1. **Eye Tracking Techniques.**

As previously mentioned, there are different eye tracking techniques that cater to different system implementation needs. The three predominant techniques according to [Majaranta and Bulling (2014)](#_ENREF_4) are:

* **Video-oculography (VOG)**

This technique involves the use of video based images of the eyes. The location of the eye and the center of the pupil are detected and this information is used for estimating the gaze – point of the user on the screen. The configuration (video camera) may be head mounted or remote.

This technique suffers from a) Sensitivity to head movements due to the pupil being the only reference to calculate gaze-point. b) Sensitivity to light, droopy eyelids, squinting of the eyes, etc. c) the camera requires an unobstructed view of the eye.

Eye trackers employing this type of eye-tracking technique do not provide accurate gaze-point estimation but is a viable option for systems that require only ‘gaze-awareness’ where accuracy is not as important.

* **Video based Infrared – Pupil Corneal Reflection (IR-PCR)**

This technique is similar to VOG (i.e., uses a video feed of the eyes) but is more accurate and much more robust to head movement. This is achieved by adding another reference point, the ‘corneal-reflection’ by using an Infrared light on- or off- axis to the eye. The position of this reflection remains constant and so by measuring the change in the location of the pupil relative to this reflection, the gaze-direction can be determined.

One drawback of this technique is that it is sensitive to ambient light and so it does not work well in outdoor settings.

Most current remote eye tracking hardware such as the Pupil Pro, Tobii EyeX, Tobii Pro Glasses, etc. employ this technique.

* **Electro-oculography (EOG)**

This technique is based on the fact that the human eye can be modeled as a dipole with the positive pole at the cornea and the negative pole at the retina([Majaranta and Bulling, 2014](#_ENREF_4)).

The head mounted gear consists of two surface electrodes placed in periorbital positions around the eyes. When a user’s eye moves, there is a change in the dipole orientation that causes a change in the electric potential field which is measured to track the relative eye movement of the user.

This technique is not sensitive to head movements or ambient light but it is not effective when accurate gaze-point estimation is required and is only applicable when relative eye movement is required.

**Comparison of eye tracking techniques.**

|  |  |  |  |
| --- | --- | --- | --- |
|  | VOG | IR-PCR | EOG |
| Accuracy | Limited | High | No |
| Sensitivity to head movement | Yes | Highly resilient | No |
| Sensitivity to ambient light | Yes | Yes | No |

*Table 1: Comparison of the different eye tracking techniques.*

1. **Gaze as an independent input modality**

Gaze can be used independently as a means of interaction. This method requires the user to deliberately perform actions with their eyes that are considered ‘not normal’. For example, normal fixations last for about 200 – 600ms ([Majaranta and Bulling, 2014](#_ENREF_4)), So, for a system to use fixations (or dwell time) as an interaction method, the fixation must be sufficiently long so as to avoid the Midas Touch problem (i.e., greater than normal fixations) but not too long so as to exhaust the user. There are two predominant eye movements around which gaze as an independent input modality is based on. They are discussed below:

* 1. **Fixations or Dwell Time**

Gaze typing is a popular service that employs dwell time to type. In gaze typing, a Keyboard is displayed on the screen, users point at the keys they want to type using gaze and fixates on the key until it is typed. According to [Majaranta et al. (2009)](#_ENREF_3), there was an increase in the typing speed from 6.9 wpm (words-per-minute) to 19.9 wpm when the user was given control of the dwell time. However, the use of dwell time is usually slower than confirming selection using other inputs such as touch and it can be tiring for the user.

In the above mentioned implementation, an accurate gaze-point of the user is needed so as to map the gaze-point to a particular key on the virtual keyboard. Hence, an Eye tracker employing IR-PCR technique would be appropriate and an initial calibration of the eye tracker for each user would be necessary.

* 1. **Smooth Pursuit**

Smooth pursuit is another popular method employed to capture intended interaction using gaze. It is not sensitive to accidental activation as this type of eye movement cannot be faked.

In a smart environment setting, [Velloso et al. (2016)](#_ENREF_11) demonstrated how a user could activate ambient objects that displayed a moving target by using smooth pursuit. Each object exposes a moving target that is either virtual or mechanical and displayed either internally or externally. Following these targets activates the object associated with it. A pupil pro eye tracker was used and its world view camera was modified so that it could capture infrared light. The objects would also give off infrared light so the system could detect the approximate direction the user was looking at by using the modified world view camera.

Smooth pursuit interaction was also used with smart watches([Esteves et al., 2015](#_ENREF_1)) whereby controls such as volume could be adjusted by following an orbiting dot around the control. In this method, the eye tracker was calibrated to the smart watch screen; fixation on a particular control would activate a point that would orbit around the control and smooth pursuit would be used to confirm selection.

Both the above implementations show that using gaze as an independent input modality is feasible and not sensitive to accidental activation. They both use relative eye tracking to map to the trajectory of the moving target. Due to this, each moving target must be different in either phase, speed or trajectory shape([Velloso et al., 2016](#_ENREF_11)). This limits the maximum number of targets but is more comfortable to use than dwell time as eyes are naturally drawn towards moving targets. Since only relative movement of the eyes are used, calibration of the eye tracker is not necessary. Also, it is easier for the above systems to work with devices that have a display; objects without a display need external projections or mechanical enhancements to display moving targets which would require additional setups.

The use of relative eye tracking gives the developer an option to choose between any of the eye tracking techniques. VOG and IR-PCR can be used to map consecutive coordinates into a trajectory while EOG can be used directly to obtain relative eye movement information.

1. **Gaze as a complementary input modality**

Gaze has enormous potential when used to support other input modalities such as touch, pen, etc. We use our visual perception before performing various task in our daily lives, this indicates that gaze is naturally complementary to other means of interaction and this observation can be used as an advantage when designing new interaction systems.

To select distant objects on a multi touch enabled tabletop, [Mauderer et al. (2013)](#_ENREF_5) combines gaze and flicking gesture (Throwable object). Two hypothetical lines are drawn, one from the flicking gesture and one from the user’s gaze that intersect at the target object. They observed that some users would divert their gaze to the throwable object at the time of flicking causing an error in selection. Another problem with this approach is that only a single user can use the tabletop at a time. They propose using multiple eye trackers to track many users but no concrete evidence was provided that this approach was feasible.

[Pfeuffer et al. (2015)](#_ENREF_7) explored direct and indirect input using gaze and touch/pen to provide functionality such as Drawing, Tracing of lines, etc. A switching mechanism was used to switch between direct and indirect input mode. Redirecting input to the user’s gaze location attempts to solve the issues of occlusion of the target, unreachable targets, etc. A threshold (T) is used to determine whether input was direct or indirect i.e., if input falls under a gaze of radius T, direct input takes place and indirect input takes place when input falls outside this gaze area. Since, no form of dwell time or smooth pursuit was used to confirm object selection, accidental activation of objects may occur. This work provides us with two interesting insights. Firstly, our area of attention within our field of vision is very small and eye trackers today are not able to determine precisely where we are looking at. Secondly, using pen or touch along with gaze enables the user to manipulate objects indirectly with a fair degree of precision.

In another implementation, gaze was used to select target and touch was used to indirectly manipulate it([Pfeuffer et al., 2014](#_ENREF_6)). This method is useful when the user’s hands should not obstruct the users view. [Pfeuffer et al. (2016)](#_ENREF_8) also investigated and compared the former mentioned indirect gaze-touch method to a pen-touch method. They concluded that users were more accurate using gaze-touch method than direct touch or the pen-touch method for the tests that they set up. The comparison does bring to light the fact that gaze and touch can be used as a good alternative for direct touch method.

All of the above gaze supported techniques use accurate gaze-point and therefore would require eye trackers employing IR-PCR techniques and calibration of the eye tracker prior to use. Calibration may be unfavourable and time consuming especially when there is more than one user.

An Interesting application of gaze comes in the form of “gaze – aware” systems. These systems do not require high accuracy from eye trackers and are only interested in the general area of focus of the user. For example, A smart television that turns on display only if the user is looking at the screen.

[Voelker et al. (2015)](#_ENREF_12) observed that Horizontal touch surfaces were much easier to work with than the vertical touch surfaces in an interactive desktop workspace. Here, both surfaces were also used as displays but only the horizontal surface was used for input, Gaze was used to switch between direct input to the horizontal display and indirect input to the vertical display. They concluded that users were more efficient using a relative mapping between the horizontal and vertical touch surfaces as opposed to absolute mapping. However, their tests were specific (Tapping, Dragging and Cross Dragging) and does not support if the method is applicable for other purposes, say, a user who is not familiar with the QWERTY keyboard may look down occasionally to the horizontal surface and accidentally change the active surface.

This implementation uses an Eye-tracker employing an IR-PCR technique namely the Dikablis Glasses. The choice of IR-PCR over VOG is because the system does require accurate gaze-estimation when selecting object using touch when working indirectly with the vertical display. If no selection was required, VOG could be used to just determine which screen was currently active.

1. **Gaze in a Social/Attentive context**

Eyes are subtle yet powerful tools of communication that we use in our daily lives without much thought. They can give use cues as to whether we are curious or attracted to what we see, for example, pupil dilation can signify attraction while pupil contraction can indicate anger or dislike.

([Hoppe et al., 2015](#_ENREF_2)) used curiousity questionnaires and eye movements to determine the curiousity of a user in objects during a shopping scenario.

[Sugano et al. (2016)](#_ENREF_9) proposed AggreGaze; a calibration free method to estimate the collective visual attention of a group on a public display. This method uses Appearance-based([Tan et al., 2002](#_ENREF_10)) gaze estimation and provides error correction for a wide range of illuminations and viewing angles using on-site training data collection. By using the aggregate, uncalibrated gaze estimates of a group of users, the system can compute the attention distribution across the display.

**References**

ESTEVES, A., VELLOSO, E., BULLING, A. & GELLERSEN, H. 2015. Orbits: Gaze Interaction for Smart Watches using Smooth Pursuit Eye Movements. *Proceedings of the 28th Annual ACM Symposium on User Interface Software &#38; Technology.* Daegu, Kyungpook, Republic of Korea: ACM.

HOPPE, S., LOETSCHER, T., MOREY, S. & BULLING, A. 2015. Recognition of curiosity using eye movement analysis. *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers.* Osaka, Japan: ACM.

MAJARANTA, P., AHOLA, U.-K. & ŠPAKOV, O. Fast gaze typing with an adjustable dwell time. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2009 Boston, MA, USA. ACM, 357-360.

MAJARANTA, P. & BULLING, A. 2014. Eye tracking and eye-based human–computer interaction. *Advances in physiological computing.* London: Springer.

MAUDERER, M., DAIBER, F. & KRÜGER, A. Combining touch and gaze for distant selection in a tabletop setting. Paper presented at CHI’13 Workshop: Gaze Interaction in the Post-WIMP World, 2013/4/27 2013 Paris, France.

PFEUFFER, K., ALEXANDER, J., CHONG, M. K. & GELLERSEN, H. Gaze-touch: combining gaze with multi-touch for interaction on the same surface. Proceedings of the 27th annual ACM symposium on User interface software and technology, 2014 Honolulu, Hawaii, USA. ACM, 509-518.

PFEUFFER, K., ALEXANDER, J., CHONG, M. K., ZHANG, Y. & GELLERSEN, H. Gaze-shifting: Direct-indirect input with pen and touch modulated by gaze. Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology, 2015 Daegu, Kyungpook, Republic of Korea. ACM, 373-383.

PFEUFFER, K., ALEXANDER, J. & GELLERSEN, H. Partially-indirect Bimanual Input with Gaze, Pen, and Touch for Pan, Zoom, and Ink Interaction. Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 2016 Santa Clara, California, USA. ACM, 2845-2856.

SUGANO, Y., ZHANG, X. & BULLING, A. AggreGaze: Collective Estimation of Audience Attention on Public Displays. Proceedings of the 29th Annual Symposium on User Interface Software and Technology, 2016. ACM, 821-831.

TAN, K.-H., KRIEGMAN, D. J. & AHUJA, N. Appearance-based eye gaze estimation. Applications of Computer Vision, 2002.(WACV 2002). Proceedings. Sixth IEEE Workshop on, 2002. IEEE, 191-195.

VELLOSO, E., WIRTH, M., WEICHEL, C., ESTEVES, A. & GELLERSEN, H. AmbiGaze: Direct Control of Ambient Devices by Gaze. Proceedings of the 2016 ACM Conference on Designing Interactive Systems, 2016 Brisbane, QLD, Australia. ACM, 812-817.

VOELKER, S., MATVIIENKO, A., SCH, J., NING & BORCHERS, J. 2015. Combining Direct and Indirect Touch Input for Interactive Workspaces using Gaze Input. *Proceedings of the 3rd ACM Symposium on Spatial User Interaction.* Los Angeles, California, USA: ACM.