TECHNOLOGY AND APPLICATION OF GAZE BASED CONTROL

Joshua Borah
Applied Science Laboratories
175 Middlesex Turnpike
Bedford, MA 01730

1. SUMMARY

This lecture reviews the potential use of gaze measurement as a means of human interaction with computers and other systems, especially in the military aerospace environment. It addresses the reasons for considering gaze control, reviews techniques for measuring gaze; and discusses physiological, behavioral, and practical considerations for design of gaze based controls.

2. REASONS FOR CONSIDERING GAZE BASED CONTROL

Use of gaze based control is intended to exploit the naturalness, speed and accuracy of visual fixation and tracking. Gaze may be used for explicit control such as designating targets in the external world (E.g., for off-boresight weapons) and selecting items on cockpit displays. It may also be used for implicit control functions such as providing context information for voice or gesture controls, or allowing enhanced resolution of just the local area ("area of interest") being viewed within a display. Gaze control may allow quicker control action while reducing the load on other control resources such as the hands, and in other cases may allow enhanced capabilities (such as the area of interest display) that would not otherwise be possible.

Where line of sight is measured relative to the headgear, it is also necessary to measure the headgear position and orientation in order to compute the eye line of sight with respect to the airframe. In comparison with head pointing alone, gaze based control offers potential advantages of speed, ability to cover a wider angular envelope, and the possibility of less performance deterioration under turbulence-induced vibration or during high-g combat maneuvering.

Many control actions must begin with a visual fixation no matter what the control modality, and there is obvious advantage to exploiting this natural behavior. The natural behavior, however, may not always include a prolonged fixation or an extremely precise fixation, and may include other brief fixations at non related objects. Care is required to maximally exploit natural looking behavior and to minimize requirements for difficult eye gaze actions that must be learned.

3. GAZE TRACKING METHODS

Although mature as laboratory research instrumentation, the current generation of gaze measurement devices has probably not yet reached the level of true practicality for applied use in aerospace cockpit environments. This does, however, appear to be a reachable horizon in the reasonably near term.

Gaze trackers measure line of gaze and/or point of gaze. *Line of gaze* is the imaginary straight line extending from the center of the fovea (the high acuity section of the retina), through the center of the eye lens and out to infinity. *Point of gaze* refers to the point whose image actually forms at the center of the fovea. It is the intersection point of the line of gaze with a visible surface. A gaze measurement system usually includes several subsystems (see Figure 1). An *eye tracker* determines pointing direction of the eyeball with respect to a sensor. If the sensor is head mounted, the system must include a *head tracker* to determine position and orientation of the head with respect to the environment; and if the system is to determine point of gaze rather than just line of gaze, it must also include a processor with knowledge of where visible surfaces are in the environment.

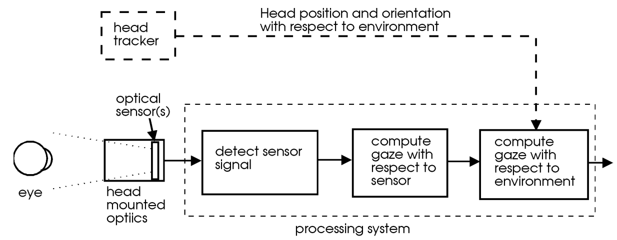


Figure 1. Schematic showing the functional components of a typical gaze tracker

It is important to note that gaze trackers cannot measure what someone is attending to, but rather can only measure aim point of the eye's high acuity area (fovea). This is usually the same as *point of regard* (the part of the visual field that the subject is actually paying attention to), but not always.

Eye tracker performance is often described in terms of the following parameters. *Accuracy* is the expected difference between measured eye line of gaze and true eye line of gaze, usually expressed in terms of visual angle. *Precision* (repeatability) is the expected difference in repeated measurements of the same true eye line of gaze. *Linearity* is the degree to which a change in the measurement is proportional to the actual change in eye angle, and is usually expressed as a percent of the eye angle change being measured. Stated another way, linearity is the amount that a plot of measured values versus actual values is expected to deviate from a straight line. *Resolution* is the smallest change in eye angle that can be reported by the device. *Range* is the amount of eye motion that can be measured, usually specified in degrees visual angle. Range may be specified with respect to the head gear or with respect to the external environment (e.g. airframe), depending upon the device reference frame. *Update rate* is the frequency with which data samples are measured and reported, usually as "samples/second". *Transport delay* is the amount of time that it takes data to travel through the system and become available for use. *Latency* (or *throughput*) usually refers to the amount of time required to accurately reflect a change in the quantity being measured. It is influenced by pure transport delay and also by dynamic operators (for example, a low pass filter) in the signal path. *Bandwidth* is the range of sinusoidal input frequencies that can be processed by the system without significant attenuation or distortion.

The component of a gaze measurement system that is currently most critical to achieving practical gaze tracking in operational aerospace environments is the eye tracker, the

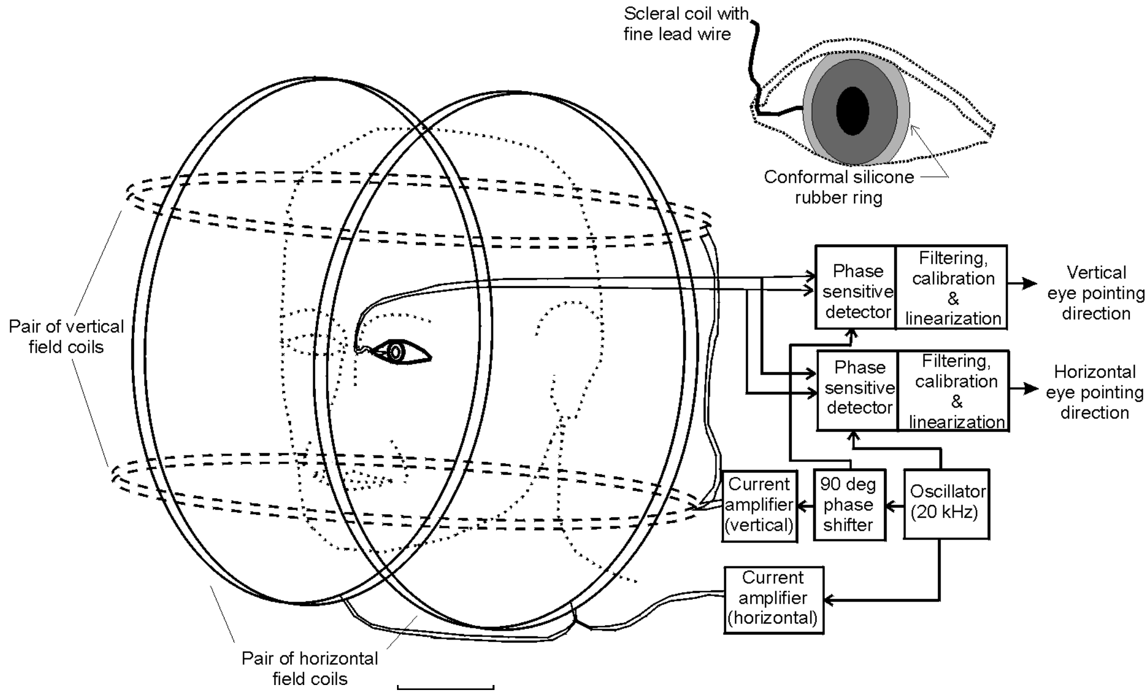


Figure 3. Schematic illustrating the sceral search coil method of measuring eye motion

Figure 2. Schematic illustrating electro-oculographic (EOG) technique for measureing eye motion

instrument that attempts to measure the pointing direction of the eye ball. The predominant eye tracking techniques are discussed below. They can be classified as *electro-oculographic*, *scleral coil*, and *optical* methods.

ELECTRO-OCULOGRAPHY

The retina at the back of the eye develops a small negative electrical charge relative to the front surface of the cornea, probably as a result of its higher metabolism [1]. Electro-oculography uses skin electrode pairs placed on either side of the eye, and above and below the eye to measure the direction of this electric dipole. The corneo-retinal dipole induces zero differential voltage when the dipole axis is about midway between electrodes. A change of about 20 μV/° results when the eye is rotated towards one of the electrodes. Rather than making independent measurements for each eye it is common to mount a single pair of electrodes at the outer canthi of both eyes, as shown in Figure 2, to sense their combined horizontal effect. To minimize drift, a "reference" electrode, sometimes sited at the center of the forehead, is usually connected to the amplifier ground.

Small commercial (silver)+(silver chloride) skin electrodes, normally used for monitoring the heart functioning in babies, are commonly employed to minimize electro-chemical artifacts. The skin is cleaned and de-greased with an alcohol swab. Then the contact surface is wetted with conductive saline gel and the electrode is fixed in place using the adhesive backing ring. The short leads are connected to high gain, high (>1 [illegible]) impedance, low noise, low drift differential amplifiers having a bandwidth from zero to about 100 Hz. Calibration is required to scale and map the EOG signals to coordinates of gaze with respect to the head. The resulting measurement has high temporal bandwidth and provides an excellent measure of eyeball dynamics, but determination of absolute line of gaze with respect to the head is significantly compromised by drift. In order to keep the signals within the dynamic range of the amplifiers, it is usually necessary to periodically re-zero the output when the subject is known to be looking straight ahead. Extrapolating from laboratory measurements by Shackel [2] and in flight tests conducted in a Jaguar aircraft [3] it seems reasonable to conclude that EOG measures in a cockpit environment might allow inference of eye pointing direction relative to the head with an expected error between about 3° and 7°, assuming some form of filtering and frequent re-zeroing.

SCLERAL COIL

The scleral coil, first developed by Robinson [4], requires attachment of a sensing element to the subject's eye. A very fine induction coil is embedded in a shallow ring of silicone rubber, the inner surface of which is slightly hollow, so that it adheres to the limbus (the boundary between the clear cornea and the white sclera) by capillary action and suction and remains concentric with the corneal bulge. Thin wires connecting the coil to sensing electronics are usually brought out across the nasal corner of the eye and taped securely to the side of the nose. The subject sits with his head inside two sets of orthagonally oriented Helmholts coils (see Figure 3). One pair of excitation coils creates an oscillating horizontal magnetic field. This field induces a current in the sensor coil proportional to the sine of the horizontal angle between the scleral coil axis and the field. The other set of excitation coils, energized with a signal at 90° phase shift to the first, creates an orthogonal field resulting in induced sensor coil current that is similarly related to vertical angle. Phase sensitive detection is used to find an induced signal that is exactly in phase with each set of excitation coils. If both eyes are outfitted with scleral coils, the pointing direction of both eyes can be measured with respect to the fixed excitation coils. The sclera ring can also be equipped with an orthaginal sensing coil allowing measurement of eye torsion (rotation about the line of sight).

Complete scleral coil systems are available commercially [5]. Scleral coil systems are extremely accurate, fast, and dependable. Following a simple calibration to define the initial reference orientation of the eye, the rotations can be measured to a resolution of about 1 arcmin over a range of about ±15° to an accuracy of about 1% of the range, with a typical bandwidth of 0 to 200 Hz. Scleral coil tracking is also distinctly invasive and requires Helmholz coils that will probably be difficult to integrate on aircrew helmets or affix to the cockpit.

OPTICAL EYE TRACKING TECHNIQUES

Optical eye tracking techniques make use of optically detectable eye features and geometry to determine the orientation of the eye ball.

The following features, illustrated in Figure 4 are most commonly used :

- *Limbus* -- the boundary between the colored iris and white sclera.
- *Pupil* -- the opening in the iris (aperture of the eye)
- *Corneal reflection (CR), or first Purkinje image* -- mirror reflection of an external source from the outer surface of the cornea
- *4th Purkinje image (4PI)* – mirror reflection of an external source from the rear surface of the eye lens.

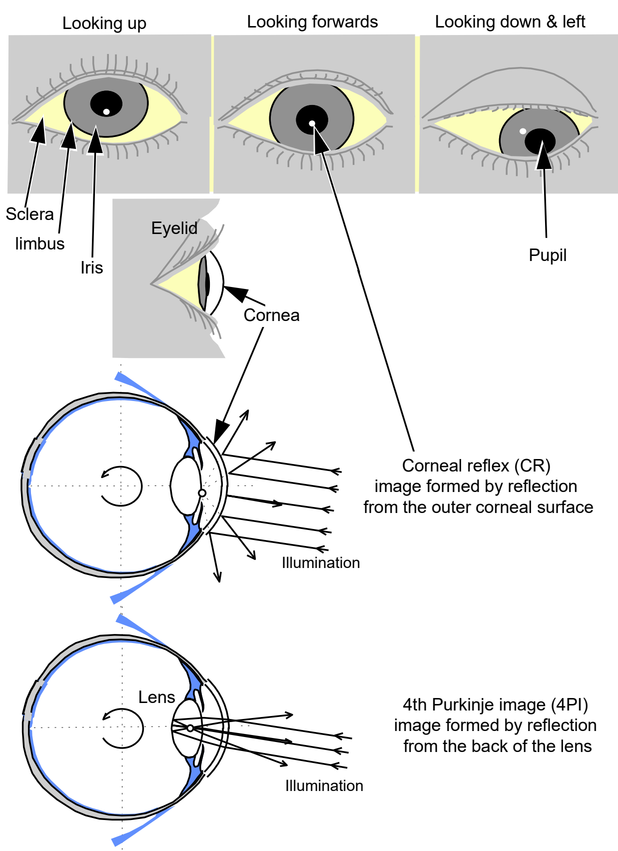


Figure 4. Eye features exploited by optical eye tracking systems

Eye ball orientation can be computed from the position of a single feature if the sensor is assumed to be rigidly fixed to the head or if sensor position with respect to the head can be independently measured.

Position of a single feature alone will not distinguish rotation of the eye ball from movement of the sensor. Multiple landmarks located at different radii from the center of the eye ball will appear to move with respect to one another as the eye rotates, but will move together when the sensor translates. By differentiating between eye rotation and translation with respect to a sensor, dual feature techniques minimize errors due to shifting of head mounted optics, and also allow use of non head mounted optics. Dual feature systems usually use the pupil and corneal reflection, or the corneal reflection (CR) and 4th Purkinje image (4PI). The pupil forms a landmark near the eye ball surface (about 9.8mm from the eye ball center), the CR behaves as would a land mark at the same radius from eye ball center as the corneal center of curvature (about 5.6 mm), and the 4PI appears to move the same amount as the posterior lens surface center of curvature (about 11.5 mm from eye ball center).

Single features or groups of features on the surface of the eye, often dominated by the boundary between the dark iris and white sclera, can be tracked with small numbers of photosensors. An example is shown in Figure 5. This allows very high temporal bandwidth and fine spatial resolution, using essentially analog processing, but results in poor static accuracy because movement of the optics with respect to the eye cannot be distinguished from eye rotation. Very small relative movements are confused with relatively large rotations.

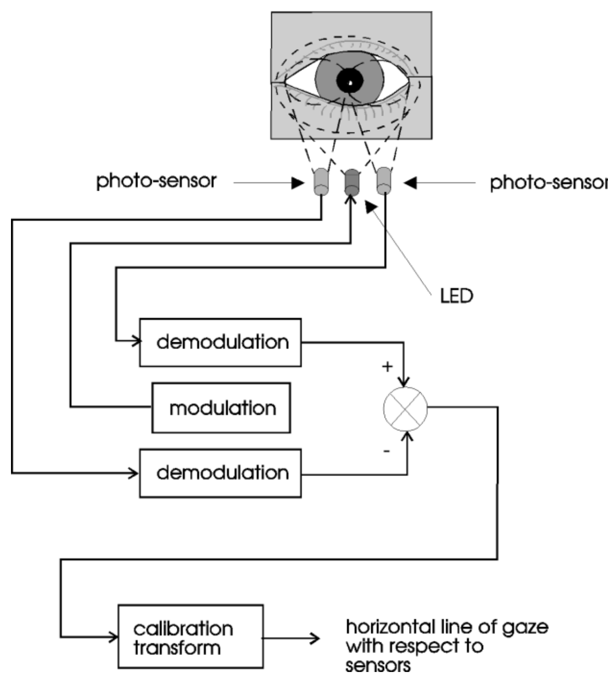


Figure 5. Schematic showing a simple photo sensor system for tracking horizontal position of the limbus. As the eye rotates, proportionately more white sclera is exposed to one photo-sensor than to the other.

The relative positions of the corneal reflection and 4th Purkinje image ("dual Purkinje image technique") can be used to track eye rotation with excellent precision and accuracy (on the order of 20 arc seconds). The only commercially available dual Purkinjie image system images the two features onto separate quadrant detectors, and uses separated closed-loop servo-controlled pathways to keep the features centered on the detectors. The servo-control signals are a measure of the feature positions. This is also an essentially analog process allowing very high temporal bandwidth.. Range, however, is relatively small (typically ±10°) because the 4PI is visible only within the lens opening and is obscured when it moves behind the iris. The only currently available system is engineered as a large bench mounted optical assembly that is impractical for airborne application.

The most prevalent optical technique, and the one used by systems that currently come the closest to being practical for airborne use, tracks the relative position of the corneal reflection and the center of the pupil. CR/Pupil devices are not as precise or accurate as the dual Purkinjie image device, but can be far less obtrusive and easier to operate primarily because the pupil is easier to detect the 4PI.

Generally the eye area is illuminated by a near infra red source (or multiple sources) and a solid state video camera captures an image of the eye. The camera is typically filtered to receive only light of the wavelength produced by the eye tracker's near infra red source.

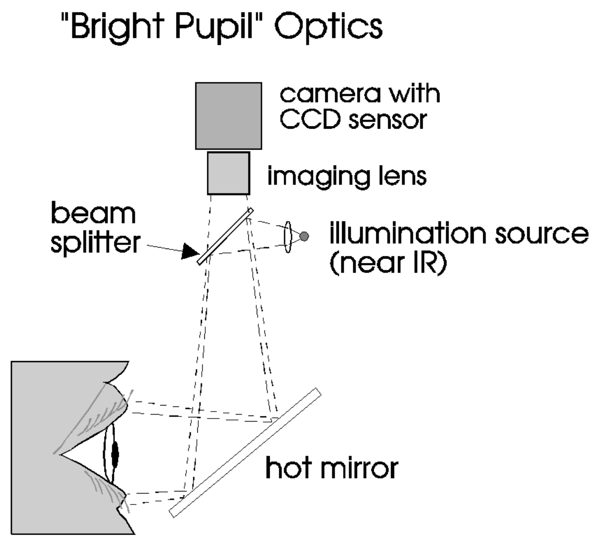
If the optics (camera, illuminator, and lenses) are mounted to the user's head gear, a hot mirror (beam splitter that reflects IR and transmits visible wavelengths) is usually used to reflect near IR light to the optics while still allowing unobstructed vision for the wearer. This is illustrated is Figure 6. Alternately, non head mounted optics may use a moving mirror or moving camera platform to follow head motions.

The eye acts as a retro-reflector. If the eye illumination beam is coaxial with the camera, light reflected back from the retina is captured by the camera making the pupil appear to be a bright circle. This accounts for the red eye effect sometimes produced by flash photography. Off axis illumination produces the more familiar dark (black) pupil image (see Figure 6). Dark pupil images provide easier pupil detection in very bright environments (e.g. sunlight), whereas bright pupil images provide easier detection in darker environments.

The corneal reflection is significantly brighter than any other visible feature, including a bright pupil image, and is relatively easy to detect so long as it is not obscured by the eye lids or confused with corneal reflections from some external sources.

Real time image analysis is used to identify the pupil and corneal reflection and find their centers. Relative feature brightness is often a primary discrimination criterion, but more and more sophisticated pattern recognition techniques are being used as the amount of readily available digital processing power increases. This makes it possible to recognize the features of interest in real conditions and cope with extraneous reflections, partial eye lid occlusion and motion-induced blur [6]. Calibration is required to account for individual eye geometry and optics placement.

Range for CR/Pupil systems is limited to about ±25° by CR excursion within the cornea for systems with a single illumination source, but can be extended considerably if multiple sources are used, especially on the horizontal axis. Vertical range in the downward direction is often limited, by eye-lid occlusion of the pupil, to 5-10° below the line of gaze that would look directly at the eye camera. For this reason the optics are usually set so that the camera views the eye from about 5° below the nominally straight ahead direction. Note that for head mounted optics these range limits apply to eye motion with respect to the head. Measurable line of gaze with respect to the airframe is limited only by head motion.



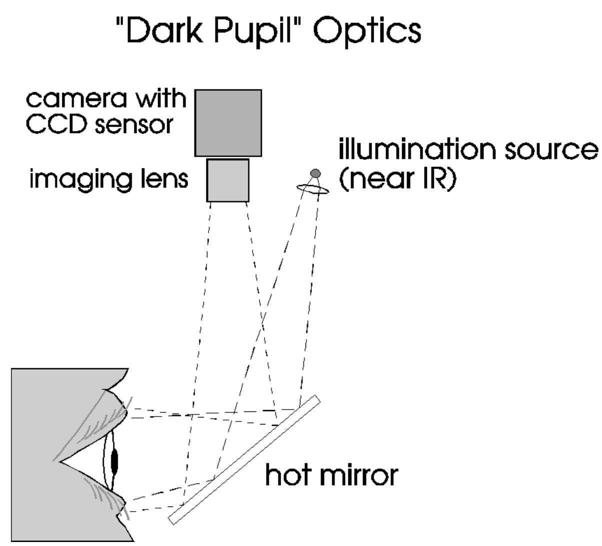


Figure 6. "Bright pupil" and "dark pupil" optics for CR/pupil tracking

Accuracy is usually on the order of 1° visual angle for CR/Pupil systems. It may vary from 0.5° or better during very careful laboratory tests with selected subjects, to > 2° under difficult conditions or with difficult subjects. Precision and resolution are usually in the range of 0.1° to 0.5° depending upon the particular system and upon operating conditions. Frequency response is usually limited by video frame rates to 50 or 60 samples/sec. Higher update rate sensors are available, but require sacrifices in spatial resolution, physical size, and sensitivity.

There are numerous developed CR/pupil systems, but currently no militarized systems, and no systems robust enough for operational military use. Current systems have difficulty handling bright sunlight which tends to saturate the entire camera image and degrades effective contrast, are not able to automatically make adjustments for changing ambient conditions, have not yet been properly integrated with military aviation head gear, and are not quite reliable enough for operational service. It is reasonable, however, to think that the necessary advances to CR/pupil systems can be made in the near term.

Illumination of the eye by optical eye trackers must, of course, remain within safe levels. Safety standards are published by numerous agencies including *The American Conference of Governmental Industrial Hygienist (ACGIH), The American National Standards Institute (ANSI)*, the *Federal Food and Drug Administration* (FDA), and *The International Electrotechnical Commission (IEC).* The standards vary, with the IEC standard currently being the most restrictive. Under IEC standards, for example, the source must be safe even if viewed from the closest mechanically possible distance through a magnifying glass.

CALIBRATION

All eye tracking methods require a transformation to convert the measured quantity to the desired quantity. For example separation between the pupil and CR must be converted to point of gaze coordinates on a surface, or a line of gaze vector in a particular coordinate frame. For all eye tracking methods discussed, with the possible exception of the scleral coil, the transformation parameters vary between subjects, optics placement, and other conditions.

Calibration refers to a procedure for gathering data, and for using the data to compute transformation parameters. The procedure usually consists of asking the subject to look at number of pre-defined points, while storing samples of the measured quantity (e.g. pupil and CR position).

The transformation can either be a form of interpolation, a set of continuous equations, or some combination of these. The details vary widely among available systems. Theoretically, the transformation can remove any systematic error that is a function of the measured quantities. In practice, there is a limit to the amount of data that it is reasonable to gather with a calibration procedure, and therefore a limit to the complexity of the transformation.

The accuracy and linearity of eye tracker measurements are limited by the underlying precision (repeatability) of the measured quantity. Up to that limit, accuracy and linearity are determined by the quality of the calibration and transformation scheme. Adding calibration target points, and using the additional data to add interpolation points or to increase the order of a polynomial transform, usually improves accuracy, but with diminishing returns. Typical calibration schemes require 5 or 9 pre-defined points, and rarely use more than 20 points.

An example of a 2 dimensional interpolation scheme can be found in references [7] and [8]. A cascaded polynomial curve fit method is described in reference [9]. Many other variations of these schemes are possible.

PROGNOSIS FOR AIRBORNE OPERATIONAL USE

Eye tracking is a relatively mature technology only in the R&D domain. No currently available eye tracking systems are dependable enough or automatic enough for operational flight applications, nor are there any current systems available in a militarized configuration. All current devices require a skilled equipment operator (other than the person being measured) for optimal use.

Scleral coil and differential PI tracking seem likely to remain laboratory techniques unless some major leaps in the technology occur. These techniques are by far the most accurate of the major techniques in use, but they both present major practical problems. Scleral coil tracking is invasive and requires a Helmholz coil that will probably be difficult to integrate on aircrew helmets. Differential PI tracking has too restrictive a range limitation and too complex an optics package to be easily helmet mounted and ruggedized.

EOG may very well have a place in aircraft as a back up measurement system, an enhancement to add temporal bandwidth to another type of eye tracker, or for use in some control function that requires only detection of eye movement, rather than absolute line of gaze. The accuracy of EOG alone is never likely to be adequate for line of gaze determination.

Differential CR/pupil tracking systems are generally the most unobtrusive eye tracking systems available, and, with head mounted optics, currently come closest to being appropriate for operational use in flight. Those systems, using dark pupil optics along with some form of illuminator strobing and sensor shuttering, are currently best able to operate in daylight and under vibration. The static accuracy (about 1° visual angle) and range (50° horizontal and 40° vertical field with single illumination source) of current CR/pupil tracking devices is adequate for implementing or assisting a variety of tasks in the aerospace environment, although increased accuracy would certainly expand the potential use for eye tracking.

No currently available CR/pupil systems are yet nearly robust enough, automatic enough, or properly integrated with aircrew head gear and military electronics. Robustness must be significantly improved to ensure dependable operation for different users under varying light conditions in operational environments. Automatic operation must be significantly enhanced so that the user can don the equipment and calibrate the system without help, and there-after depend upon proper operation with no intervention by a second person. Optics must be integrated with the appropriate head gear and head mounted display systems, and both optics and electronics must be hardened and militarized. Work is underway in all of these areas, and there is no reason to think that such enhancements cannot be realized with currently available optics, sensor, and processing technology.

Significant improvement of CR/pupil system accuracy is clearly possible, but far less certain, especially in operational environments. Improvements can theoretically be made by using increased processing power to more accurately find the center of the oval pupil in the presence of partial image occlusions, and ragged edges; use of higher order calibration schemes to remove more of the systematic error; use of additional variables in calibration, such as pupil diameter, to further account for systematic effects; use of precision sensor arrays, or sensor arrays that are mapped and compensated for spatial non-linearity's; etc. Such gains may be more than counter-balanced, however, by additional error introduced under the rigors of operational use. Furthermore, the lengthy, careful calibration procedures probably required for exquisite accuracy may be contrary to operational imperatives for quick and easy set-up. Designers may want to consider eye tracking tasks tailored to require less rather than more accuracy and precision in order to improve the chances of meeting robustness and ease of use requirements.

Major eye tracking techniques are summarized in table I

Table 1. Summary of most prevalent eye tracking techniques

Method	**Typical Applications**	**Typical Attributes**	**Typical Reference Frame(s)**	**Typical Performance**
• EOG	• Dynamics of saccades • smooth pursuit • nystagmus	• High bandwidth • Eyes can be closed • In expensive • Drift problem (poor position accuracy) • Requires skin electrodes - otherwise unobtrusive	• head	• static accuracy: ~3°-7° • resolution: with low pass filtering & periodic re-zero, virtually infinite • bandwidth: ~100 Hz
• Scleral Coil	• Dynamics of saccades, smooth pursuit, nystagmus • Miniature eye movements • Point of gaze • Scan path	• Very high accuracy and precision • Invasive • Very obtrusive	• Room	• accuracy:~15 sec arc • resolution: ~1arc min. • range: ~ 30° • bandwidth: ~200 Hz
• Limbus (using small number of photo sensors)	• Dynamics of saccades, smooth pursuit, nystagmus • Point of gaze • Scan path	• High bandwidth • Inexpensive • Poor vertical accuracy • Obtrusive (sensors close to eye) • Head gear slip errors	• head gear	• accuracy: varies • resolution: 0.1° (much better res. possible) • range: ~30° • update rate: 1000 samples/sec
• CR/Pupil	• Point of gaze • Scan path	• Minimal head gear slip error • Unobtrusive • Low bandwidth • Problems with sunlight	• Head gear • Room (airframe)	• accuracy: ~1° • resolution: ~ 0.2° • hor. range:~50° • vert. range: ~40° • update rate: 50 or 60 samples/sec
• CR/4PI	• Dynamics of saccades, smooth pursuit, nystagmus • Miniature eye movements • Point of gaze • Scan path • Image stabilization on retina • Accommodation	• Very high accuracy and precision • High bandwidth • Obtrusive (large optics package, restricted head motion) • Limited range	• Room	• prec: min. of arc • range: ~20° • update rate: 500 samples/sec

4. HUMAN PHYSIOLOGICAL AND BEHAVIORAL CONSIDERATIONS

Even if instrumentation could make perfect point of gaze and line of gaze measurements, human physiological and behavioral characteristics impose certain constraints, and it is very important to keep these characteristics in mind when considering gaze based control.

Scanning behavior is described by a series of fixations (stopping points), saccades (extremely rapid jumps between fixation points) and smooth pursuits. When examining a stationary scene, both lines of sight are simultaneously held steady for short periods (usually 200 - 600 msec), called fixations, to bring a feature of interest within the approximately 1° angular range of the fovea. Miniature eye movements, of up to several minutes of arc do occur during the periods of "fixation", but are not perceived.

Rapid jumps, called saccades, move the eye between fixations.. Saccades usually reach velocities of 400-600 °/sec, and last 30 - 120 msec. Vision is significantly suppressed during this period. Although saccades can be as large as 50°, they are more commonly 1 - 20°. If a target appears in peripheral vision, it takes a minimum of about 100 msec for a saccade towards the target to be initiated.

When observing a slowly moving object the lines of sight usually track smoothly, but this pursuit reverts to fixations and saccades when the object moves faster than about 30°/sec. Without specific training, smooth eye movements are only possible when following a smoothly moving target or compensating for head movement. A thorough review of eye movement behavior can be found in [10].

Even if we could measure direction of the visual axis perfectly we would not have perfect knowledge of point of regard. Visual acuity is best on the foveal region of the retina, and within the fovea is best near the very center. People therefore direct the visual axes of the eyes (axes passing through the center of each eye lens and fovea) to objects that they want to see clearly; however, there may be a foveal "dead zone" or "indifference threshold" on the order of about 0.3 degrees visual angle for fixation of stationary targets [11]. Attention can be shifted within the foveal region, and even outside of the foveal region if the target of interest falls within acuity limits of peripheral vision [12, 13, 14]. Furthermore, foveation accuracy falls off markedly if a person attempts to maintain fixation for several seconds [15], when tracking a moving target [10, 16, 17, 18], or during rapid head movements [19]. If a person consciously attempts to fixate a small, stationary, target, for a short time, while holding their head steady, we can probably assume the visual axis to be within 0.3 degrees of the target. This should not be confused with visual acuity. People can visually determine whether one object is aligned with another (a traditional aiming task) with precision on the order of arc minutes.

Eye movement with respect to the head (rotation of the eye ball within the socket) has a maximum range of about ±50° horizontally, and about +40°, -60° vertically , but normally remains within about ±15°-20° [20, 21]. Gaze shifts beyond the central 20° field are usually, although not always, accompanied by head rotation. Horizontal eye rotation with respect to the head of more than about 40° from the central position becomes quite uncomfortable if maintained for several seconds.

The normal fixation/saccade pattern of visual scanning can be thought of as a continual series of snap shots that are used to create a mental image of the visual environment; however this is usually an unconscious process. Perception of the environment is of the "single picture" formed in the brain. People are normally not aware of their scan pattern (although they can be if they make a special effort), but rather are normally aware only of the resulting mental image of the environment.

People are not accustomed to consciously controlling their gaze. Unintentional actions initiated by unintentional glances is sometimes called the "Midas touch" problem of gaze control.

It is difficult and annoying, although possible, to maintain steady fixation on a single target for significantly more than a second. Fixations of several hundred milliseconds are most natural. There is also a strong tendency to make quick glances at other nearby targets during unnaturally long fixations.

The eye is drawn to features, and it is very difficult to fixate a blank spot.

Secondary visual feedback (presentation to a person of their own gaze point as measured by an eye tracker) must be handled carefully. Continuous feedback of measured gaze position, if not very accurate and up to date, can sometimes be distracting instead of helpful. If the displayed indicator is slightly displaced from the central line of gaze there may be a tendency to continually try to look at it, leading to a positive feedback loop; however, it requires only minimal conscious effort to avoid this.

The vestibular system helps the brain to stabilize the perceived visual field in the presence of head motion and vibration, and to partially, but not completely, stabilize line of gaze with respect to the visual environment. Steinman and Collewijn [19] showed, for example, that in the case of voluntary head rotations of 2.5-5 Hz during fixation on a distant target, eye motion sometimes compensates for only about 80% of head motion and the two eyes do not move equally, although vision remains clear and fused.

Eye movement appears to remain unaffected under G_z loading that is sufficient to make head motion difficult. This is supported by only a small amount of empirical data [22], but is consistent with anecdotal evidence and mechanical analysis. The eye is well supported, and has a relatively small moment of inertia. Because of its roughly spherical, homogeneous structure inertial forces would not be expected to produce large rotational moments. Furthermore, eye movements do not cause the disorienting motion sensations that can be produced by moving the head, and hence the vestibular sensors, in the presence of high inertial forces.

5. DESIGN CONSIDERATIONS FOR GAZE BASED CONTROL

FIXATION FILTERING

As previously discussed Scanning behavior is described by a series of fixations, saccades, and smooth pursuits. When using gaze measurement to determine point of regard, as often required for gaze based control functions, it may be desirable to recognize fixations while filtering out saccades, pursuits, and measurement noise. This is usually done by looking for periods of longer than some threshold time during which gaze remains within a threshold (maximum) area or during which eye movement velocity remains below a threshold. Typical threshold values would be 100 msec, 1° visual angle (circular radius), and 10°/sec, respectively.

USE OF EYE POSITION FEEDBACK

It is usually important to provide some sort of feedback so that the user of a gaze control system knows that the system is functioning properly and can make adjustments if necessary. In its simplest form this feedback is a cursor superimposed on the visual scene continually showing measured gaze point In theory such feedback, often called secondary visual feedback, can improve system accuracy by allowing the user to adjust his gaze point to correct measurement errors. There is empirical evidence that precise secondary visual feedback can improve visual smooth pursuit performance [23, 24]. In practice, this type of feedback may sometimes prove more annoying than useful depending upon the amount of error, noisiness, and latency in the gaze measurement [25, 26, 27]. In such cases it may help to present measurements that have been filtered by a fixation detection algorithm (feed back cursor moves only when a new fixation is detected), and/or to display a transparent cursor that is at least as large as the expected gaze measurement system error. Alternately it may sometimes be best to present feed back by highlighting the object that the system computes gaze to be indicating, for example, a display icon or object, rather than by continually displaying measured gaze point. If the gaze control task is to designate an external target such as an opposing aircraft or ground target, then some form of continual measured gaze point feed back may be required.

ACTION CONFIRMATION

Even if gaze is available as a control input, not all fixations will be intended as explicit control actions. There will still be present the semi conscious pattern of fixations ("snap shots") that normally form our perception of the visible environment. To avoid a "Midas touch" affect, any explicit gaze control system must include means to differentiate fixations that are intentional control actions from glances intended only to acquire visual information. Confirmation protocols can range from requiring a slightly unusual gaze behavior (longer than average fixations, a sequence of blinks, or a long blink) to manual action. Control actions that are more consequential or more difficult to undo should require a more reliable mode of confirmation than less consequential actions. Citing an example from Jacob [25], if gaze information is used to cause menus on a display to automatically "pull down" when the menu title is fixated, and "roll up:" again when the menu is no longer being viewed, the consequence of unintentional activation of this feature is minimal. A sensible activation criterion might be a fixation on the menu title that is slightly longer than the average 250 msec fixations that characterize scanning behavior. Further increasing the required fixation time makes unintentional activation less likely at expense of longer task execution times and increasingly unnatural behavior.

An example at the other extreme would be use of gaze to designate external targets to a weapon delivery system. In this case very high precautions must be taken against unintentional activation, and manual confirmation, in the form of a trigger pull or button press is probably warranted. Furthermore, the pilot must first have good feedback providing assurance that the intended target has indeed been selected.

Jacob [25] used a hierarchical set of techniques to confirm display manipulation actions depending on the consequences of inadvertent action. Actions that were benign and easily reversible required only short fixations for activation. Actions that were not as easily reversible required longer fixations or manual confirmations. He found that when the eye tracker was working accurately and dependably it felt as though the system were "reading the user's mind", but when eye tracker performance was not stable enough or not accurate enough it was extremely frustrating to the user.

GAZE CONTROL TASKS

Explicit gaze control tasks usually involve some type of gaze designation, including designation of external targets from within a cockpit, selection of real or virtual switches within a cockpit, and a range of "cursor control" type tasks (designation of icons, menu labels, screen locations, etc.) associated with fixed or head mounted displays.

Unambiguous gaze designation requires that targets be separated by at least twice the maximum expected error (E.g., 95% confidence interval) of the measurement, so that confidence intervals drawn about adjacent targets will not overlap. With current state of the art for unobtrusive eye trackers, this would correspond to separations of at least 2° visual angle (to accommodate 1° errors), and probably somewhat more in demanding environments. If gaze measurement is to be used to designate an external target to a "lock before launch" weapon system, the capture field of the "locking" system must correspond to a similar visual angle. If icons on a display are to be designated by gaze alone, the same limit (twice the expected error) defines the minimum space between the borders of adjacent icons, or at least between central fixation targets within each icon.

Eye designation has been investigated in the lab, and generally has been found to be as fast or faster than manual designation, and faster than head designation, so long as the eye tracker is working dependably and so long as the task employs large enough targets to be well within the accuracy capability of the eye tracker [26, 27, 28, 29, 30]. Applied use of eye designation to date has been primarily restricted to systems that facilitate communication for people with motor control disabilities, and there are several commercially available systems that specifically support this application. Performance of eye designation tasks may sometimes be significantly enhanced by use of fixation filtering algorithms, but in general, accuracy of unobtrusive eye tracking systems does not permit as fine a control capability as mouse, trackball, or other manual techniques.

Multi mode control can be used to achieve greater precision, again at the expense of longer task execution times and a requirement for less natural behavior. For example, gaze can be used to position a display cursor near the point of gaze, followed by use of a manual control for fine positioning. Especially in the case of a large cluttered display it is sometimes time consuming simply to find the cursor. In this case, use of gaze to quickly position the cursor to within easy view has the potential to save time over manual control alone; although it would not be as fast or natural as designating a large enough target with gaze alone. Note that besides the manual (or other mode) control for fine cursor positioning, the multi mode technique requires a switch to designate the mode change, and some learned behavior to properly sequence actions (*gaze > mode switch > manual control*).

Gaze measurement can be an excellent tool for "context disambiguation" of voice commands. For example in response to the verbal command *"zoom"*, point of gaze can be used to determine which display or display section to expand. Citing an example from a mission planning interface developed by Hatfield [29], the *command "nav designate steer point"* sets the steer point to the radar display position being fixated at the time the verbal referent *"designate"* was detected. In this way some operations that would otherwise be sequential can be made concurrent. Although context disambiguation takes advantage of natural behavior (E.g., a person is likely to be looking at the display that they want to "zoom") this type of control is not entirely implicit. A particular behavior is required which, although likely, is not certain without explicit intent. Use of gaze to provide context and position information for verbal commands has been tested fairly successfully in several laboratory studies [25, 31, 32, 33, 34].

Gaze measurement has been used in military aircraft simulation to create computer generated, out-the-window displays which have high resolution only in the immediate area about the pilot's gaze point. This "area of interest" display technique was originally motivated by the difficulty in producing computer graphics with both the wide field and high resolution desired. Since high visual acuity is present only in the area near the gaze point, only this area need have high resolution at any given time in order for the entire display to be perceived as being rich in detail. In future operational cockpits, a similar moving area of interest concept might prove valuable with respect to head mounted displays in order to increase the richness of the display about the region of the current gaze point. This may, for example, involve slaving an external sensor with a very narrow cone of sensitivity to follow gaze, and displaying resulting information in the corresponding field within a head mounted display.

Use of gaze measurement for external sensor slaving and/or area of interest display function constitutes an implicit control function. It works entirely in response to normal gaze behavior. For the simulator application described above it proved important to minimize latency and to smoothly blend the boundary between high and low resolution sections in order to preserve the illusion of a detailed out-the-window world. Total latency (gaze tracker plus display) probably needs to remain below 50 msec to preserve the illusion, and this proved a significant problem for the simulation application. Future operational application of this concept is more likely to have the purpose of making additional information available rather than creating an illusion of realism, and although the blending and latency factors will still be of importance in order to avoid annoying or distracting the user, the criteria may be somewhat less stringent.

6. CONCLUSIONS

Gaze measurement may enable a range of potentially useful explicit and implicit control functions. The technology is not yet mature enough for operational airborne application, but the necessary advances can probably be made in the near term. When considering gaze control functions, designers need to consider natural human gaze behavior and the tradeoffs between performance requirements imposed (precision, accuracy, latency, etc.) and robustness, simplicity, and ease of use in difficult environments.

7. AKNOWLEDGEMENTS

This presentation is based on work done as part of AGARD Working Group 25, and contains contributions from all members of that group. It relies especially heavily on the efforts of two other members of that group: Dr. Don Jarrett, DERA, Farnborough, U.K,; and Dr. Alain Leger, Sextant Avionic, Bordeaux, France.

8. REFERENCES

1. Young L. R. and Sheena D., "Eye Movement Measurement Techniques", in Webster (Ed) "Encyclopedia of Medical Devices and Instrumentation", New York, John Wiley & Sons, 1988
2. Shackel B., "Eye movement recording by electro-oculography", in "A manual of Psychophysiological methods", North-Holland Publ. Co., 1967.
3. Vivash J. P., Belyavin A. J., "Eye movements under operational conditions", in Waters M. and Stott J. R. R. (Eds) "Journal of Defence Science", 1, 2, 1996.
4. Robinson D. A., "A method for measuring eye movement using a scleral search coil in a magnetic field", IEEE Transactions on Biomedical Electronics, BME-10, 1963, pp 137-145.
5. Ferman L., Collewijn H., Jansen T. C. and Van den Berg A. V., "Human gaze stability in the horizontal, vertical and torsional direction during voluntary head movements, evaluated with a three-dimensional scleral coil technique", Vision research, 27, 1987. pp 818-828
6. Brinicombe, A. M., Boyce, J. F. and Durnell, L., "Direction of regard determination", in Delogne, P., (Ed) "Proc. Intl. Conf. on Image Processing", Lausanne, Switzerland. IEEE Signal Processing Society, 1996.
7. McConkie, G. W., "Evaluating and reporting data quality in eye movement research", Behaviu Research Methods & Instrumentation, 13, 2, 1981, pp 97-106.
8. Kliegle, R. and Olson, R. K., "Reduction and calibration of eye monitor data", Behavior Research Methods & Instrumentation, 13, 2, 1981, pp 107-111.
9. Sheena, D. and Borah, J., "Compensation for some second order effects to improve eye position measurements", in Fisher, D. F., Monty, R. A., and Senders J. W., (Eds) "Eye Movements: Cognition and Visual Perception", Hillsdale, Lawrence Erlbaum Associates, 1981.
10. Hallet,P.E., "Eye Movements", in Boff, K. R., Kaufman, L., and Thomas, M. P. (Eds) "Handbook of Perception and Human Performance - Vol. 1", New York, John Wiley and Sons. 1986.
11. Young, L.R., "The sampled data model and foveal dead zone for saccades", in Zuber B. L. (Ed) "Models of Oculomotor Behavior and Control", Boca Raton, CRC Press, 1981.
12. Julez, B., Gilbert, E. N., Shepp, L. A., and Risch, H. L., "Inability of humans to discriminate between visual textures that agree in second-order statistics - revisited", Perception, 2, 1973, pp 391-404.
13. Bergin, J. R. and Julez, B., "Parallel versus serial processing in rapid pattern discrimination" Nature, 303, pp 696-698.
14. Scinto, L. F. M., "Retinal inhomogeneity and the allocation of focal attention during fixation", in "The Annual Meeting of the Applied Vision Association" St. John's College, Oxford. 1988.
15. Yarbus A. L., "Eye Movements and Vision", New York, Plenum Press, 1967.

16. St.Cyr G. L. and Fender, D. H., “Non-linearities of the human oculomotor system: Gain”, Vision Research, 9, 1969, pp 1235-1246.

17. Michael, J. A., Melvill Jones, G., “Dependence of visual tracking capability upon stimulus predictability”, Vision Research, 6, 1966, pp 707-716.

18. Kowler, E. and McKee, S., “Sensitivity of smooth eye movement to small differences in target velocity”, Vision Research, 27, 6, 1987, pp 993-1015.

19. Steinman, R. M. and Collewijn, H., “Binocular Retinal Image Motion During Active Head Rotation”, Vision Research, 20, 1980, pp 415-429.

20. Yamada. M., “Head and eye coordination analysis and a new gaze analyzer developed for this purpose”, in d'Ydewalle G. and Van Rensbergen, J. (Eds) “Visual and oculomotor functions”, Elsevier, 1994, pp 423-434.

21. Bahill, A. T., Adler, D. and Stark, L. “Most naturally occurring human saccades have magnitudes of 15 degrees or less”, Investigative Ophthalmology, 14, 1975, pp 468-469.

22. Sandor, P. B., Hortolland, I., Poux, F., and Leger, A., “Orientation du regard sous facteur de Charge Aspects methodologiques Resultats preliminaires”, in “AGARD Meeting on Virtual Interfaces: Research and Applications”, October, 1993.

23. Peli, E. and Zeevi, Y. Y., “Multiple visual feedback loops in eye movement control”. in “XII International Conference on Medical and Biological Engineering”, Jerusalem, 1979.

24. Kenyon, R. V., Zeevi, Y.Y., Wetzel, P. A., and Young, L. R., “Eye movement in response to single and multiple targets”, US Air Force report AFHRL-TR-84-29, 1985.

25. Jacob, R. K., “Eye Tracking in Advanced Interface Design”, in Barfield, W. and Furness, T. A. (Eds.) “Virtual Environments and Advanced Interface Design”, New York, Oxford University Press, 1995.

26. Calhoun, G. L. and Janson, W. P., “Eye Line-of-Sight Control Compared to Manual Selection of Discrete Switches”, US Air Force report AL-TR-1991-0015, NTIS: AD-A273 019, 1991.

27. Borah, J., “Investigation of Eye and Head Controlled Cursor Positioning Techniques”, US Air Force report AL/CF-SR-1995-0018, September 1995.

28. Calhoun, G. L., Janson, W. P. and Arbak, C. J., “Use of eye control to select switches”, in “Proceedings of the Human Factors Society - 30th Annual Meeting”, 1986, pp 154-158.

29. Calhoun, G. L. and Janson, W. P., “Eye control interface considerations for aircrew station design”, in “Sixth European Conference on Eye Movements”, Leuven, Belgium, 1991.

30. Ware, C. and Mikaelian, H. T., “An evaluation of an eye tracker as a device for computer input”, in by Carroll, J.M. and Tanner, P. P., (Eds) “Proceedings of Human Factors in Computing Systems and Graphics Interface Conference”, Toronto, Canada. 1987, pp. 183-188.

31. Hatfield, F., Jenkins, E. and Jennings, M. W., “Eye/Voice Mission Planning Interface (EVMPI)”, US Air Force report TR-J103-1, 1995.

32. Glenn, F. A. Harrington, N., Iavecchia, H. P., and Stokes, J., “An Oculometer and Automated Speech Interface System”, in “Analytics, Technical Report 1920”, Analytics, Willow Grove, PA, May 1984.

33. Bolt, R. A., “The Human Interface: Where People and Computers Meet”, Lifetime Learning Publications, London, UK. 1984.

34. Koons, D. B., Sparrell, C. J. and Thorisson, K. R., “Integrating Simultaneous Input from Speech, Gaze, and Hand Gestures,” in Maybury, M. T. (Ed.) “Intelligent Multimedia Interfaces”, Menlo Park, AAI Press/The MIT Press, 1993.