Towards an open-hardware open-software toolkit for robust low-cost eye tracking in HCI applications

David Winfield[12], Dongheng Li[2], Jason Babcock[2], Derrick J. Parkhurst[12]
The Department of Psychology[1]
The Human Computer Interaction Program[2]
Iowa State University

April 18th, 2005

Iowa State University Human Computer Interaction Technical Report ISU-HCI-2005-04
Communicated by Derrick Parkhurst.

Abstract

Eye tracking has long held the promise of being a useful methodology for human-computer interaction ranging from explicit control of computer interfaces to adaptive interfaces based on the user's attentional state. A number of barriers have stood in the way of the integration of eye tracking into everyday applications, including the intrusiveness, robustness, availability, and price of eye-tracking systems. To lower these barriers, we have developed the openEyes toolkit. The toolkit consists of an open-hardware design for a digital eye tracker that can be built from low-cost off-the-shelf components and a set of open-source software tools for digital image capture, manipulation, and analysis in eye-tracking applications. We expect that the availability of this toolkit will facilitate the development of eye-tracking applications and the eventual integration of eye tracking into the next generation of everyday human computer interfaces. We discuss the methods and technical challenges of low-cost eye tracking as well as the design decisions that produced our current system. Finally, we discuss the benefits of an open-hardware and open-software approach as well as ways we can facilitate the integration of eye-tracking techniques into the next generation of human computer interfaces.

1. Introduction

Eye tracking has been primarily used in research systems to investigate the visual behavior of individuals performing a variety of tasks (for review see [2]). However, for some time now, research has also been underway to examine the use of eye movements in human computer interfaces [6]. Only to some very small degree has eye-tracking research been integrated into consumer products. This lack of a more widespread integration of eye tracking into consumer-grade human computer interfaces can be attributed to the significant intrusiveness, lack of robustness, low availability, and high price of eye-tracking technology. It is important that these obstacles be overcome because HCI research clearly indicates the potential of eye tracking to enhance the quality of everyday human-computer interfaces. For example, eye tracking can allow users to control a computer though the use of eye movements [6]. Eye typing is one such application. Users with movement disabilities look at keys on a virtual keyboard in order to interact with a computer that they would otherwise be unable to use manually [8]. Moreover, icon selection in everyday graphical interfaces could potentially be speeded with eye tracking given that when users intend to select an icon, they typically locate it using eye movements beforehand [13]. Eye movements can also be used to adapt an interface to the user's needs. For example in video transmission and virtual reality applications, gaze-contingent variable-resolution display techniques actively track the viewer's eyes and present a high level of detail at the point of gaze while sacrificing level of detail in the periphery [10, 11].

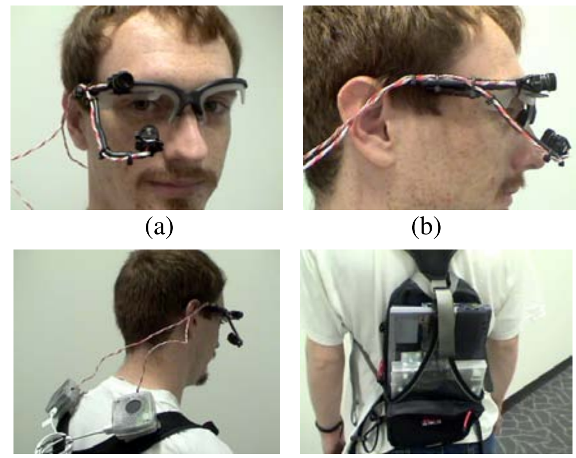


Figure 1: Eye tracker: Generation 1. (a&b) CCDs and lens mounts are mounted on a to a pair of modified safety glasses. (c) The CCDs are separated from the processing board via wires such that the processing boards can be located on a back pack. (d) A laptop modified to fit within a small backpack. The laptop captures and processes video from the cameras.

Although various eye-tracking technologies have been available for many years (for review, see [15]), these techniques have all been limited in a number of important ways. The primary limitation, especially for application in consumer products, is the invasiveness of eye-tracking systems. Some techniques require equipment such as special contact lenses, electrodes, chin rests, bite bars or other components that must be physically attached to the user. These invasive techniques can quickly become tiresome or uncomfortable for the user. Video-based techniques have minimized this invasiveness to some degree. Video-based techniques capture an image of the eye from a camera either mounted on head gear worn by the user or mounted remotely. The recent miniaturization of video equipment has greatly minimized the intrusiveness of head-mounted video-based eye-trackers [12, 1]. Furthermore, remotely located video-based eye-tracking systems can almost be completely unobtrusive (e.g., see [5, 9]), although at some cost to the robustness and quality of the eye tracking.

The cost and availability of eye-tracking technology has also limited its application. Until only recently, eye trackers were custom made upon demand by a very few select production houses. Even today, eye-tracking systems from these sources range in price from 5,000 to 40,000 US dollars, and thus limit their application to high-end specialty products. It is important to note however that the bulk of this cost is not due to hardware, as the price of high-quality camera technology has dropped precipitously over the last ten years. Rather the costs are mostly associated with custom software implementations, sometimes integrated with specialized, although inexpensive, digital processors, to obtain high-speed performance. Moreover, customer support can also contribute significantly to these final purchase prices.

It is clear that to reap the potential benefits of eye tracking in everyday human-computer interfaces, the development of inexpensive and robust eye-tracking systems will be necessary. Towards this goal, we have undertaken the development of an eye tracker that can be built from low-cost off-the-shelf components. We have iterated through a number of system designs and in this paper we describe these systems as well our successes and failures in this process. Although our system is ever evolving, we have arrived at a minimally invasive, digital head-mounted eye tracker capable of an accuracy of approximately one degree of visual angle. Aside from a desktop or laptop computer to processes video, the system costs approximately 350 US dollars to construct. As part of the openEyes toolkit, a step by step tutorial is provided that details the construction of our open-hardware design. In this paper, we also describe our eye tracking software. We make this software freely available in the form of an open-source package as part of the openEyes toolkit. We hope that the availability of software, ease of construction and open design of our eye-tracking system will enable interface designers to begin exploring the potential benefits of eye tracking for human computer interfaces. Furthermore, the flexibility provided by our open approach should allow system designers to integrate eye tracking directly into their system or product, an option not typically feasible with equipment purchased from production houses that use proprietary methods and software. We expect that the availability of the openEyes toolkit will significantly enhance the potential that eye tracking will be incorporated into the next generation of human-computer interfaces.

2. Video-based eye tracking

Two types of imaging processes are commonly used in eye tracking, visible and infrared spectrum imaging [4]. Visible spectrum imaging is a passive approach that captures ambient light reflected from the eye. In these images, it is often the case that the best feature to track in visible spectrum images is the contour between the iris and the sclera known as the limbus. The three most relevant features of the eye are the pupil - the aperture that lets light into the eye, the iris - the colored muscle group that controls the diameter of the pupil, and the sclera, the white protective tissue that covers the remainder of the eye. Visible spectrum eye tracking is complicated by the fact that uncontrolled ambient light is used as the source, which can contain multiple specular and diffuse components. Infrared imaging eliminates uncontrolled specular reflection by actively illuminating the eye with a uniform and controlled infrared light not perceivable by the user. A further benefit of infrared imaging is that the pupil, rather than the limbus, is the strongest feature contour in the image; both the sclera and the iris strongly reflect infrared light while only the sclera strongly reflects visible light. Tracking the pupil contour is preferable given that the pupil contour is smaller and more sharply defined than the limbus. Furthermore, due to its size, the pupil is less likely to be occluded by the eye lids. The primary disadvantage of infrared imaging techniques is that they cannot be used outdoors during the day time due to the infrared component of sun light.

Infrared eye tracking typically utilizes bright-pupil or dark-pupil techniques (however, see [9] for the combined use of both bright and dark pupil techniques). Bright-pupil techniques illuminate the eye with a source that is on or very near the axis of the camera. The result of such illumination is that the pupil is clearly demarcated as a bright region due to the photoreflective nature of the back of the eye. Dark-pupil techniques illuminate the eye with an off-axis source such that the pupil is the darkest region in the image, while the sclera, iris and eye lids all reflect relatively more illumination. In either method, the first-surface specular reflection of the illumination source off of the cornea (the outer-most optical element of the eye) is also visible. This vector between the pupil center and the corneal reflection is typically used as the dependent measure rather than the pupil center alone. This is because the vector difference limits sensitivity to head movements.

The innovative work of Jeff Pelz and colleagues [12, 1] at the Rochester Institute of Technology on the construction of low-cost minimally invasive head-mounted eye trackers is particularly noteworthy. In their system, analog cameras are mounted onto safety glasses (in a similar configuration as that shown in Figure 1) and video of the user's eye and the user's field of view are interleaved in a single interlaced video frame and recorded using a mini-DV camcorder stowed in a backpack. Point of gaze computation is then performed off-line using proprietary hardware and software purchased from a production house. Given our goal to integrate eye movement measurements into human computer interfaces, this depedence on high-cost proprietary equipment is a serious limitation of their approach. Furthermore, the off-line nature of the system is another limitation as some degree of real-time performance will be necessary in many HCI applications. However, their innovation in head gear design and low-cost approach is laudable and we adopt both in our own efforts.

3. openEyes toolkit

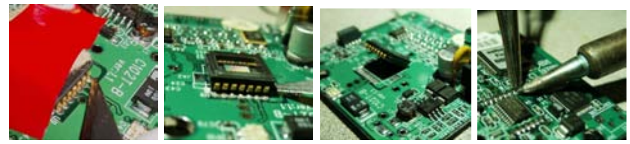
The motivation for the creation of the openEyes toolkit stems from the recognition in the eye-tracking and human computer interaction communities of a need for robust inexpensive methods for eye tracking. The openEyes toolkit addresses this need by providing both an open-hardware design and a set of open-source software tools to support eye tracking. These components are described below. The openEyes toolkit is distributed under the GNU General Public License (GPL) as published by the Free Software Foundation (http://www.fsf.org) and can be downloaded from the openEyes toolkit website (http://hcvl.hci.iastate.edu/openEyes).

4. Open-hardware design

In this section, we begin by describing the basic hardware design principles behind our system. Four generations of the head gear are described. For each generation, we discuss the design improvements as well as the system limitations. Rather than describing in detail the construction of each head gear, we limit our description to the most significant construction details. We provide a much more extensive description of the construction on the openEyes toolkit website. This includes a step by step tutorial on head-gear construction as well as a detailed parts list accompanied by hyperlinks to vendor web sites.

The first design consideration after having chosen to use a head-mounted system was the configuration of the head gear. The most significant issue was where to mount the cameras. Given that until recently cameras were quite large, a number of other head-mounted systems have the cameras placed either above the eyes, on top of the head or above the ears, primarily for ergonomic reasons. This placement necessitates that a mirror or prism be inserted somewhere in the optical path to the eye. Instead of taking this approach, we adopt the solution developed at RIT of placing the eye camera on a boom arm such that there is a direct line of sight between the camera and the eye (see Figure 1). The primary advantage of this design is that it avoids the need for expensive optical components. Half-silvered infrared-reflecting mirrors or prisms can be very expensive and glass components can pose significant danger of eye damage in near-eye applications. Furthermore, we were unable to locate an inexpensive source of half-silvered infrared-reflecting mirrors constructed of plexiglass. Such mirrors are typically used by production houses and must be purchased in bulk. The primary disadvantage of this design is that a portion of the visual field is blocked by the camera. Given the small extent and peripheral positioning of the camera/boom, we view this as an acceptable compromise. In fact, because these components are attached to the head gear and thus static in the user's visual field, they are easily ignored in a similar way that the frames of normal eye glasses are not overly distracting.

The second design consideration concerned finding a way to capture and process digital images for real-time application. The RIT system used inexpensive low-resolution CMOS cameras to generate analog video output. The cameras they use are among the smallest available on the market and, in general, analog cameras are available in much



(a) (b) (c) (d)

Figure 2: How to remove the CCD without damage. (a) Tape over the sensor to prevent it from being damaged during CCD removal. (b) Cut all the way through the pins on one side of the sensor as close as possible to the camera board. The sensor is cut from the board because it is heat sensitive and would be damaged if de-soldered. (c) Score the pins on the other side of the sensor a few times with the knife. Lift the free side of the sensor up until the pins on the other side break and the sensor detaches from the board. (d) Flip board and remove remaining pins using a high-temperature soldering iron and forceps. Minimize the time that the iron is in contact with the board to avoid heat damage.

smaller packages than digital cameras. We considered a number of analog image-capture solutions to use in combination with an RIT head gear, but all such solutions were overly expensive (i.e. many hundreds of dollars), would require considerable fabrication expertise (e.g., the use of an A/D chip), or were not applicable in the mobile context (i.e. required a desktop computer). We therefore considered only solutions that utilized digital cameras with a readily available means of capture to a standard laptop computer. For example, a number of small inexpensive USB web cameras were investigated but the resolution and frame rates were limited by the bandwidth of USB. We failed to find any inexpensive USB 2.0 compatible web cameras that utilized the bandwidth of USB 2.0. Ultimately, we settled upon using inexpensive IEEE-1394 web cameras. The bandwidth of these cameras (400Mbit/sec) is sufficient to capture video from two cameras at an 8-bit resolution of 640x480 pixels with a frame rate of 30hz. Two additional benefits of IEEE-1394 cameras include the fact that cameras on the same bus will automatically synchronize themselves and that the IEEE-1394 standard is well supported under Linux with the 1394-based DC Control Library.

We examined a number of inexpensive IEEE-1394 cameras available on the market. Initially, the Apple I-sight camera was considered because of its unique construction. The optics have an auto-focus feature and the CCD is mounted on a flat flex cable approximately one inch long that leads to the main processing board. However, after much investigation, we failed to find a way to extend this cable in a reasonable way; any modifications would have required extremely difficult soldering of surface mount connectors. We finally settled on using the comparably priced Unibrain Fire-i IEEE-1394 web camera. One advantage of using this camera for our application is that more than one can be daisy chained together and share a single power source (see Figures 1c and 3d). The disadvantage with this camera is that the CCD sensor is soldered directly to processing board and without removal, the entire board would be too cumbersome to mount on a head gear. Therefore a technique was developed to detach the CCD sensor from the camera board and solder a multi-conductor cable of some length between the board and the chip. Shown and described in Figure 2 is the method that we used to remove the sensor. When done carefully, the sensor remains undamaged and the lens and mount can be re-attached so that the camera functions as before. Note however that a degree of noise is induced in the captured images. Much of our work subsequent to our initial design has been to find a way to reduce this noise (see below).

The final design consideration was to make the eye-tracking system completely mobile. While mobility was important we also wished to do all our image processing on general-purpose hardware platform. Furthermore, we wanted to be sure that we had sufficient processing power to perform all the image processing on-board in real-time. Therefore we purchased a "desktop-replacement" class Sony Laptop. The laptop was configured with a 3.1Ghz Pentium processor, 512 megs of memory, three USB 2.0 ports, a 4pin IEEE-1394 port and a 802.11 a/b/g wireless card. To reduce the weight and profile of the laptop, the LCD and keyboard were removed and the battery was relocated. The battery life for the system actively eye tracking (i.e., wireless networking active and powering an infrared illuminater) is on the order of approximately one hour. Although the modified laptop still weights a few pounds, it is easily carried in a small mesh backpack (See Figures 1d and 3b).

4.1. Generation 1

Our first generation eye tracker is shown in Figure 1 and, as can be seen, the profile is small and unobtrusive. The Sony CCD and lens mount assembly standard with the Fire-i camera were extended from the camera processing boards and mounted on a pair of modified safety glasses that had the plastic lenses cut mostly away. Very fine unshielded wire was used to extend the CCD and when routed above the ear and back to the processing boards mounted on the backpack, its presence was hardly noticeable. Moreover, the lightness of the lenses and boom arm did not add to the perceivable weight of the glasses when worn. The presence of the eye tracker was not disturbing despite the fact that the camera occluded a portion of the visual field.

The design of the first generation system had three major limitations. First, the CCDs for this system were removed using a soldering iron. Given the small size of the chip and the proximity of other components on the board, this was a procedure that we believe damaged the chips and/or board. Second, the thin unshielded wire lead to significant noise in the captured images when both cameras were operated simultaneously. The amount of noise was amplified when the 14 lines for each CCD were run adjacent to each other down to the processing boards on the backpack. Although

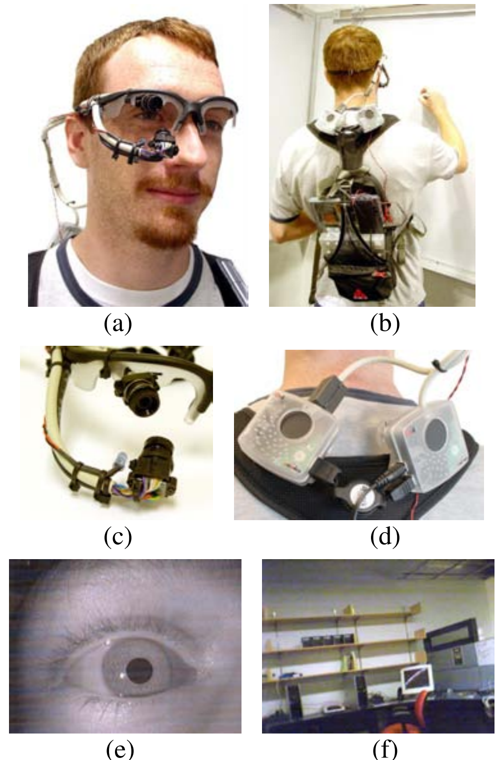


Figure 3: Eye tracker: Generation 2 (a&c) Head gear: note the shielded cable and use of the CS mount & lens. (b) Laptop computer in small backpack (d) closeup of daisy chained camera housings (e) IR-illuminated eye image (f) scene image. Note the line noise inducing a color shift.

the noise was tolerable, it was unpredictable and tended to change as the wearer shifted their head and body. The final limitation of this approach was that we employed visible spectrum imaging. Due to the low sensitivity of these consumer-grade cameras, we were often unable to image the eye with the user indoors. Furthermore, the presence of specular reflections from various ambient sources made extracting a reliable measure of eye movements particularly difficult.

4.2. Generation 2

In our second generation eye tracking system, we attempted to redress many of the limitations of the first generation system. Most significantly, we moved to an infrared imaging approach. As can be seen in Figure 3, we placed an infrared LED on the boom off axis with respect to the center of eye camera. This configuration produces a dark-pupil pattern of illumination. The LED was powered from a free USB port on the laptop. Unfortunately, this design decision also required that we switch the lens mount assembly on the eye camera. The Fire-i cameras come with a small, non-standard mount and lens combination which has an infrared cut-filter coated on the sensor side that could not be removed. To solve this problem, we salvaged the somewhat larger lens mount and lens from an OrangeMicro i-

Bot web camera. The infrared blocking filter was easily removed from this lens and replaced with an 87c Wratten filter to block visible light and pass only infrared. The image captured using infrared illumination can be seen in Figure 3e. Note that the infrared illumination strongly differentiates the pupil from the the iris in the image. Also note the presence of a specular reflection of the LED. This is an important benefit as the corneal reflection can be tracked and used to compensate for head gear slippage.

The second major modification that we made to the system was to use shielded cables between the CCD and the processing boards in order to reduce the noise. The added diameter of these cables can be seen in Figure 3c. Although we could still maintain a relatively ergonomic cable configuration, the cables extending over the ear were much more noticeable to the user than in the previous generation. Furthermore, the additional stiffness of the cables sometimes induced the head gear to shift when the user turned their head. To minimize this slippage of the head gear, we employed the use of an elastic head band specially designed for glasses (see Figure 3b). In addition to the shielded cables we also employed an electrical choke to the cable of the eye camera. While the interference noise was reduced to some degree, its presence was still noticeable and continued to depend on the positioning of the cables. In our final configuration, both cables were joined together along their length, which resulted in the least noise. Unfortunately, a second type of strong noise appeared in this system which was much more problematic although it was sporatic. For example, when the head gear was nudged, touched or the user turned their head abruptly, line noise was induced. We suspect that because we had soldered to the CCD and boards a number of times that these components were damaged. Furthermore, it is possible that weak solder joints on the chip lead to this increased sensitivity.

4.3. Generation 3

Having produced the second generation camera that was capable of infrared eye tracking (albeit with a large degree of noise which induced frequent tracking errors), we were encouraged to produce another iteration. The system that we ultimately arrived at is shown in Figure 4. The basic design is fundamentally the same, however, a number of important modifications were made. First, we focused on ways to reduce the noise. Thin but double-shielded cables were employed to this end. These cables added a significant degree of stiffness and consequentially the only reasonably ergonomic configuration of the head gear was for the scene camera to be mounted on the left side of the glasses (see Figure 4a and b). The consequence of this modification is that the parallax is increased between the tracked eye and the scene camera. This increase of parallax limits the extent of the region in which the eye tracking calibration will be valid. Further exploration will be needed to determine

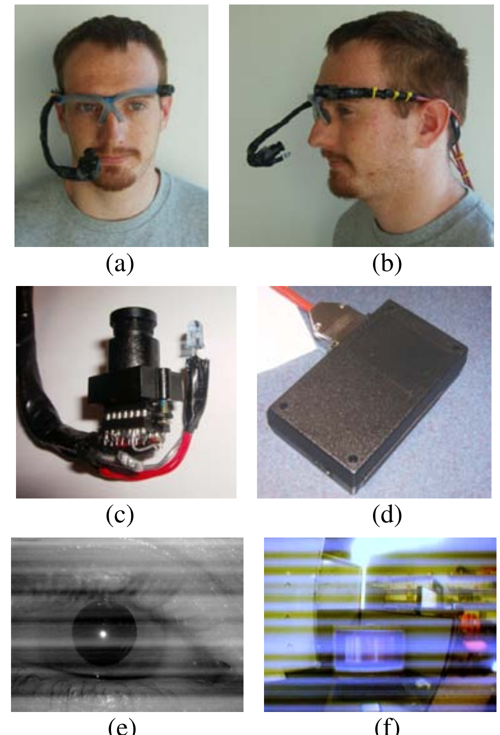


Figure 4: Eye tracker: Generation 3. (a&b) Headgear with extended boom arm, new eye camera lens and repositioned scene camera. (c) Exposed view of the eye camera showing the IC socket configuration allowing easy exchange of CCD/lens mount and the more centrally positioned LED. (d) The housing for both camera boards using a single connector leading to the scene and eye cameras. Note that the firewire cables that extend to the computer are not shown. (e) Eye image. Note the reduced depth of field in the eye image due to the zoom lens. (f) Scene image. Note the radial distortion attributable to the wide field of view lens.

the seriousness of this artifact. Another modification in the design was to switch to the Unibrain monochrome Fire-i board level camera in order to take advantage of its overall greater sensitivity to infrared light and thus allow us to use a lower level of infrared illumination. Given the significant additional price (approximately 50 US dollars), it is unclear to us whether this design decision was worth while. Our experience was that image quality (aside from the noise) was comparable to that obtainable with a standard Fire-i camera. In any case, we used a new monochrome camera for the eye camera and a color camera for the scene camera. We extracted the CCDs from the processing boards using the technique described earlier and shown in Figure 2. The final design decision was to develop an interlocking socket assembly on which to mount the CCD sensors in order to minimize any joint stress on the chip. We did this in the hopes of eliminating the sporatic noise in the second generation eye tracker. The integration of the CCD and a standard IC socket is shown in Figure 5a&b and the socket assembly attached to the boom arm is shown in Figure 4c. Together, these modifications completely eliminated the sensitivity of

(a) (b) (c) (d)

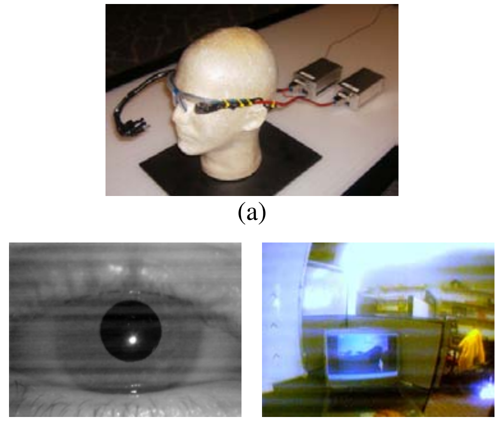
Figure 5: (a & b) CCD mounted to IC socket via short solid-conductor wires. (c) The eye camera lens (12 mm) in a CS mount is attached securely to the CCD socket via a small plastic strip screwed to the mount from behind the socket. (d) The scene camera lens (1.9 mm) in a smaller CS mount similarly attached to the CCD socket.

the camera to spurious noise during head movements or adjustments to the head gear.

In the second generation system, we used the I-bot 4.5 mm lens and as can be seen from Figure 3e, the portion of the image that was occupied by the eye was quite small. Given that the quality of eye tracking is related to the size of the eye in the image, we decided to employ a 12mm zoom lens in the third generation system. As can be seen from Figure 4e, a much closer image of the eye is obtained. While this is clearly beneficial for achieving high-accuracy eye measurements, this design decision lead to a number of consequences. First, the depth of field in the image is smaller and consequentially more attention in necessary to obtain a correct focus. Furthermore, more attention must be paid to the alignment of the camera. The restricted field of view of the system also means a greater sensitivity of head gear slippage. In the worst case, the pupil may slip out of view of the camera. The zoom lens also has a longer minimal focusing distance which required that the boom arm be extended. While a longer boom arm might be more prone to movement and vibration during vigorous activity and lead to image blur, a benefit is that the profile of the eye camera in the visual field of the user is significantly minimized.

We also improved the modularity of the system by housing both processing boxes in a single plastic case and separating the head gear from the housing using a single connector. This can be seen in Figure 4d. However, this design decision was a serious misstep. As can be seen in Figures 4e and 4f, we experienced even worse noise levels than previously. We concluded that this was due entirely to interference between the cameras because when either camera was used alone, the images were entirely noise free. There was little if any difference between the images captured before and after the CCD removal. However, when both cameras were operated simultaneously, significant line noise was induced into the captured images. Fortunately, this line noise is periodic and can be minimized through image averaging over time. In spite of the high degree of noise, we can still obtain high accuracy measures of eye movements (see Section 6).

A few further, but minor modifications worth mentioning were made to the design. First, a socket assembly was constructed for the LED as well to facilitate easy replacement. The LED was moved to a more central location to keep the corneal reflection from being off of the cornea. A scene camera with wider field of view was also used. A different type of safety glasses were also used because the previous type went out of production.



(a)

(b) (c)

Figure 6: Eye tracker: Generation 4. (a) The scene and eye camera processing boards are now housed in separate, aluminum boxes to reduce interference between the cameras. (b&c) Eye and scene images with significantly reduced line noise compared to generation 3.

4.4. Generation 4

To reduce the interference noise in our next generation eye tracker, a minor modification was needed that lead to a major improvement. We separated the camera processing boards into individually shielded metal cases and used separate shielded metal connectors. As can be seen in Figure 6 we observed a significant reduction in line noise. The primary disadvantage of this approach is the the cases must be mounted somewhat aukwardly on the mobile backpack. Furthermore, the relatively short length of the cables in combination with the aukward mounting limits the extent of head movements that the user can make. Longer and more flexible cable will be used in the future so that the cases can be mounted in the backpack and not overly restrict movement.

5. Open-source software

5.1. cvHAL: computer vision Hardware Abstraction Layer

cvHAL is a Linux-based open-source computer vision software package that is currently under development in our lab. The purpose of cvHAL is to provide an automated system for discovery, configuration, networking, and "smart camera" image processing. The software allows users to focus on computer vision algorithm development by abstracting away from hardware-specific camera issues. cvHAL is an always-on daemon that processes requests for video streams from clients on the network. While there is other similar video server software available, cvHAL is targeted at the computer-vision community by implementing advanced functionality such as automatic multi-camera calibration and synchronization, color-format transformations and the ability to provide server-side preprocessing on video streams. A major advantage of cvHAL is that with the recent availability of low-cost gigabit networking and high-speed wireless networking, consumer-grade off-the-shelf cameras can be easily turned into "smart cameras" by connecting them to any networked computer. Smart cameras solutions can cost an order of magnitude more than a cvHAL-based solution. cvHAL provides camera abstraction for the openEyes eye-tracking project and can be downloaded from the openEye website.

5.2. Off-line eye-movement analysis

We use an eye-tracking algorithm that we developed recently known as the Starburst algorithm to perform off-line analyze of the video captured from the eye tracker. This algorithm combines feature-based and model-based image processing approaches to achieve a good trade-off between run-time performance and accuracy for images of the eye captured using a dark-pupil systems [7]. The algorithm is particular adapt at handling noisy images. The algorithm extracts the location of the pupil center and the corneal reflection so as to relate the vector difference between these measures to coordinates in the scene image. The algorithm begins by locating and removing the corneal reflection from the image. Then the pupil edge points are located using an iterative feature-based technique. An ellipse is fitted to a subset of the detected edge points using the Random Sample Consensus (RANSAC) paradigm [3]. The best fitting parameters from this feature-based approach are then used to initialize a local model-based search for the ellipse parameters that maximize the fit to the image data. To calculate the point of gaze of the user in the scene image, a mapping between locations in the scene image and an eye position must be determined. The typical procedure in eye-tracking methodology is to measure this relationship through a calibration procedure [14]. During calibration, the user is required to look at a number of scene points for which the positions in the scene image are known. While the user is fixating each scene point, the eye position are measured. Then a mapping between the two sets of points is generated using a linear homography. The user's point of gaze in the scene for any frame can then be established. Our implementation of this algorithm is made available as part of the openEyes toolkit.

5.3. Real-time eye tracking

We have also implemented a real-time eye-tracking package known as cvEyeTracker that is capable of tracking eye movements at over 20 frames per second. cvEyeTracker software uses Intel's OpenCV libraries in conjunction with cvHAL to process captured video. The algorithm used in this application is relatively simple to afford it speed. Consequently, it is also likely the eye movement measures are less accurate than that obtained with the off-line analysis described above. First, to deal with noise, frame normalization is employed. Then a high and a low threshold are applied. The region above the high threshold is taken as the corneal reflection and an ellipse is fit to this region. The region below the low threshold is taken as the pupil and an ellipse is fit to this region. The vector difference between the centers of these two regions is calculated. cvEyeTracker uses a similar calibration procedure as described above with the exceptions that bi-cubic nonlinear interpolation [14] is used to map the vector differences into the scene coordinate frame. From within a graphical user interface, the user can view the eye and scene images, proceed with the calibration and alter the parameters of the algorithm to maximize accuracy. The implementation of this program is made available as part of the openEyes toolkit.

	1st	2nd	3rd
Generation 3	1.16	1.38	1.71
Generation 4	0.52	1.20	1.57

Table 1: Eye tracking accuracy (degrees of visual angle)

6. Validation Study

An evaluation was conducted in order to test the system. Video was recorded from the third and fourth generation eye trackers while three of the authors viewed two movie trailers presented on a laptop computer. Prior to and after the viewing of each trailer, each user placed their head in a chin rest and fixated a series of nine calibration marks on a white board positioned approximately 60 cm away. The video captured during the evaluation is available for viewing on the openEyes web site. Shown in Table 1 are the accuracy estimates derived from the first, second and third viewings of the calibration grid separately. The starburst algorithm was applied to estimate the points of gaze. Accuracy is measured as the distance between the estimated point of gaze and the actual location of the calibration marks in the scene image averaged over all nine calibration points. The first viewing of the grid is used to calibrate the eye tracker. The results show that the average eye-tracking error is very low in spite of the noise levels induced by extending the CCD sensor away from the board. This degree of accuracy is easily on par with much more expensive, commercially available eye tracking systems. A small decrease in accuracy is seen over the course of the experiment, which can be attributed to some slippage of the head gear. Note the improvement for the fourth generation eye tracker which corresponds to the observed decrease in image noise.

7. Discussion

In future iterations, a number of improvements are will be made to improve the current system design. First, the entire system could be made more mobile with the use of a smaller lighter weight computer. Given that we can now measure the computational demands of on-board real-time eye tracking, we are in a better position to optimize the trade-off between cost, size and computational power of the computer. We will also further investigate the best eye camera lenses. While the first and second generation eye trackers used lenses that we felt had a field of view that was too large, the zoom lens used on later generations required a more delicate alignment and suffered from a smaller depth of field. There is clearly a trade-off between the size of the eye in the image, the depth of field, and the quality of eye tracking that needs to be explored. For ergonomic reasons, further consideration will also be given to selecting thin and flexible cable that has sufficient shielding. We have recently found that it is possible to extend the shielded cable to a length of approximately four feet without a significant change in noise. Increasing the cable length and using more flexible cable should increase the users comfort wearing the system and will allow stowing of the camera boxes in the backpack. We expect that these considerations will also help minimize a degree of head gear slippage that can alter the alignment between the eye camera and the eye.

8. Acknowledgments

JB was supported by the 2004 Summer Program for Research in Human Computer Interaction at Iowa State University sponsored by Applied Science Laboratories. JB is now at the Interactive Telecommunications Program, New York University and can be contacted at babcock@nyu.edu.

References

[1] J. Babcock and J. Pelz, "Building a lightweight eyetracking headgear," in *Eye Tracking Research & Applications Symposium*, San Antonio, TX, USA, March 2004, pp. 109–114.

[2] A. Duchowski, "A breadth-first survey of eye-tracking applications," *Behavior Research Methods, Instruments and Computers*, vol. 34, no. 4, pp. 455–470, 2002.

[3] M. Fischler and R. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[4] D. Hansen and A. Pece, "Eye tracking in the wild," *Computer Vision and Image Understanding*, vol. 98, no. 1, pp. 155–181, 2005.

[5] A. Haro, M. Flickner, and I. Essa, "Detecting and tracking eyes by using their physiological properties, dynamics, and appearance," in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, June 2000, pp. 163–168.

[6] R. Jacob, "The use of eye movements in human-computer interaction techniques: what you look at is what you get," *ACM Transactions on Information Systems*, vol. 9, no. 2, pp. 152–169, 1991.

[7] D. Li, D. Winfield, and D. J. Parkhurst, "Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches," in *Proceedings of the IEEE Vision for Human-Computer Interaction Workshop at CVPR*, June 2005, pp. 1–8.

[8] P. Majaranta and K. Raiha, "Twenty years of eye typing: systems and design issues," in *Proceedings of the symposium on Eye tracking research and applications*, New Orleans, Louisiana, USA, March 2002, pp. 15–22.

[9] C. Morimoto, A. Amir, and M. Flickner, "Detecting eye position and gaze from a single camera and 2 light sources," in *Proceedings. 16th International Conference on Pattern Recognition*, 2002, pp. 314–317.

[10] D. Parkhurst and E. Niebur, "Variable resolution displays: a theoretical, practical and behavioral evaluation," *Human Factors*, vol. 44, no. 4, pp. 611–29, 2002.

[11] D. J. Parkhurst and E. Niebur, "A feasibility test for perceptually adaptive level of detail rendering on desktop systems," in *Proceedings of the ACM Applied Perception in Graphics and Visualization Symposium*. New York, NY: ACM, November 2004, pp. 105–109.

[12] J. Pelz, R. Canosa, J. Babcock, D. Kucharczyk, A. Silver, and D. Konno, "Portable eyetracking: A study of natural eye movements," in *Proceedings of the SPIE, Human Vision and Electronic Imaging*, San Jose, CA, USA, 2000, pp. 566–582.

[13] L. Sibert and R. Jacob, "Evaluation of eye gaze interaction," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, The Hague, The Netherlands, April 2000, pp. 281–288.

[14] D. M. Stampe, "Heuristic filtering and reliable calibration methods for video-based pupil-tracking systems," *Behavior Research Methods, Instruments, and Computers*, vol. 25, no. 2, pp. 137–142, 1993.

[15] L. Young and D. Sheena, "Survey of eye movement recording methods," *Behavior Research Methods and Instrumentation*, vol. 7, pp. 397–429, 1975.