Human-Computer Through Eye Tracking: Current Status and Future Prospects

Nimmy Prakash
Izone Infosoft solution
I T Company
Irinjalakkuda

Abstract—Eye-movement tracking is a model that is increasingly being employed to study usability issues in HCL contexts. The objectives are First, we introduce the reader to the basics of eye-movement technology, and also present the key aspects of practical guidance to those who might be interested in using eye tracking in HCL research, whether in usability-evaluation studies, or for capturing peoples eye movement as an input mechanism to drive system interaction. Second ,we examine various ways in which eye movements can be systematically measured to examine interface usability .Third ,the various opportunities for eye-movement study in future HCL research, and detail some of challenges that need to be overcome to enable the effective application of the technique in studying the complexities of advanced interactive system use. The aim of ongoing research is to develop an application to replace a computer mouse for a people with physical impairment. The physically impaired users can't handle the traditional input devices such as mouse, keyboard etc the alternate for this category of users must be available .Speech is another promising technology to achieve this. The first approach researches estimation of eye gaze point as pointing device .The second approach researches for the combination of both.The application is based on an eye gaze estimation algorithm and assumes that the camera and the head position are fixed.The system after successful development will able to interact user with specific application.

Index Terms— Eye-tracking, ROI ,speech recognition, eye gaze-point

I. INTRODUCTION

Eye tracking is a technique whereby an individual's eye movement are measured so that the researcher knows both where a person is looking at any given time and the sequence in which their eyes are shifting from one location to another.Tracking people's eye movement can help HCI researches understand visual and display-based information processing and the factors that may impact upon the usability of system interfaces.In this way,eye-movement recordings can give an objective sources of interface-evaluation data that can inform the design of improved interfaces. Eye movements can also be captured and used as control signals to enable people to interact with interfaces directly without the need for mouse or keyboard input, which can be a major advantage for certain populations of users such as disable individuals.Insearching for betterinterfacesbetweenusers and their computers,an additional mode of communications between the two parties would be of great use. The problem of human computer interaction can be viewed as two powerful information processors(human and computer) attempting to communicate with each other via a narrow bandwidth, highly constrained interface faster, more natural, more convenient(and, particularly, more parallel, less sequential) means for users and computers to exchange information are needed to increase the useful bandwidth across that interfaces. On the user's side, the constraints are in the nature of the communication organs and abilities with which humans are endowed; on the computer side, the only constrained is the range of devices and interaction techniques that we can invent and their performance.

II. METHODOLOGY

A. Design and Implementation of Eye Tracking Modules

There aretwo type of eye tracking systems head mounted and remote.The head mounted eye trackers commonly use reflected light to track the eyes .Camera suspended from a contraption mounted on the head capture video of the eye.The eye position is determined by shining a light source at the eyeball and measuring the distance between the light reflection and a feature of the eye (eg.pupil). Head mounted trackers are accurate but can be intrusive and feel unnatural to wear, although component miniaturization has extended their utility .Remote eye trackers use multiple fixed cameras in the environment fixating on the face.The remote cameras are responsive enough to capture images of the eyes and head position .Advances in image processing and computational power have made remote tracking more popular,although the quality of the tracking is inferior to head-mounted tracking.Real time eye tracking is achieved with various steps are follows

- Step 1 –Image Acquisition: the eye image was acquired through the iball night vision web camera.
- Step 2 –Image Pre-processing: after image acquisition pre-processing is required to convert the acquired image into grayscale and further binary image.
- Step 3 –Eye and PupilDetection: the upper and lower threshold was applied to extract the pupil from the eye image.
- Step 4 –Pupil Localization: the new Region of interest (ROI) is considered for the future

Special Issue - 2015

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NSRCL-2015 Conference Proceedings

processing to detect location of the pupil in real time.
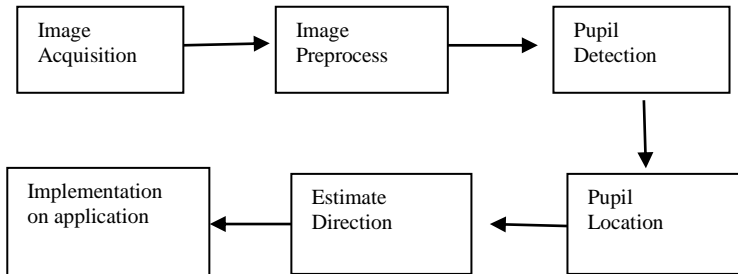
- Step 5 –Estimation of Direction:

Figure A.1: Eye tracking module

The current direction of the tracking module is estimated on the basis of the location of the pupil at theparticular place in the ROI of the eye image.

B. Speech Recognition Module

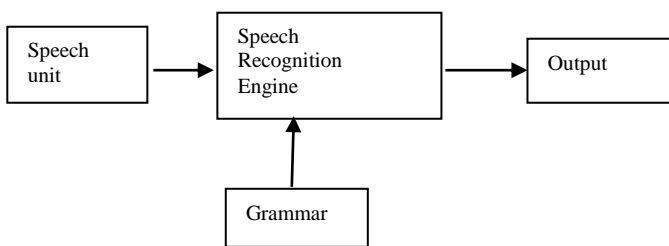

Figure b.1: Speech Recognition module

Speech is natural mode of communication for people.The accuracy of any system – can vary along the following dimensions:

- Vocabulary size and confusability: As a general rule ,it is easy to discriminate among a small set of words, but error rates naturally increase as the vocabulary size grows.for example ,the 10 digit "zero" to "nine" can be recognized essentially perfectly ,but vocabulary sizes of 200,5000 or 100000 may have error rates of 3%,7% or 45%.
- Speaker dependence vs. independence:Speaker independence is difficult to achieve because a systems parameters become turned to speakers that it was trained on, and those parameters tend to be highly speaker-specific.
- Isolated, discontinuous or continuous speech: Isolated speech means single words and discontinuous speech means full sentences in words are artificially separated by silence and continuous speech means naturally spoken sentences.
- Read vs. spontaneous speech: Systems can be evaluated on speech that is either read from

prepared scripts, or speech that is spoken spontaneously. Spontaneous speech is vastly more difficult, because it tends to be peppered with disfluencies like "um" and "uh",false starts, incompletesentences, stuttering, coughing etc.

- Adverse conditions: It include environmental noise (noise in a motor); acoustical distortions(eg. echoes);different microphone (tele phone) etc.

C. Fusion of Eye Gaze Point and Speech Recognition

Fusion is the process of joining two or more things together to form a single entity. It started as follows

- Levels of Fusion: The most commonly used strategy to follow is to fuse the information at the feature level,which is non as early fusion.The other approach is decision level fusion or late fusion which fuses multiple modalities in the semantic space.A combination of these approach is also practiced as the hybrid fusion approach.
- How to Fuse: There are several methods that are used in fusing different modalities.These methods are particularly suitable under different settings.The discussion also include how the fusion process utilizes the feature and decision level correlation among the modalities,and how the contextual and the confidence information influences the overall fusion process.
- When to Fuse: The time when the fusion should take place is an important consideration in the multimodal fusion process.Certain characteristics of media,such as varying data capture rates and processing time of media, poses challenges on how to synchronize the overall process of fusion. Due to asynchrony and diversity among streams and due to the fact that different analysis tasks are performed at different granularity levels in time, the identification of these designed points ie,when the fusion should take place is a challenging issue.
- What to Fuse: The different modalities used in a fusion process may provide complementary or contradictory information and therefore knowing which modalities are contributingtowards accomplishing an analysis task needs to be understood. This is also related to finding the optimal number of media streams [16] or feature sets required to accomplish an analysis task under the specified constraints.If the most suitable subset is unavailable,can one use alternate streams without much loss of cost-effectiveness and confidence?

D. Neural Network for Fusion

We are using neural networkapproach for datalevel fusionof eye gaze point and speech. We are use both AND/OR situation for the fusion. AND type can be used in a situation where both the input must be ture viz . applications used for security purposes . OR type fusion can be used to interact

disable users with the computer system. The users can work the application through eye or speech. The perceptron model used for the fusion of eye gaze point and speech.

III. APPLICATION

A. *Experimental Setup*

The experimental modal arranged through proposed model.

B. *Image Acquisition*

Eye image is captured by the modified web camera(iBall night vision webcam).The web camera is fixed on the cap.

C. *Eye Tracking*

When pupil is detected the tracking is monitored and estimated by Gaze point estimation algorithm.In the proposed system voting scheme is used to estimate the localization of the pupil.The ROI (Region of Interest) is considered for the estimation of the localization of the pupil.[8][9][11] The Gaze directions such as Left,Right,Up and Down is determined on the basis of the change in point of the pupil in real time.The capture model is responsible for providing an image of the eye to the computer.it was created with iBall night vision webcam.The software uses algorithm based on the image obtained in the infrared light.Available webcams work in the visible spectrum.It is necessary to modify the camera and mount a suitable filter that allows capturing images in infrared light.The first step in modify a webcam is complete disassembly of the outer casing.At the back of the camera is a screw,unscrew it and then release normal LEDs.The key operation of the capture module is to move images took in the infrared to computer. In the camera lens should be installed filter that stops the rays of visible light and transmit infrared rays.It allows web cam to capture images in a way similar to the human eye. The filter was detached by undermining its banks by thin knife.In place of it is added infrared filterProfessional IR filters are relatively expensive. Foundation to create head mounted eye tracking system to gaze tracking was minimize the cost according to needs.Infrared filter in construction of the capture module was created from negative film.

IV. CONCLUSION

Our contention is that eye-movement tracking represents an important, objective technique that can afford useful advantages for the in-depth analysis of interface usability.Eye-tracking studies in HCI are beginning to burgeon,and the technique seems set to become an established addition to the current battery of usability-testing methods employed by commercial and academic HCI researchers.The continued growth in the use of the method in HCI studies looks likely to continue as the technology becomes increasingly more affordable,less invasive and easier to use .The future seems rich for eye tracking   HCI.The paper demonstrate the various techniques used for implementation of the fusion of Eye gaze point and speech in real time.The required hardware and modification in the camera according to the requirement for the pupil finding and localization. Eye tracking technology also needs to be improved to increase the validity and reliability of the recorded data. Eye tracking systems need to become cheaper in order to make them a viable usability tool for smaller commercial agencies and research labs.Once eye tracking achieve the improvements in the technology ,methodology,and cost it can eye tracker as an input device is far from "perfect " ,in the sense that a mouse or keyboard is,and that is caused both by the limitations of current equipment and more importantly by the nature of human eye movements.

REFERENCES

[1] R.J.K.Jacob,"Human-computer Interaction ," pp.383-388 in Encyclopedia of Artificial Intelligence, ed. S.C. Shapiro, John Wiley,New York(1987).

[2] M.A. Just and P.A. Carpenter, "A Theory of Reading : From Eye Fixations to Comprehension, "Psychological Review 87 (4) pp. 329-354 (1980).

[3] J.L.Levine,"An Eye-Controlled Computer," Research Report RC-8857,IBM Thomas J.Waston Research Center ,Yorktown Heights, N.Y.(1981).

[4] J.L. Levine , "Performance of an Eyetracker for Office Use," Comp C.Schmandt,N.S.Ackerman, and D.Hindus, "Augmenting a Window System With Speech Input," IEEE Computer 23(8) pp.50-56 (1990).23.

[5] Starker and R.A. Bolt , "A Gaze-Responsiv e Self-Disclosing Display," Proc. ACMCHI'90 Human Factors in Computing Systems Conference pp.3-9 Addison –Wesley/ACM Press(1990).

[6] H.M. Tong and R.A. Fisher,"Progress Report on an Eye-Slaved Area-of-Interest Visual Display," Report No. AFHRL-TR-84-36, Air Force Human Resources Laboratory

[7] Albert,W. (2002). Do web users actually look at ads? A case study of banner ads and eye-tracking technology. In Proceedings of the Eleventh Annual Conference of the Usability Professionals' Association.

[8] Altonen,A.,Hyrskykari,A.,&Raiha,k.(1998).101 Spots,or How do users read menus? In proceedings of CHI'98 Human Factors in Computing Systems(pp.132-139). NY:ACM Press