ph.D.,
C.G.L. Cao[3], M.Sc., and A.J. Lomax[1], M.D.,
Simon Fraser University[1],
DSI Datotech Systems Inc.[2],
University of Toronto[3]

We describe two recent research projects: the Virtual Hand Laboratory, and Remote Manipulation in Endoscopic Surgery. The Virtual Hand Laboratory (VHL) is a prototype experimental tool for investigating human visuomotor coordination for object manipulation in augmented and virtual environments. The Remote Manipulation in Endoscopic Surgery (RMES) project examined surgeon's use of viewing and manipulating technologies in laparoscopic surgery, both in clinical and experimental settings. Current research brings together these two parallel research projects (VHL and RMES), for applications in planning and real-time execution of surgical procedures as well as for surgical training. We outline our research directions and detail current activities on superposition of display space on the workspace for the surgeon's hands.

1. Introduction

Goal-directed human activities include skilled hand movements [1,2,3]. The planning and control of hand movements have been a primary research interest, examined recently in two application domains: human-computer interaction, and endoscopic surgery. In this report, we overview two recent, independent research projects at Simon Fraser University: Virtual Hand Laboratory, and Remote Manipulation in Endoscopic Surgery. We note parallels and similarities, and describe current directions to dovetail these two paradigms in the pursuit of user-centred, intelligent tools for surgery.

2. Virtual Hand Laboratory (VHL)

The Virtual Hand Laboratory (VHL) shown in Figure 1 is a prototype experimental tool for investigating human visuomotor coordination underlying object manipulation in augmented and virtual environments. This is motivated by a natural prehension metaphor [2], and views of direct manipulation in human motor control. Part of a strategic project on Hand-centred Input, we married high-end 3-D graphics workstations with leading 3-D motion analysis systems to study 3-D interaction. Calibrated to physical grasping space, the system provides stereoscopic, head-coupled displays, superimposed on interactive workspace for the hand. The system uses an OPTOTRAK 3-D motion analysis system (Northern Digital Inc.) to track head, eyes, hands and objects with spatial precision of 0.1 mm. These data drive on-line, in real-time, the Silicon Graphics display with temporal lag of 17 to 25 ms (about 1 frame, at 60 Hz). There are two, daisy-chained monitors. The display from one monitor (placed upside down), is reflected in a half-silvered mirror, to create graphic objects in the virtual workspace below. Subjects, wearing CrystalEYES shutter goggles, can interact with physical, virtual or augmented (both physical and graphical) target objects in the workspace.

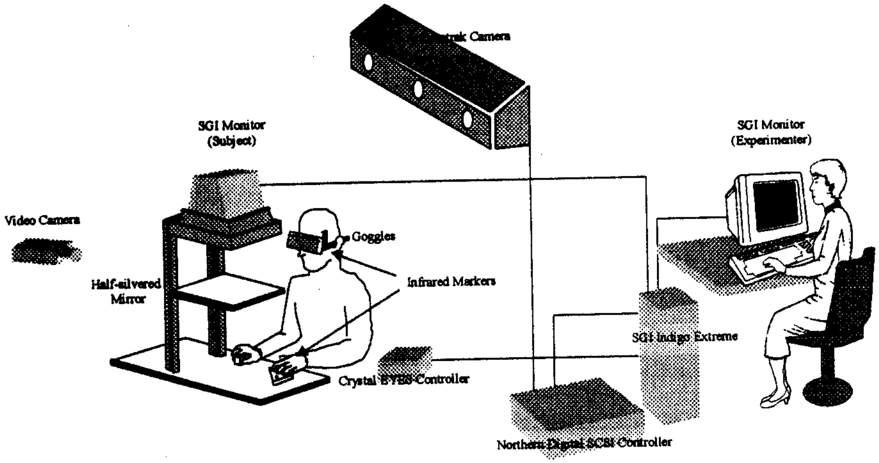

Figure 1. The Virtual Hand Laboratory at Simon Fraser University marries 3-D graphics and 3-D motion analysis technologies. The subject interacts with physical, virtual or augmented objects while viewing a stereoscopic, head-coupled display, driven by the subject's and object's movements, detected on-line in real-time. See text for system details.

The VHL has greater spatial and temporal resolution than most other virtual environment systems; most subjects experience a compelling sense of presence.

Using as our foundation theory and research on natural human prehension [2], we have been conducting simple, systematic, controlled experiments on pointing, grasping and object manipulation using the Virtual Hand Laboratory since 1995.[4,5,6,7,8] In one series of experiments, we examined 3-D kinematics of pointing movements in: the standard human-computer interaction (HCI) setup, the virtual hand lab setup (Virtual), and natural pointing at physical targets on the table top (Physical). Figure 2 shows how display space is related to the workspace for the hand in the three conditions. From a constant starting location, subjects made pointing movements with the extended index finger to match a cursor to graphic targets on the monitor (for HCI), superimposed on the table surface (for Virtual Condition), or to painted, physical targets on the table surface (for Physical). The cursor was a red arrow on the monitor, driven graphically by 3-D position data of infrared markers on the index finger. In the physical condition, a cardboard red arrow was physically mounted on the index finger, and the subject viewed this directly.

The Index of Difficulty (ID in bits) of the pointing movements was determined by Fitt's formulation, and comprised various combinations of movement amplitude and target width. Movement times (MT) followed Fitts' Law, and were longest in the HCI pointing configuration, then the virtual pointing condition which was only slightly longer than physical pointing, as shown in Figure 2. Further, increases in the movement time for smaller targets were primarily reflecting a longer time spent in the deceleration phase of the movements. The experimental data in Figure 2 were best fit by a two-part model of MT, available in [5]. In addition, these experiments showed a greater sensitivity of MT to manipulations of movement amplitude than target width. This movement behaviour of the hand, rather than the cursor leads to more effective modelling of human performance [4,5,6].

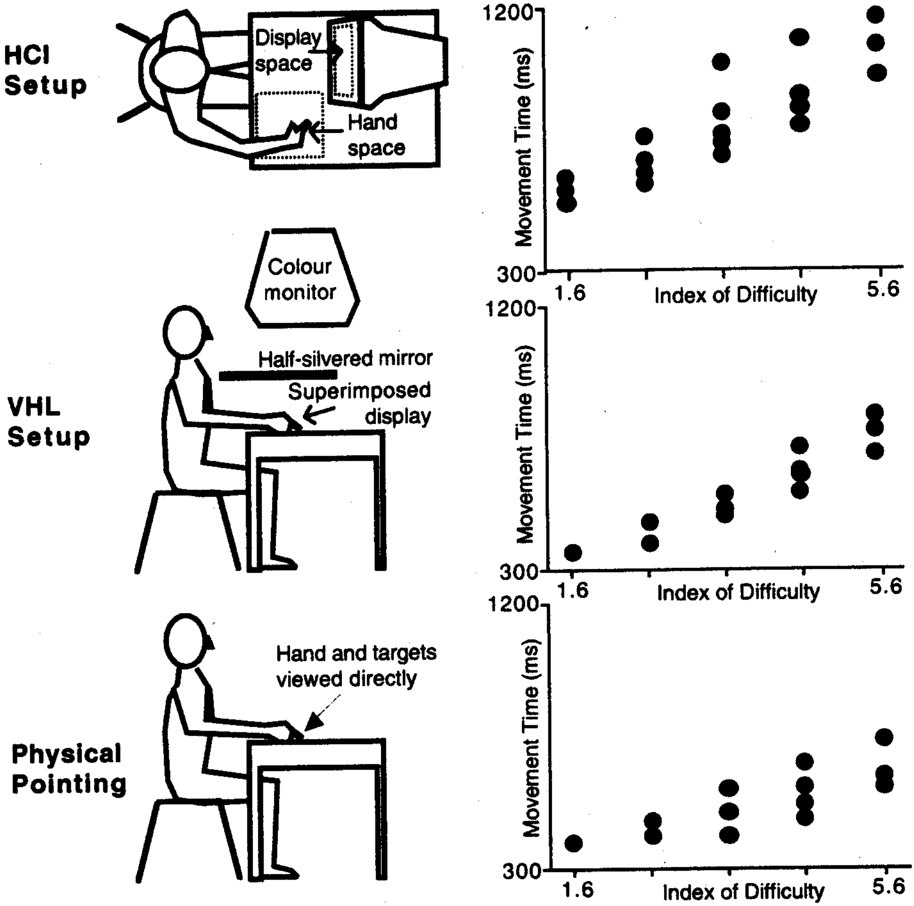

Figure 2. Controlled experiments on pointing on computer displays [Adapted from 6]. Aiming to targets was examined in the standard Human-computer interaction (HCI) setup, with a variety of gains or magnifications. Shown and discussed here for only gain = 1. As well, aiming to targets was compared in the Virtual Hand Laboratory setup (VHL) to natural physical pointing (Physical). The VHL setup superimposed display space on the workspace of the hand, most like natural pointing.

We have contrasted prehension in virtual and augmented environments to natural environments, and situations where display space is separated from the hand's workspace. In a large number of experiments, we have shown that time, errors and 3-D kinematics for pointing and grasping in the VHL setup are very similar to natural pointing and grasping. We are interested in understanding the differences. [6,7,8] When display space is separated from the hand's workspace, there is a significant decrement in performance, suggesting additional visuomotor and spatial transformations.

Two important concluding comments about the VHL research to date. First, our performance measures (time, accuracy, 3-D kinematics and derived parameters) of humans interacting in the VHL parallel very closely the same measures taken for pointing and grasping in a natural, physical environment. This is in marked contrast to interaction measures in other Virtual Environments in the literature. For example, the completion time for a docking task requiring object translation and orientation by the human hand can be three to ten times longer than in the VHL [7,8]. We do not know which components/aspects of the VHL system provide this advantage, but it may be the high temporal and spatial resolution as well as the rigorous calibration steps in the system. Second, to date we have used simple tasks and "impoverished environments", limited to simple geometric shapes for objects (e.g., cubes, pointers, wireframe targets), in the computer-generated images. This is in contrast to the complex surgical tasks and rich camera-based images in laparoscopic surgery.

3. Remote Manipulation in Endoscopic Surgery (RMES)

In endoscopic surgery, the surgeon manipulates tools outside of the body to manipulate target tissue by the tools' end-effectors inside the body, based on viewing the internal milieu with special endoscopic cameras. This difficult task may be viewed as a form of remote manipulation. Like working at the standard human-computer interaction (HCI) workstation, vision is indirect because display space is dissociated from hand space (look at a monitor, and manipulate elsewhere). As well, in forms of minimally invasive surgery, manipulation is indirect, i.e., the surgeon is manipulating tools outside the body, the end effectors manipulate target tissues inside the body.

The Remote Manipulation in Endoscopic Surgery (RMES) project examined surgeon's use of viewing and manipulating technologies in laparoscopic surgery, both in clinical and experimental settings. We performed task analysis, hierarchical decomposition of surgical procedures, and technology assessments in laparoscopic training workshops and operating theatres of local hospitals in B.C. Other papers in this volume report the hierarchical decomposition of selected laparoscopic procedures, and associated gaze patterns of surgeons. [9, 10]. In parallel with our clinical observations, we conducted controlled experiments at SFU with surgeons and students performing pseudosurgical tasks, using the OPTOTRAK to measure 3-D kinematics of tool motion, during different phases of the task. We focus briefly on experiments here; we have examined learning, task complexity, different tools and viewing conditions.

In one experiment, we compared direct vision with 2-D and 3-D imaging systems to perform the same tasks with an endoscopic grasper. Our hypothesis was that surgeons would perform best with direct vision of the surgical site, and better with 3-D than 2-D endoscopic imaging. Using the Telepresence 3DC™ system (International Telepresence Corp.), we had skilled laparoscopic surgeons perform ten trials on each of two tasks (in counterbalanced order) with an endoscopic grasper in a surgical space within an Endotrainer™ box: an aiming task, and a grasping task. For aiming, the surgeon touched the tip of five, metal pins placed in star formation (the pins were 2 mm. in diameter and different heights from 5 - 30 mm.); for grasping, the surgeon grasped, pulled out and dropped the five pins. Surgeons viewed a) the surgical space directly (DV), or b) the endoscopic camera's view (at 90 degrees) of the surgical space on a video monitor using either the 2-D option or c) 3-D option of the 3DC™ system. We measured timing in the task and port-centred 3-D kinematics from infrared markers placed on the tool.

Results only partially supported our hypothesis. As expected, task completion times for surgeons' performance were shorter with Direct Vision (DV) than with the 3-D imaging and 2-D imaging for all tasks. Total time for the pointing task showed that performance with DV was better when compared with the 3-D and 2-D pointing trials. However, 3-D and 2-D task times were not different. In contrast, for the grasping task, performance with DV was better than with 3-D which was reliably faster than with the 2-D imaging system. Looking at the grasping task only, the 'time on pin' and 'time to pin' measures also showed the same difference between direct viewing and endoscopic viewing. However, the advantage of 3-D over 2-D was only apparent in the 'time on pin' measure, suggesting that depth perception was beneficial for better manipulation of the pins by the end effector of the tool in the task space. When the performance time measures were divided by grasping phases using 3-D kinematics and contact data [after human prehension, 2], the two phases (grasp object phase and manipulate object phase) which comprised the 'time on pin' showed differences between 3-D and 2-D imaging; the advantage of 3-D was more pronounced in the manipulate object phase.

Thus the advantages of 3-D over 2-D imaging using the Telepresence 3DC™ system were apparent only in the more difficult grasping task; it appears that the additional depth information provided selective advantage to the compliant motion or manipulation phases, when the end effector of the tool is in contact with target objects, not during the approach or free motion phases. However, viewing the task space on a monitor is never as effective as direct vision.

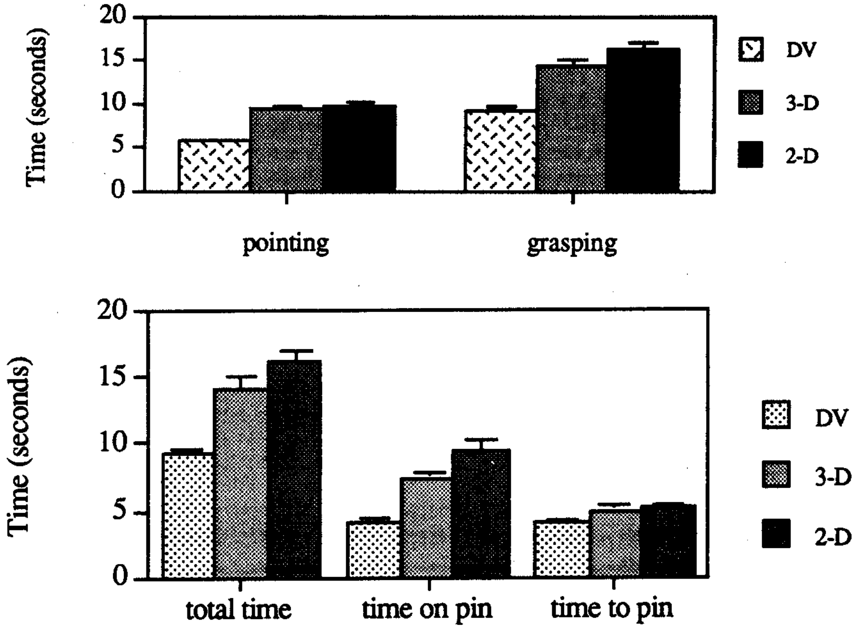

Figure 3. RMES experiment comparing direct vision, 3-D and 2-D viewing systems. Above: Task completion times (seconds) for the pointing and grasping tasks. Below: Grasping task only. Time to pins and Time on pins, showing greatest differences between the viewing conditions on manipulation phases during "time on" measure. Data are based on n = 6 surgeons, performing 10 trials in each condition.

4. Bridging VHL and RMES: Toward Intelligent Tools for Surgery

With current funding from Canada's IRIS3 project, "Intelligent Tools for Healthcare", research is bringing together these two parallel projects (VHL and RMES), for applications in planning and real-time execution of surgical procedures, for surgical training, and design of user-centred, intelligent tools in endoscopic surgery.

From HCI experimental results, we predict that superposition of display space on the workspace for the surgeon's hands would give rise to performance advantages. We have been investigating this possibility recently with the ViewSite™ (Karl Storz Endoskopy), using the same experimental tasks, time and 3-D kinematic analyses. The ViewSite™ is a first generation display system which takes the endoscopic video image and superimposes this on a sterile, disposable plate that can be placed close to the workspace for the surgeons' hands in the operating room. This would avoid the need for frequent changes in direction of gaze [10]. Recently, Hanna et al [11] manipulated location of image display and concluded that endoscopic task performnce improves when the image display is placed in front of the operator, at a level below the head and close to the hands. We are also planning to place our Telepresence video monitor in the VHL setup shown in Figure 1, to directly superimpose the display over the workspace of the hands, comparing 2-D and 3-D imaging conditions.

In stark contrast to the simple computer graphics generated on-line, in real-time for human interaction in the Virtual Hand Laboratory, are the rich, complex, real-time camera-based images captured in Endoscopic Surgery. Having studied goal-directed human manipulation and remote manipulation in these two juxtaposed domains for the last several years, the chasm between camera-based image processing and computer-generated image rendering is forefront in our challenges. Our plans are to move towards research involving composite images or overlays of camera-based and computer-generated images. This could include teaching/training [e.g., 13] or, alternatively, the more demanding real-time usage of composite images during execution of surgical procedures (computer-integrated and image-guided surgery). We plan to design and build an Enhanced Virtual Hand Laboratory, with graphics, auditory and haptics displays for real time interaction.

We wish to collaborate with others to extend the capabilities of the Virtual Hand Laboratory to training surgical skills in a variety of specialties using minimally invasive procedures. We are currently surveying across surgical specialties to understand the unique constraints, and the hierarchical decomposition of procedures in each specialty, as well as attempt two and five year technology lookaheads from practising specialty surgeons.

Our research had as its foundation basic research in human motor control, with focus on underlying processes for manipulation skills, including perceptual-motor integration through visuomotor and haptic channels. Medical robotics, surgical assist systems and haptic tools are also of interest but beyond the scope of this paper. Medical robotics requires perceptual robotics. Given the human knowledge, expertise and skills of specialized surgeons, we need to think very carefully about allocation of function in the operating rooms of the future as we develop intelligent tools...what do surgeons do best, and what tasks are best assigned to evolving technologies?

Future research will examine techniques for improving endoscopic surgery, evolution for redefining surgical tasks and redesigning intelligent tools, and investigate ways in which virtual and augmented environments can be used to train surgeons and other health care professionals.

Acknowledgements

The Virtual Hand Lab project was supported by Strategic and Research grants from the Natural Sciences and Engineering Research Council of Canada (NSERC). The Remote Manipulation in Endoscopic Surgery Project was supported by a Health Services institutional grant from the British Columbia Health Research Foundation (BCHRF). Currently this research is funded by NSERC and the Institute for Robotics and Intelligent Systems (IRIS3), a National Centre of Excellence in Canada. All research was approved by the Research Ethics Board at Simon Fraser University. We thank Kelly Booth, John Dill, Jennifer Ibbotson, Regan Mandryk, Andrea Mason, Shahram Payandeh, Yanqing Wang and Siobhann Williamson for ideas and contributions to bridging the two research areas. We thank our industrial partner, Northern Digital Inc., as well as International Telepresence Corp., and Karl Storz Endoskopy for their cooperation and support.

References

[1] Napier, J.R. (1980) *Hands*. New York: Pantheon Books.

[2] MacKenzie, C.L. and Iberall, T. (1994) *The Grasping Hand*. Advances in Psychology, Volume 104. Amsterdam: North-Holland (Elsevier Science) (500 pages).

[3] MacKenzie, C.L. and Marteniuk, R.G. (1985) Motor skill: Feedback, knowledge and structural issues. *Canadian Journal of Psychology, 39(2)*, 313-327.

[4] Graham, E.D. and MacKenzie, C.L. (1995) Pointing on a computer display. *CHI Companion '95*, 314-316.

[5] Graham, E.D. and MacKenzie, C.L. (1996) Physical versus virtual pointing. *CHI '96: Human Factors in Computing Systems*, 292-299.

[6] Graham, E.D. (1996) Pointing on a computer display. Ph.D. dissertation. Simon Fraser University, Burnaby, Canada.

[7] Wang, Y., MacKenzie, C.L. and Summers, V. (1997) Object manipulation in virtual environments: Human bias, consistency and individual differences. Short papers of *Human Factors in Computing Systems, CHI 97*, 349-350.

[8] Wang, Y., MacKenzie, C.L., Summers, V. and Booth, K.S. (1998). The structure of object transportation and orientation in human-computer interaction. *Human Factors in Computing Systems. Proceedings CHI 98*, 18-23 April, 1998, 312-319.

[9] Cao, C.G.L, MacKenzie, C.L., Ibbotson, J.A., Turner, L.J., Blair, N.P. and Nagy, A.G. (1999). Hierarchical decomposition of laparoscopic procedures. In *Proceedings of Medicine Meets Virtual Reality (MMVR:7)*. Amsterdam: IOS Press, this volume.

[10] Ibbotson, J.A. MacKenzie, C.L., Cao, C.G.L., Lomax, A.J. (1999) Gaze patterns in laparoscopic surgery. In *Proceedings of Medicine Meets Virtual Reality (MMVR:7)*. Amsterdam: IOS Press, this volume

[11] Hanna G.B., Shimi, S.M., and Cushieri, A. (1998) Task performance in endoscopic surgery is influenced by location of the image display. *Annals of Surgery*, 227(4): 481-484.

[12] Taylor, R.H., Lavalee,S., Burdea, G.C. and Mosges, R. (Eds.) (1996) *Computer-integrated Surgery: Technology and Clinical Applications*. Cambridge, Mass: MIT Press.

[13] Tendick, F., Downes, M., Cavusoglu, M.C., Gantert W. and Way, L.W. (1998) Development of virtual environments for training skills and reducing errors in laparoscopic surgery. SPIE, *Proceedings of Surgical-Assist Systems, 3262*, 36 - 44.

Author Contact:

Christine L. MacKenzie, Ph.D., Simon Fraser University, School of Kinesiology, Burnaby, B.C., CANADA V5A 1S6. Phone: +1 604 291-3004 Fax : +1 604 291-3040 Email: christine_mackenzie@sfu.ca
http://www.fas.sfu.ca/css/members/mackenzie.html