

## Chapter 6

# Head-Mounted System Hardware Installation

Several types of eye trackers are available, ranging from scleral coils, EOG, to video-based corneal reflection eye trackers. Although each has its benefits and drawbacks (e.g., accuracy vs. sampling rate), for graphical or interactive applications the video-based corneal reflection tracker is arguably the most practical device. These devices work by capturing video images of the eye (illuminated by an infra-red light source), processing the video frames (at video frame rates), and outputting the eye's  $x$ - and  $y$ -coordinates relative to the screen being viewed. The  $x$ - and  $y$ -coordinates are typically either stored by the eye tracker itself, or can be sent to the graphics host via serial cable. The advantage of the video-based eye tracker over other devices is that it is relatively noninvasive, fairly accurate (to about  $1^\circ$  visual angle over a  $30^\circ$  viewing range), and, for the most part, not difficult to integrate with the graphics system. The video-based tracker's chief limitation is its sampling frequency, typically limited by the video frame rate, 60 Hz. Hence, one can usually expect to receive eye movement samples at least every 16 ms (typically a greater latency should be expected because the eye tracker needs time to process each video frame, and the graphics host needs time to update its display).

### 6.1 Integration Issues and Requirements

Integration of the eye tracker into a graphics system depends chiefly on proper delivery of the graphics video stream to the eye tracker and the subsequent reception of the eye tracker's calculated 2D eye coordinates. In the following description of system setup, a complete graphics system is described featuring two eye trackers: one a table-mounted, monocular eye tracker set underneath a standard television display, and the other a binocular eye tracker fitted inside a Head-Mounted Display (HMD). Both displays are powered by the same graphics host. Such a laboratory has been set up within Clemson University's Virtual Reality Laboratory, and is shown



(a) HMD-mounted binocular eye tracker (panorama photostitch image)



(b) Table-mounted monocular eye tracker

**Fig. 6.1** Virtual Reality Eye Tracking (VRET) lab at Clemson University

in Fig. 6.1. In Fig. 6.1b, the portion of the lab is shown where, on the right, the TV display is installed with the monocular table-mounted eye tracker positioned below. In Fig. 6.1a, the HMD helmet is resting atop a small table in front of three d.c. electromagnetic tracking units. The eye tracker unit (a PC) is situated between the d.c. tracking units and the dual-head graphics display monitors to the right in the image. The PC's display is a small flat panel display just left of the dual graphics display monitors. Both HMD and TV (and graphics) displays are driven by the graphics host, which is out of view of the image (on the floor beneath the desk in front of the visible chair). The four small TV monitors atop the eye tracking PC in

Fig. 6.1a display the left and right scene images (what the user sees) and the left and right eye images (what the eye tracker sees).

The following system integration description is based on the particular hardware devices installed at Clemson's Virtual Reality Eye Tracking (VRET) laboratory, described here for reference. The primary rendering engine is a Dell Dual 1.8 GHz Pentium 4 (Xeon) PC (1 GB RAM) running Fedora Core 4 and equipped with an nVidia GeForce4 FX5950U graphics card. The eye tracker is from ISCAN and the system includes both binocular cameras mounted within a Virtual Research V8 (high resolution) HMD as well as a monocular camera mounted on a remote pan-tilt unit. Both sets of optics function identically, capturing video images of the eye and sending the video signal to the eye tracking PC for processing. The pan-tilt unit coupled with the remote table-mounted camera/light assembly is used to noninvasively track the eye in real-time as the subject's head moves. This allows limited head movement, but a chin rest is usually used to restrict the position of the subject's head during experimentation to improve accuracy. The V8 HMD offers  $640 \times 480$  resolution per eye with separate left and right eye feeds. HMD position and orientation tracking is provided by an Ascension 6 Degree-Of-Freedom (6 DOF) Flock Of Birds (FOB), a d.c. electromagnetic system with a 10 ms latency per sensor. A 6 DOF tracked, handheld mouse provides the user with directional motion control. The HMD is equipped with headphones for audio localization.

Although the following integration and installation guidelines are based on the equipment available at the Clemson VRET lab, the instructions should apply to practically any video-based corneal reflection eye tracker. Of primary importance to proper system integration are the following,

1. Knowledge of the video format the eye tracker requires as input (e.g., NTSC or VGA)
2. Knowledge of the data format the eye tracker generates as its output

The first point is crucial to providing the proper image to the user as well as to the eye tracker. The eye tracker requires input of the scene signal so that it can overlay the calculated point of regard for display on the scene monitor which is viewed by the operator (experimenter). The second requirement is needed for proper interpretation of the point of regard by the host system. This is usually provided by the eye tracker manufacturer. The host system will need to be furnished with a device driver to read and interpret this information.

Secondary requirements for proper integration are the following.

1. The capability of the eye tracker to provide fine-grained cursor control over its calibration stimulus (a crosshair or other symbol)
2. The capability of the eye tracker to transmit its operating mode to the host along with the eye  $x$ - and  $y$ -point of regard information.

These two points are both related to proper alignment of the graphics display with the eye tracking scene image and calibration point. Essentially, both eye tracker capabilities are required to properly map between the graphics and eye tracking viewport coordinates. This alignment is carried out in two stages. First, the operator

can use the eye tracker's own calibration stimulus to read out the extents of the eye tracker's displays. The eye tracker resolution may be specified over a  $512 \times 512$  range, however, in practice it may be difficult to generate a graphics window that will exactly match the dimensions of the video display. Differences on the order of one or two pixels will ruin proper alignment. Therefore, it is good practice to first display a blank graphics window, then use the eye tracker cursor to measure the extents of the window in the eye tracker's reference frame. Because this measurement should be as precise as possible, fine cursor control is needed. Second, the eye tracker operates in three primary modes: reset (inactive), calibration, and run. The graphics program must synchronize with these modes, so that a proper display can be generated in each mode:

- *Reset*: display nothing (black screen) or a single calibration point.
- *Calibration*: display a simple small stimulus for calibration (e.g., a small dot or circle) at the position of the eye tracker's calibration stimulus.
- *Run*: display the stimulus required over which eye movements are to be recorded.

It is of the utmost importance that proper alignment is achieved between the eye tracker's calibration stimulus points and those of the graphics' display. Without this alignment, the data returned by the eye tracker will be meaningless.

Proper alignment of the graphics and eye tracking reference frames is achieved through a simple linear mapping between eye tracking and graphics window coordinates. This is discussed in Chap. 7, following notes on eye tracking system installation and wiring.

## 6.2 System Installation

The two primary installation considerations are wiring of the video and serial line cables between the graphics host and eye tracking systems. The connection of the serial cable is comparatively simple. Generation of the software driver for data interpretation is also straightforward, and is usually facilitated by the vendor's description of the data format and transmission rate. In contrast, what initially may pose the most problems is the connection of the video signal. It is imperative that the graphics host can generate a video signal in the format expected by both the graphics display (e.g., television set or HMD unit) and the eye tracker.

In the simplest case, if the host computer is capable of generating a video signal that is suitable for both the stimulus display and eye tracker, all that is then needed is a video splitter which will feed into both the stimulus display and eye tracker. For example, assume that the stimulus display is driven by an NTSC signal (e.g., a television), and the host computer is capable of generating a display signal in this format. (This was possible at the Clemson VRET lab because the SGI host can send a copy of whatever is on the graphics display via proper use of the `ircombine` command.) If the eye tracker can also be driven by an NTSC signal, then the installation is straightforward. If, however, the stimulus display is driven by VGA video (e.g., an

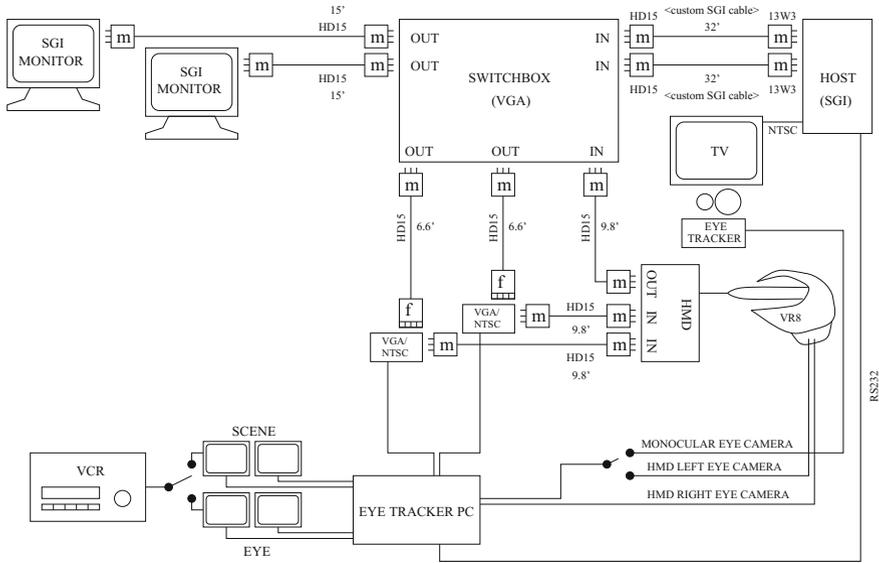


Fig. 6.2 Video signal wiring of the VRET lab at Clemson University

HMD), but the eye tracker is driven by NTSC video, then the matter is somewhat more complicated. Consider the wiring diagram given in Fig. 6.2. This schematic shows the dual display components, the HMD and TV, used for binocular eye tracking in a virtual reality environment, and monocular eye tracking over a 2D image or video display. The diagram features the necessary wiring of both left and right video channels from the host to the HMD and eye tracker, and a copy of the left video channel sent to the TV through the host’s NTSC video output.

The HMD is driven by a horizontal-synch (h-sync) VGA signal. A switchbox is used (seen in Fig. 6.1a just above the eye tracker keyboard) to switch the VGA display between the dual-head graphics monitors and the HMD. The HMD video control box diverts a copy of the left video channel through an active pass-through splitter back through the switchbox to the left graphics display monitor. The switchbox effectively “steals” the signal meant for the graphics displays and sends it to the HMD. The left switch on the switchbox has two settings: monitor or HMD. The right switch on the switchbox has three settings: monitor, HMD, or both. If the right switch is set to monitor, no signal is sent to the HMD, effectively providing a biocular display in the HMD (instead of a binocular, or stereoscopic display). If the right switch is set to HMD, the graphics display blanks out because the HMD does not provide an active pass-through of the right video channel. If the right switch is set to both, the right video channel is simply split between the HMD and the monitor, resulting in a binocular display in both the HMD and on the monitors. This last setting provides no amplification of the signal and hence both the right LCD in the HMD and the right graphics monitor displays appear dim. This is mostly used for testing purposes.

The entire video circuit between the graphics host, the switchbox, and the HMD is VGA video. The eye tracker, however, operates on NTSC. This is the reason for the two VGA/NTSC converters that are inserted into the video path. These converters output an NTSC signal to the eye tracker and also provide active pass-throughs for the VGA signal so that, when in operation, the VGA signal appears undisturbed. The eye tracker then processes the video signals of the scene and outputs the signal to the scene monitors, with its own overlaid signal containing the calculated point of regard (represented by a crosshair cursor). These two small displays show the operator what is in the user's field of view as well as what she or he is looking at.

The eye images, in general, do not pose a complication since this video signal is exclusively processed by the eye tracker. In the case of the eye tracker at Clemson's VRET lab, both the binocular HMD eye cameras and the table-mounted monocular camera are NTSC and the signals feed directly into the eye tracker.

### 6.3 Lessons Learned from the Installation at Clemson

The eye tracker at Clemson contains hardware to process dual left and right eye and scene images. It can be switched to operate in monocular mode, for which it requires just the left eye and scene images. In this case, a simple video switch is used to switch the signal between the eye image generated by the left camera in the HMD and the camera on the table-mounted unit.

The first installation at Clemson used display monitors from Silicon Graphics, Inc. (SGI) that were driven by either VGA video, or the default R, G, B video delivered by 13W3 video cables. To drive the HMD, VGA video was required, connected by HD15 cables. To connect the video devices properly, special 13W3-HD15 cables were needed. Although this seems trivial, custom-built cables were required. These cables were not cheap, and took a day or two to build and deliver. If timing and finances are a consideration, planning of the system down to the proper cabling is a must! (Today this is less of a concern although properly matching cabling is still an issue, this time not so much concerning the video signal as concerning the Keyboard Video Monitor or KVM switch; see Chap. 9.)

A problem that was particularly difficult to troubleshoot during the Clemson installation was the diagnosis of the format of the VGA signal emitted by the SGI host computer. Initially, before the eye tracker was installed, the HMD was tested for proper display. The output of the switchbox was used directly to drive the HMD. Everything functioned properly. However, after inserting the HMD into the video circuit, the eye tracker would not work. It was later found that the problem lay in the VGA/NTSC converters: these converters expect the more common VGA signal which uses a timing signal synchronized to the horizontal video field (h-sync signal; the horizontal and vertical sync signals are found on pins 13 and 14 of the VGA HD15 cable). The SGI host computer by default emits a sync-on-green VGA signal leaving pins 13 and 14 devoid of a timing signal. The VR HMD contains circuitry that will read either h-sync or sync-on-green VGA video and so functions quietly given either

signal. The fault, as it turned out, was in an improper initial wiring of the switchbox. The switchbox was initially missing connections for pins 13 and 14 because these are not needed by a sync-on-green VGA signal. Once this was realized, the entire cable installation had to be disassembled and the switchbox had to be rewired. With the switchbox rewired, the custom 13W3 cables had to be inspected to confirm that these components carried signals over pins 13 and 14. Finally, a new display configuration had to be created (using SGI's `ircombine` command) to drive the entire circuit with the horizontal sync signal instead of the default sync-on-green. The moral here is: be sure of the video format required by all components, down to the specific signals transmitted on individual cable pins!

## 6.4 Summary and Further Reading

This chapter presented key points from an eye tracker's installation and its integration into a primarily computer graphics system. Although perhaps difficult to generalize from this particular experience (a case study if you will), nevertheless there are two points that are considered key to successful installation and usage of the device:

- Signal routing
- Synchronization.

Signal routing refers to proper signal input and output, namely concerning video feeds (input to the eye tracker) and (back then rs232) serial data capture and interpretation (output from the eye tracker). Synchronization refers to the proper coordinate mapping of the eye tracker's reference frame to the application responsible for generating the visual stimulus that will be seen by the observer. Proper mapping will ensure correct registration of both eye tracker calibration and subsequent data analysis. The actual data mapping is carried out in software (see the next chapter), however, the hardware component that facilitates proper software mapping is the eye tracker's capability of measuring the application's display extents in its own reference frame. This is usually provided by the eye tracker if it allows manual positioning and readout of its calibration cursor.

There are two primary sources where further information can be obtained on system installation and setup. First, of course, is the manufacturer's manual that will typically be included with the eye tracking device. Second, the best resource for installation and usage of the equipment is from the users themselves. Users of eye trackers typically report, in a somewhat formal way, what they use and how they use it in most technical papers on eye tracking research. These can be found in various journal articles, such as *Vision Research*, *Behavior Research Methods, Instruments, and Computers (BRMIC)*, and conference proceedings. There are various conferences that deal with eye tracking, either directly or indirectly. For example, conferences that deal with computer graphics (e.g., SIGGRAPH, EuroGraphics, or Graphics Interface), human-computer interaction (e.g., SIGCHI), or virtual reality (e.g., VRST),

may include papers that discuss the use of eye trackers, but their apparatus description may be somewhat indirect insofar as the objective of the report typically does not deal with the eye tracker itself, rather it concentrates more on the results obtained through its application. Examples of such applications are given in Part III of this text. At this time, there are two main conferences that deal more directly with eye tracking: the European Conference on Eye Movements (ECEM), and the U.S.-based Eye Tracking Research and Applications (ETRA). Finally, the eye movement email listservs *eye-movements* and *eyemov-l* are excellent on-line “gathering places” of eye tracker researchers. Here, questions on eye trackers may be answered directly (and usually promptly) by users of similar equipment or even from other vendors.