

Annotating with Light for Remote Guidance

Doug
Palmer¹

Matt
Adcock¹

Jocelyn
Smith^{1,2}

Matthew
Hutchins¹

Chris
Gunn¹

Duncan
Stevenson^{1,2}

Ken
Taylor¹

¹ CSIRO ICT Centre
GPO Box 664
Canberra ACT Australia
+61 2 6216 7000

Firstname.Lastname@csiro.au

² Department of Computer Science
Australian National University
Canberra, Act Australia
+61 2 6125 5111

Firstname.Lastname@anu.edu.au
Jocelyn.Smith@alumni.cs.ubc.ca

ABSTRACT

This paper presents a system that will support a remote guidance collaboration, in which a local expert guides a remotely located assistant to perform physical, three-dimensional tasks. The system supports this remote guidance by allowing the expert to annotate, point at and draw upon objects in the remote location using a pen and tablet-based interface to control a laser projection device. The specific design criteria for this system are drawn from a tele-health scenario involving remote medical examination of patients and the paper presents the software architecture and implementation details of the associated hardware. In particular, the algorithm for aligning the representation of the laser projection over the video display of the remote scene is described. Early evaluations by medical specialists are presented, the usability of the system in laboratory experiments is discussed and ideas for future developments are outlined.

Categories and Subject Descriptors

H5.2 [Information interfaces and presentation]: *User Interfaces. - Graphical user interfaces.*

General Terms

Design, Human Factors

Keywords

Laser, pointing, annotation, remote, network

1. INTRODUCTION

Effective direction of someone located remotely is an area of ongoing research [2,10,15]. Simple video and associated audio provide a flat, two-dimensional environment, suitable for one-to-one discussion. When more than one person is present in such an environment, the natural cues people use to indicate attention, such as gaze, become distorted by the medium [4].

Similarly, the natural mechanisms people use for direction: deixis (pointing), guidance and demonstration can become either damaged or absent. Remote deixis and annotation (in the sense of placing illustrative information on an object) in a physical space has applications to any environment where an individual is attempting to guide a remote collaborator [8].

Our approach to providing remote guidance information is to use a laser projector to project both deixis and more complex annotations directly into the physical space of the remote collaboration partner. The laser projector that we are using allows complex use of the laser because it can display vector graphics data rather than just a single spot. This gives the person controlling the laser projector (the local collaborator) a wide visual vocabulary while retaining the focus of interaction by directly annotating the subject of discussion. Figure 1 shows examples of the pointing and annotation forms possible with the projector implementation.



Figure 1 Clockwise from top left: Pointing with cursor, Graphic Annotation on forehead plus freehand sketch and Sketch Annotation

The remote pointing and annotation are controlled by the guide with a two-dimensional interface containing a live video view of the remote workspace on a tablet display. By touching and drawing over the displayed video with an electronic pen the guide can direct the laser projector to display pre-defined graphics or free-form line sketching. A video camera is mounted with the laser projector so that the most direct view of the remote

OzCHI 2007, 28-30 November 2007, Adelaide, Australia. Copyright the author(s) and CHISIG. Additional copies are available at the ACM Digital Library (<http://portal.acm.org/dl.cfm>) or can be ordered from CHISIG(secretary@chisig.org)

OzCHI 2007 Proceedings, ISBN 978-1-59593-872-5

workspace is provided to the guide, though the software architecture will support video streams from any video camera which shows the relevant part of the remote scene.

Since the video source and projector do not share a common position, the resulting annotations need to be aligned. Rather than attempt to provide accurate and dynamic modelling of the target space, we provide an approximate alignment and make use of the video feedback to allow the guide to correct any errors.

This paper discusses the technical features of the remote projection system. Although we include some discussion of the overall user experience, a more comprehensive examination of the user experience is the subject of another paper [22].

1.1 Related Work

A simple approach to remote deixis is to allow the remote collaborator to use a laser spot, an idea that dates back to Schlossberg's patent of 1981[19]. This approach has been used in the GestureLaser[26] system, which uses a mobile laser pointer mounted on a car or person, and the WACL [18] system, which uses a laser mounted co-axially with a shoulder-worn camera. As an alternative to an actual pointer, NETMAN uses an augmented reality headset to provide pointing [9]. Experiments with GestureLaser and NETMAN have shown that remote pointing can improve performance in completing guided physical tasks [2,14].

In addition to simple pointing, some efforts have also included annotation: projecting an icon or diagram over a physical object. Pinhanez has used a steerable projector for turning a surface into a projection screen [16]. Butz and Schmidt use a projector to annotate real-world objects with labels [3]. Miyashita uses an LCD projector, with a local remote control to control a pointer [12]. Raskar *et al* use projectors to locally enhance 3D models with shading and other visual effects [17].

Shared drawing facilities, such as shared whiteboards, allow complex, remote annotation [20]. The DOVE system allows drawing over video with a number of the features, such as fade rates that we have implemented [6]. However, such non-collocated annotations can interrupt the flow of direction by forcing those local to the subject under discussion to remove their attention from that subject.

Laser lightshow systems, used in entertainment, can also be used to project annotations. Iwasaki and Okuyama have used the vector drawing facilities of a laser to annotate printed circuit boards [7]. Despite using two mirrors, in the same configuration as a laser projector, GestureLaser only uses single laser spot, although the idea of using the laser to draw more complex objects was mentioned in one paper [11]. The GestureLaser pointer is light-weight and designed to operate on a car or person, with the result that more complex drawings are difficult to achieve.

1.2 Design Criteria

The laser drawing system described in this paper is a sub-system of a tele-health system which is intended to support pre- and post-surgical consultations between the surgeon and a remotely located patient. In this tele-health scenario the patient is accompanied by a clinic assistant (a trained health worker) who is guided by the surgeon to facilitate the consultation, including conducting any necessary physical examination of the patient. Our laser drawing system allows the surgeon to point, draw and annotate directly

onto the patient, thus providing very direct real-time guidance for the clinic assistant.

This scenario creates a set of design criteria which were used to develop the laser drawing system. These criteria are drawn from observations of actual outpatient clinics at the Royal Children's Hospital (Melbourne, Australia), from design discussions with the surgeons themselves and from independent progress reviews by two medical specialists (a cranio-facial surgeon and an intensive-care specialist). These criteria reflect the complex nature of paediatric surgical consultations, in which the patient examination is embedded in a wider dialogue with the patient, the patient's parents, the clinic assistant and one or more clinicians.

These design criteria include:

- Simple deixis and the ability to annotate the patient in various ways.
 - The ability to **point**, preferably with different cursor shapes so that the consultant can indicate the subject of discussion. ("Put the instrument here.")
 - The ability to **leave icons** on the patient to mark points as a complex intensive care procedure is outlined. ("Insert a tube here, then a needle here, then ...")
 - The ability to **sketch areas** on the patient to indicate a region of interest. ("There should be a plate in the head, roughly here and this shape.")
- The drawing should be capable of covering the head and torso of the patient.
- The general focus is on conversational support, rather than precise and accurate instruction. The surgeon can be expected to adjust drawing positions via video feedback. The annotations are not expected to remain in position on the patient if the patient moves. (In the case of intensive care, where more complex instruction sequences are usual, the patient is usually static.)
- The system should not be too noisy or otherwise distracting.
- The option to view the patient from additional directions should be available and, if necessary, to control the laser system from one of these additional camera views.
- Both the surgeon and clinic assistant should be able to see the pointer located on the patient with minimal effort, under ordinary clinical lighting conditions similar to internal office lighting.
- Laser safety regulations require that the laser have a power of less than 1mW so that the blink reflex will protect the eye.
- The entire system is intended for networked clinical use in a number of locations and needs to be portable and non-intrusive. As a result, parts of the system that might otherwise be mounted on the walls or the ceiling need to stand alone.

2. SYSTEM OVERVIEW

The laser projection system consists of two major parts: the hardware that provides user interface and projection facilities and a software system that provides coordination between the various system components and hardware control. The following descriptions are couched in the tele-health scenario outlined above, where a surgeon is controlling the laser display system to

guide a clinic assistant in the examination of a child patient who is sitting in an adjustable office-style chair.

2.1 Hardware

The layout of the projector system is shown in Figure 2. A small laser projector is mounted on a pole above the patient. An overview pan-tilt-zoom video camera is mounted on the top of the projector, with a video field of view similar to the projector's display region. A second pan-tilt-zoom camera is mounted on a desk level with the patient's face, giving the surgeon a normal horizontal view of the patient. The surgeon uses a pen and tablet display to control the cameras and laser projector. A second pen and tablet display is located in the patient's room, mounted on a moveable arm. This display allows the clinic assistant to see the same view that the surgeon sees and to draw supplementary on-video annotation.

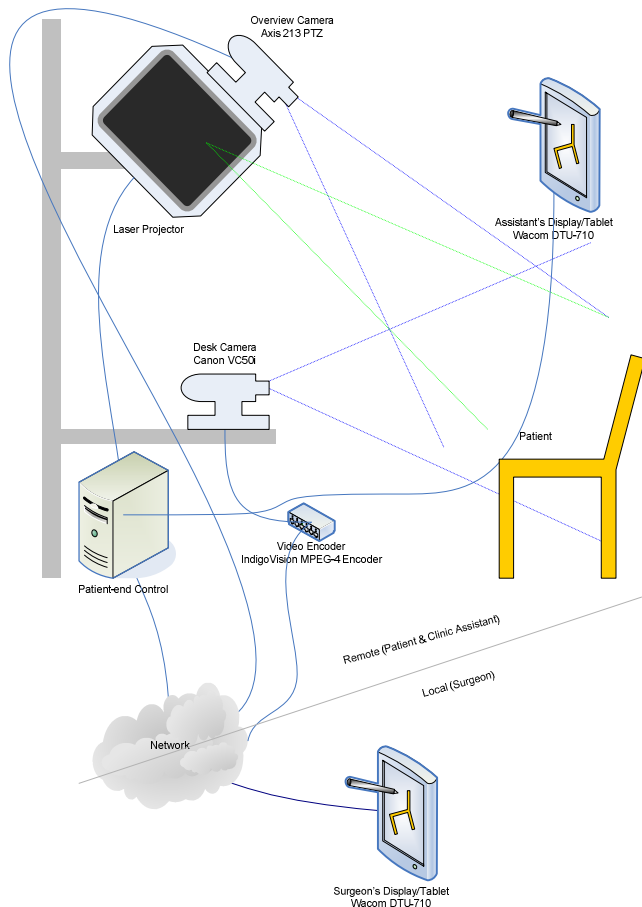


Figure 2 Laser Projector Layout

The laser projector was custom-built by Genius Laser Technology Pty Ltd. Since the projector needs to be portable, it is mounted on a pole rather than on the ceiling. Commercial laser projectors, which are generally designed for entertainment work and need to be correspondingly robust, were judged to be too heavy for pole mounting. Genius constructed a smaller, lighter version with a 1 milliwatt (1 mW) laser. Figure 3 shows the mounted projector and overview camera.

The first iteration was a single unit with an internal power supply and was found to be too large and heavy for pole mounting. In

addition, the galvanometers in the laser projector generated heat while in operation, as did the internal power supply, requiring several noisy fans. Genius then redesigned the unit. By making the power supply external, the heat and weight of the unit could be reduced to the point that a single, quieter fan could be used and the unit could be pole mounted.



Figure 3 Laser Projector and Overview Camera: left image is front view, right image is side view

The laser projector is driven by a Pangolin QM2000 projector controller PCI board. This board hosts a dedicated processor that drives the laser via a standard ILDA (International Laser Display Association) connection. The QM2000 comes with an API that allows application software to define a complete vector graphics drawing that the laser projector then displays.

The overview camera is an Axis 213 PTZ network camera. This camera was chosen for its large picture size of 768x576 pixels, pan-tilt mounting, zoom, simple HTTP interface and ability to produce both Motion JPEG and MPEG-4 video streams.

The desk camera is a Canon VC-C50i. This camera is primarily used by the larger tele-health system for face-to-face communication between surgeon and assistant, and was chosen for its video output, pan-tilt mounting and zoom. A serial connection allows control of the pan-tilt-zoom functions.

The display tablet is a Wacom DTU-710 combined tablet and LCD monitor. At 17", this display is large enough to provide a clear picture from the cameras. Pen-based input allows the surgeon to perform natural sketching and pointing actions while discussing the patient's condition.

2.2 Software

The control software for the laser projection system is part of the software architecture developed for the larger tele-health system. The architecture is agent-based and constructed from software components. Multi-agent architectures have been successfully used to manage other mixed remote/local environments, such as the Smart Classroom [21]. Such architectures allow loosely coupled components to cooperate in a peer-to-peer manner over multiple platforms.

The architecture uses the Spread toolkit messaging system [23] to allow components to communicate via messages. The messaging model is loosely based on the FIPA agent communication standards [5]. XML messages, based on XML-schema descriptions [25], provide a language-neutral form of communication. Language-specific containers manage components and mediate communications. The architecture currently supports Java and C++ components.

A simplified diagram of the software components and communications is shown in Figure 4.

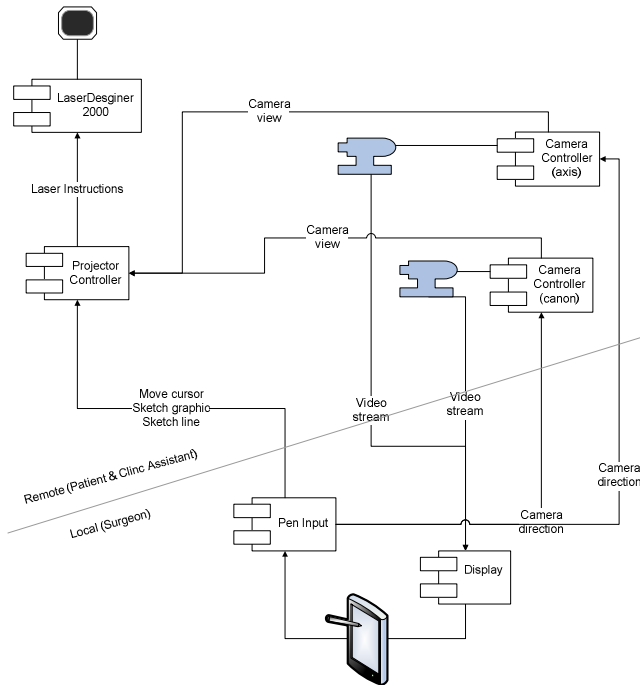


Figure 4 Software Components

The display and pen input is used by the surgeon who can select video streams from any of the cameras in the patient's room and display them on the tablet. The surgeon can then control the camera's pan-tilt-zoom facilities by sending requests to the camera controllers. The camera controllers are responsible for interpreting the direction requests and converting them into camera-specific instructions. The camera controllers then issue view descriptions for the current view of the camera.

The surgeon can also use the pen input to send drawing requests to the projector controller. The projector controller is responsible for interpreting the drawing requests and converting them into a composite drawing. Since the surgeon's perspective is that of the camera, requests coming from the pen input are from the point of view of a camera and need to be converted to the point of view of the projector before they are displayed. For this reason, the projector monitors changes in camera view.

The surgeon has access to several sorts of request to the projector: graphic management and enquiry, fade rates, floating cursor moves, static graphics and free line drawings.

2.3 Location in the larger tele-health system

The laser-projection subsystem is embedded in the larger tele-health system, which provides the other components needed to deal with the full complexity of the scenarios in question. These components include multiple static streams of video and audio in both directions to general situational and activity awareness, and conversations between multiple people. Separate hand-guided video cameras provide close-up views of the patient, including some experimental work on 3D image and video capture and display. Further experimental work includes a wearable tactile hand-guidance system.

This paper will concentrate on the laser projection sub-system and will not describe the other work in further detail.

The tele-health system provides a shared sketching space so that the surgeon and clinic assistant can both directly sketch on a shared still picture or video stream, using the Wacom pen and tablet displays. This space is similar to the DOVE system [6].

2.4 Coordinate Systems

The pointer system uses two basic coordinate systems: an absolute coordinate system based on vectors from a common origin and a relative coordinate system based on the current view of a camera or projector, treating the view as a flat screen.

A transformation from relative to absolute coordinates provides a number of architectural benefits. The key benefit is that the display and input tablet does not need to be concerned with the geometry of the system. Pen coordinates can be simply expressed as relative coordinates in a particular camera field of view. The projector controller and camera controllers are then responsible for converting the relative coordinates into actual directions and pan-tilt settings.

2.5 Camera-Projector Alignment

In an ideal configuration, the camera and laser projector would be perfectly co-located and relative camera coordinates would translate directly into projector directions. However, the need to provide a clear line of sight for both the projector and camera and the need for multiple cameras means that co-location is impossible. Some mechanism for ensuring alignment is needed.

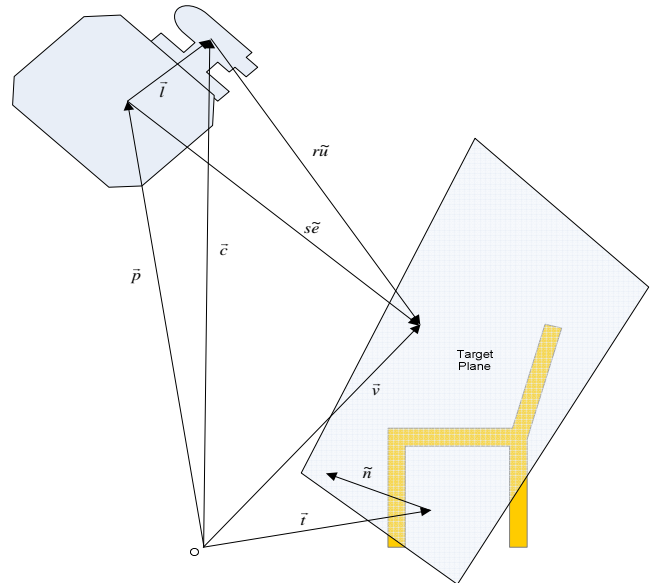


Figure 5 Camera-Projector Alignment

Figure 5 shows the current model for alignment. The current model, rather than attempting an exact model of the object to be drawn on, uses a simple target plane, specified by the position, normal pair of (\vec{t}, \vec{n}) . The desired target plane is currently specified in the system configuration. In Figure 5, O represents a common origin for position vectors. The camera, at \vec{c} , has a ray drawn in the direction of \vec{u} to the point where it intersects the

target plane at $\vec{v} = \vec{c} + r\vec{u}$ Since \vec{v} is on the target plane, $(\vec{v} - \vec{t}) \cdot \vec{n} = 0$ and, therefore,

$$r = \frac{(\vec{t} - \vec{c}) \cdot \vec{n}}{\vec{u} \cdot \vec{n}}$$

The direction \vec{e} from the projector, located at \vec{p} , is then given by $s\vec{e} = \vec{l} + r\vec{u}$ where $\vec{l} = \vec{c} - \vec{p}$ [13].

2.5.1 Error Analysis

If the actual display plane for the laser beam is not the configured target plane, then the output will appear to be shifted in the camera view. It is possible to model the effect of these errors by a shift in the target plane.

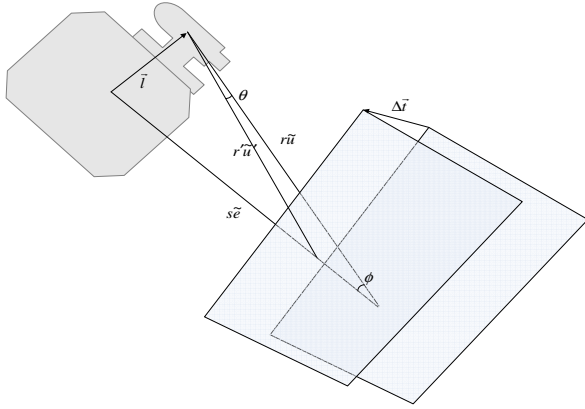


Figure 6 Target Plane Error

The effect of a shift is outlined in Figure 6. A shift of $\Delta\vec{t}$ in the target plane produces an apparent change in distance and direction from the camera to $r'\vec{u}'$ at an angle of θ from the original \vec{u} .

The sine and cosine rules give $\sin \theta / (\Delta\vec{t} \cdot \vec{e}) = \sin \phi / r'$, $\cos \phi = \vec{u} \cdot \vec{e}$ and

$$r'^2 = r^2 + (\Delta\vec{t} \cdot \vec{e})^2 - 2r(\Delta\vec{t} \cdot \vec{e})(\vec{u} \cdot \vec{e})$$

Setting $\rho = (\Delta\vec{t} \cdot \vec{e})/r$ gives an equation for θ of

$$\sin^2 \theta = \frac{\rho^2(1 - (\vec{u} \cdot \vec{e})^2)}{1 + \rho^2 - 2\rho(\vec{u} \cdot \vec{e})}$$

If $s \approx r$ then $(\vec{u} \cdot \vec{e}) \approx 1 - l^2/2r^2$.

For the overhead camera, l is 12cm and r is approximately 1m. A 10cm change in target plane will produce a change in angle of less than 1 degree – less than 2cm at 1m range. For the desk camera, l and r are approximately 1m and a 10cm change in the target plane will produce a change in angle of approximately 5 degrees.

2.6 Drawing Modes

The three modes of input discussed in Section 1.2 are reflected in three drawing modes: cursor, graphic and sketch. These modes are illustrated in Figure 1.

The cursor and graphic modes allow arbitrary pre-built graphics to be added to the laser drawing. The main difference between

cursors and graphics is that cursors are mobile and graphics are static. The sketch mode builds simple line segments. Graphics from all modes can be mixed arbitrarily. A fade function allows elements to fade and eventually disappear over a specified time, in keeping with the ephemeral, discussion-based model of interaction. Different elements can fade at different rates, with the fade rate specified by the user when the element is added. Fussell et al. discuss appropriate default fade rates [6].

Cursors and graphics are represented by scalable vector graphic (SVG) documents [24]. These documents can be transmitted across a network, allowing a consultant to add graphics to the library of shapes that can be drawn. The projector controller maintains the graphic library. SVG documents can contain the preferred dimensions of the output. Since the target plane gives a distance estimate from the projector to the target, it is possible to scale the SVG output to approximately reproduce the preferred size of the projected graphic.

Sketches are simply represented by polyline segments. New points are added incrementally, with segment-end messages indicating when the consultant wishes to start a new line.

The projector controller accumulates the graphics and then converts the collection into a single list of points for the QM2000 card to process. The QM2000 card accepts lists of point positions, each with associated corner flags and line colours. To display a new line segment a black (invisible) line is drawn between the two segments.

The corner flags cause the laser projector to dwell at points for a small time. Without corner flags, the laser projector tends to cut corners when moving from point to point, since the galvanometers' inertia leads to inaccurate paths. With no corner flag set, the laser projector produces curves that can be used to approximate arc line segments.

The SVG graphics are converted to outline shapes by the Batik SVG toolkit [1]. These shapes are then converted into points for the QM2000 card. Straight lines are easily converted into two points with corner flags. Quadratic and cubic line segments are represented by two and three line segments respectively, with no corner flags.

Laser projectors are physical systems. The blanking shutter and galvanometers have inertia and design limits that need to be managed. The QM2000 card can be set with a number of management parameters: the number of points to trace per second, the amount of time to dwell at corner points and the lead time for the blanking shutter. All of these parameters can be set by the Pangolin LD2000 API and by the card management software.

The number of points per second and the number of points being displayed determines the number of times that the projector will scan over a part of the drawing every second. If too many points are displayed, then the drawing will start to visibly fade and flicker; this effect is even more pronounced when viewing the drawing through a camera, since stroboscopic effects come into play. The maximum number of points displayed per second is set by the limits of the galvanometers in the laser projector. The number of points is set to 20,000; the limit recommended by Genius. This value corresponds to approximately 100-200 points in the drawing, with the value dependant on the individual viewer's susceptibility to flicker and the camera frame rates.

The amount of time to dwell at corner points determines how sharp the corners are on straight line segments. Too little time and the corners of a box become rounded. Too long and the number of scans becomes too low. The value for the custom laser projector was determined to be 3 points by trial and error. Such a value produces a recognisable straight-edged rectangle with sharp corners for a rectangular input to the QM2000 card; lower values produce observably rounded corners and higher values produce no recognisable improvement.

The lead time for the blanking shutter determines how accurate the start and end points of line segments are. If the shutter opens too early, the line is drawn before the galvanometers are in position and ends before complete. If too late, then the line is incomplete at the start and overshoots. The value for the custom laser projector was set at 4 points by trial and error. Such a value produces an almost closed circle for a circular input to the QM2000; higher or lower values increase the gap in the circle.

2.7 Input

Input is managed by the Wacom display and a pen input device. The pen is used to directly position the cursor on a display of camera output. This gives a direct interface to the pointer, with GUI button mode controls allowing cursor, graphic or sketch modes. Button mode controls also allow the pen to be used for camera control and on-display sketches. The Wacom display can generate up to 100 mouse movement events a second, giving smooth motion on input. Figure 7 shows the display and pen. Our software currently limits the input rate to 40 events per second to avoid swamping the system.



Figure 7 Display and Tablet Input

The pen input also controls the camera, with the user able to indicate a region of interest in the video output and have the camera move and zoom to show that area.

The optimisation of the GUI layout is an ongoing, iterative process. Even though the pen input offers a number of options (buttons, pen-touch against the display and an “eraser” input on the reverse of the pen) the number of possible actions is quite large. We have designed a number of different user interface configurations and the relative merits of each will be the subject of a subsequent user study.

3. EXPERIENCE OF PROTOTYPE

During construction, the spatial reach and operational qualities of the laser system were explored and performance tests, which varied round-trip latency and number of drawing actions per second, were conducted. These tests raised a number of issues.

3.1 Laser Projector

The laser projector was built to specifications at the start of the project and experience has demonstrated a number of areas where the design could be improved. The 40 degree angle of view of the current laser projector is too small to reliably cover the area needed for patient interaction. The cooling fan is noisy and has the potential to cause a slight distraction. The projector is also vulnerable to overheating and blown fuses, with the laser overheating if a complex drawing is shown for an extended period of time (over an hour) and fuses blowing about once every two months. The scan rate may also be too slow (see Section 3.3). A new projector, with a wider angle of view and faster galvanometers has been built and is currently being tested.

Correctly mounting and aligning the laser projector and camera has proved to be quite difficult. Currently, the exact position and alignment of the various components is treated as a known input. A procedure for deriving the positioning and alignment information from the running system, based on a known calibration pattern, would improve both set-up time and accuracy.

Our design allows any light projecting device to be used in place of a laser. As a cheaper alternative to a laser projector, a standard DLP data projector was also tried. The projector used was not able to project enough light to be reliably visible under ordinary fluorescent light, making it hard for the clinic assistant to initially locate the cursor, even though various guide techniques were used to allow the aide to find the pointer. The cursor was even more difficult to see through the camera. The projector was also distractingly noisy and generated a great deal of heat.

If a sufficiently bright, quiet and cool data projector that was light enough to be pole mounted and of a comparable price to the laser projector could be obtained, then the use of a data projector has certain advantages. The laser projector that we are using is monochrome, although multi-colour laser projectors exist, whereas an ordinary projector is inherently multi-colour. The strokes drawn by the laser are simple lines; with a data projector, the stroke information in the SVG graphics could be used - although the laser projector has the advantage that the laser beam is relatively narrow and less subject to distance effects. The laser needs to re-trace the drawing to ensure a sufficient level of brightness, limiting the complexity of the drawing; a data projector is only limited by its resolution.

3.2 Round-Trip Latency

There is a round-trip delay of approximately 250ms between the surgeon performing some action and the video of that action returning to the display. There are a number of components contributing to this delay: network and middleware latency, processing of requests by the projector controller and the time needed to assemble a video frame into Motion JPEG or MPEG-4. This delay is enough to make the movement of the cursor “as if the cursor was on the end of a rubber band,” according to one user.

Our informal experiments with artificial delays show that the immediate feedback loop becomes unusable with a round trip delay of approximately 500ms. If the round-trip delay rises during actual clinical trials over extended wide area networks, then compensatory techniques will need to be developed.

3.3 User Experience

The inaccuracies of the target plane model and alignment usually mean that the laser cursor appears at an offset to the pen position on the tablet. The cursor offset adds to the difficulty of positioning the cursor, although the pilot users seemed able to adapt within a few minutes. As a side-effect, the offset can prevent the user's hand from obscuring the cursor in the video.

The offset is typically 2-4cm and is not consistent across the tablet surface. These offsets correspond to errors of 3-6 degrees for the 17" Wacom screen and the camera angle of view of 45 degrees. This offset is larger than the errors calculated in Section 2.5.1. A likely explanation is that the extra offset is caused by the difficulties in positioning and aligning the projector and cameras discussed above, since the alignment can be improved by adjusting the zero position of the camera by manually rotating the pan-tilt mechanism. The level of offset will be revisited once a more effective alignment procedure has been devised.

Initial trials with the laser in an office setting, with overhead fluorescent light, suggests that the laser output is generally bright enough for the aide. By the time that output has been passed back through the video stream, the cursor is sometimes difficult to see on certain clothing — light pink in particular. Since 1mW output is a hard limit imposed by safety concerns, we cannot make the raw output of the laser any brighter. However, the higher scan rate of the newer projector provides a brighter image.

The laser projection system was deployed in a user-interface study which explored use of direct annotation in the workspace and sketching over a video display of the workspace when both are available. The laser projection system performed sufficiently well for the study participants, who were able to guide their remotely located study partners with it. Among the outcomes of the study was the finding that, depending upon the role adopted, direct laser projection is more or less convenient than the non-collocated, shared sketching facility. From the clinic assistant's point of view, the ability to look directly at the patient, rather than having to switch between a screen and the patient, has obvious advantages. From the surgeon's point of view, direct sketching is easier and more accurate than controlling the laser. The mix of annotation methods used by subjects during laboratory experiments, and the user interfaces that support that mix, are still areas of active research [22].

3.4 Cameras

Both the Axis and Canon cameras showed non-linear zoom behaviour. The maximum zoom levels also fall far short of the advertised specifications but are still adequate. Calibration curves were constructed for both cameras by comparing camera settings to the visible portion of a measure in the camera's video output. A consistent calibration curve takes about a day to construct.

Both cameras have an auto-focus that can be queried. We investigated using the auto-focus information as a source of information for positioning a dynamic target plane. However, the results were not encouraging. The cameras tend to hunt for focus, leading to noisy results. In our development environment, a

typical office, we have found the noise level to be greater than the change in focus between near and distant objects.

3.5 Software Interface

The designed interface between the projector controller and the QM2000 card currently requires that a complete drawing is passed to the card every time a part of the drawing changes. This unsophisticated interface was built to simplify handling the complete drawing, since the complete drawing can be quite complex with various sketch elements and SVG graphics, along with a mobile cursor, being positioned and then fading at different rates.

The cursor position or sketch lines can update at up to 40 times a second, triggering a complete redraw each time. For complex drawings, the updates become queued and the cursor or sketch begins to significantly lag behind; it was possible to get lag times of 5 seconds or more by continuously adding to the sketch. To prevent lag, the length of the sketch is currently limited to 64 elements. This is acceptable for quick annotations — the sketch in Figure 1 contains 12 elements — but will become a limitation if a more complex task is being performed. At present, we are not using the full capabilities of the QM2000 card and re-engineering of the drawing interface should improve performance.

4. CONCLUSIONS AND FURTHER WORK

We have presented a design and implementation of a laser-based system for remotely annotating the objects in a scene. This annotation, in the form of a moving cursor, a fixed graphic or a live line sketch, can enable guidance of a person located in the remote scene. This system was presented in the context of a surgeon guiding a remotely located clinic assistant in the examination of a patient. This context has provided the design criteria which the system was required to meet.

The software architecture for controlling the laser-based system which uses an open-source toolkit to message-based communication between the components was presented. This software allows the laser-based system to be controlled in real time through pen-based interactions over video display from any video camera in the remote site.

A method of aligning the laser-based system and the selected video camera, using a target-plane approach was presented, together with an error analysis for cases where the actual target plane has shifted away from the intended target plane.

Early user experience of this system shows that it performs fast enough, and with small enough latency, that users can effectively guide a person in the remote location to perform physical three-dimensional tasks in the designed-for workspace. However, limitations on the number of vector segments displayed and on the alignment calibration between the laser projector and the matching video camera will limit complex annotations.

There are two areas of future work that we would like to undertake. Firstly, we would like to develop a method for rapidly calibrating the pen-based user interface. The calibration referred to here deals with the offsets and non-linearities inherent in controlling a 3D pointer system via a 2D interface (as shown in Figure 7). This would allow for re-configuration of the actual layout of the room in which the laser-based system is deployed.

Secondly, we would like to extend the present system from its relatively static target and ephemeral laser annotation to include

situations where the target may move and the laser annotation will follow it, remaining in place on the target. An application in our tele-examination scenario would be to allow the clinic assistant to move the patient in order to better view the area under discussion, and have the laser-projected guidance marker remain on the exact place of interest.

5. ACKNOWLEDGEMENTS

The authors would like to thank Dr Wellington Davis and Dr John Lambert for their valuable advice and insight. Thanks also go to Arthur Ipsaros of Genius Laser Technology for his time and effort in designing and constructing the laser projector. This work was supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and by the CSIRO ICT Centre.

6. REFERENCES

- [1] The Apache XML Graphics Project. *Batik SVG Toolkit*. <http://xmlgraphics.apache.org/batik/> 2007
- [2] Bauer, M., Kortuem, G., and Segall, Z., "Where Are You Pointing At?" A Study of Remote Collaboration in a Wearable Videoconference System. In *3rd IEEE International Symposium on Wearable Computers (ISWC '99)*, (1999), pp. 151.
- [3] Butz, A. and Schmitz, C., Annotating Real World objects using a Steerable Projector-Camera Unit. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, (2005), pp. 109.
- [4] Daly-Jones, O., Monk, A., and Watts, L., Some Advantages of Video Conferencing Over High-Quality audio Conferencing: Fluency and Awareness of Attentional Focus *International Journal of Human-Computer Studies*, 49, (1998) pp. 21-58.
- [5] Foundation for Intelligent Physical Agents, "ACL Message Structure Specification," FIPA, SC00061G, Dec 2002.
- [6] Fussell, S. R., Setlock, L. D., Yang, J., Kramer, A. D. I., Ou, J., and Mauer, E., Gestures over Video Streams to Support Remote Collaboration on Physical Tasks *Human-Computer Interaction*, 19, (2004) pp. 273-309.
- [7] Iwasaki, E. and Okuyama, T., inventors. Laser drawing apparatus with precision scaling-correction. US. Patent no. 5980088, Issued Nov 9, 1999.
- [8] Kirk, D. S., Fraser, D., and Stanton, The Effects of Remote Gesturing on Distance Instruction. In *Conference on Computer Support for Collaborative Learning (CSCL'05)*, (2005), pp. 301-310.
- [9] Kortuem, G., Bauer, M., and Segall, Z., NETMAN: the design of a collaborative wearable computer system *Mobile Network Applications*, 4, 1, (1999) pp. 49-58.
- [10] Kraut, R. E., Fussell, S. R., and Siegel, J., Visual Information as a Conversational Resource in Collaborative Physical Tasks *Human-Computer Interaction*, 18, (2003) pp. 13-49.
- [11] Kuzuoka, H., Ishimo, G., Nishimura, Y., Suzuki, R., and Kondo, K., Can the GestureCam be a Surrogate? In *Fourth European Conference on Computer Supported Cooperative Work (ECSCW'95)*, (1995, Stockholm, Sweden), pp. 179-194.
- [12] Miyashita, K., inventor. Image projection system and a method of controlling a projected pointer. US. Patent no. 7059724, Issued Jun 13, 2006.
- [13] M. Mortenson. *Computer Graphics Handbook*, Industrial Press, 1990.
- [14] Ohta, S., Kuzuoka, H., Noda, M., Sasaki, H., Mishima, S., Fujikawa, T., and Yukioka, T., Remote support for emergency medicine using a remote-control laser pointer. *Journal of Telemedicine and Telecare*, 12, (2006) pp. 44-48.
- [15] Oviatt, S., Ten myths of multimodal interaction *Communications of the ACM*, 42, 11, (1999) pp. 74-81.
- [16] Pinhanez, C., Using a Steerable Projector and a Camera to Transform Surfaces into Interactive Displays . In *CHI '01 Extended Abstracts on Human Factors in Computing Systems*, (2001), pp. 369-370.
- [17] Raskar, R., Ziegler, R., and Willwacher, T., Cartoon Dioramas in Motion. In *2nd International Symposium on Non-Photorealistic Animation and Rendering (NPAR'02)*, (2002), pp. 7-13.
- [18] Sakata, N., Kurata, T., Kato, T., Kourogi, M., and Kuzuoka, H., WACL: Supporting Telecommunications Using - Wearable Active Camera with Laser Pointer. In *Seventh IEEE International Symposium on Wearable Computers*, (2003), pp. 53-56.
- [19] Schlossberg, H. A., inventor. Remote pointing system . US. Patent no. 4280135, Issued Jul 21, 1981.
- [20] Shae, Z.-Y., Tseng, B., and Leung, W. H., Immersive Whiteboard Collaborative System *Annals of Software Engineering*, 12, 1, (2001) pp. 193-212.
- [21] Shi, Y., Xie, W., Xu, G., Shi, R., Chen, E., Mao, Y., and Liu, F., The Smart Classroom: Merging Technologies for Seamless Tele-education *Pervasive Computing*, 2, 2, (2003) pp. 47-55.
- [22] Stevenson, D., Li, J., Smith, J. and Hutchins, M., A Collaborative Guidance Case Study. *To appear in Proc 9th Australasian User Interface Conference (AUIC2008), Wollongong, Australia (2008)*
- [23] Spread Concepts LLC. *The Spread Toolkit*. <http://www.spread.org/> 2006
- [24] SVG Working Group, "Scalable Vector Graphics (SVG) 1.1 Specification," World Wide Web Consortium, REC-SVG11-20030114, Jan 2003.
- [25] XML Schema Working Group, "XML Schema 1.0 Specification," World Wide Web Consortium, REC-xmlschema-1-20041028, Oct 2004.
- [26] Yamazaki, K., Yamazaki, A., Kuzuoka, H., Oyama, S., Kato, H., Suzuki, H., and Miki, H., GestureLaser and GestureLaser Car: development of an embodied space to support remote instruction. In *Sixth European Conference on Computer Supported Cooperative Work*, (1999), pp. 239-258.