EYE-CENTRIC ICT CONTROL

Fangmin Shi, Alastair Gale and Kevin Purdy

Applied Vision Research Centre
Loughborough University
Loughborough LE11 3UZ

There are many ways of interfacing with ICT devices but where the end user is mobility restricted then the interface designer has to become much more innovative. One approach is to employ the user's eye movements to initiate control operations but this has well known problems of the measured eye gaze location not always corresponding to the user's actual visual attention. We have developed a methodology which overcomes these problems. The user's environment is imaged continuously and interrogated, using SIFT image analysis algorithms, for the presence of known ICT devices. The locations of these ICT devices are then related mathematically to the measured eye gaze location of a user. The technical development of the approach and its current status are described.

Introduction

The availability of low cost systems which measure eye gaze behaviour has led to an increasing number of viable opportunities for utilising eye movement recording as a tool for interacting with the environment (see for instance, Istance and Howarth, 1994).

Similarly, there are an increasing number of ICT and electrically operated devices in our environment that have the potential to be operated remotely, as well as add-on products that are available for automating typical manually operated items. For the purposes of this paper all such items are simply referred to as 'objects' or 'controllable devices'.

Using eye gaze information, people can achieve efficient interaction with their surroundings (e.g. Shell et al., 2003). For instance, various commercial eye-typing systems are available, to help disabled people interact with a computer and users can be trained to achieve fast typing speed by selecting soft keys displayed on the PC screen (c.f. Ward and MacKay, 2002).

However, using eye gaze as a selection device can be problematic as sometimes what people look at is not what they are actually attending to and the issue of attention is ever intriguing (for instance, see Wood et al., 2005). Such an involuntary selection can result in a direct operation of

the target object, leading to unpredictable false reactions. We are currently investigating a selective attention control system, using eye gaze behaviour, which overcomes such an issue. It does this initially by supplying a Graphical User Interface, which allows people to confirm their intentions consciously, so that any control executions are actually based on their needs. This system is known as Attention Responsive Technology – ART (Gale, 2004).

Eye-centric control system
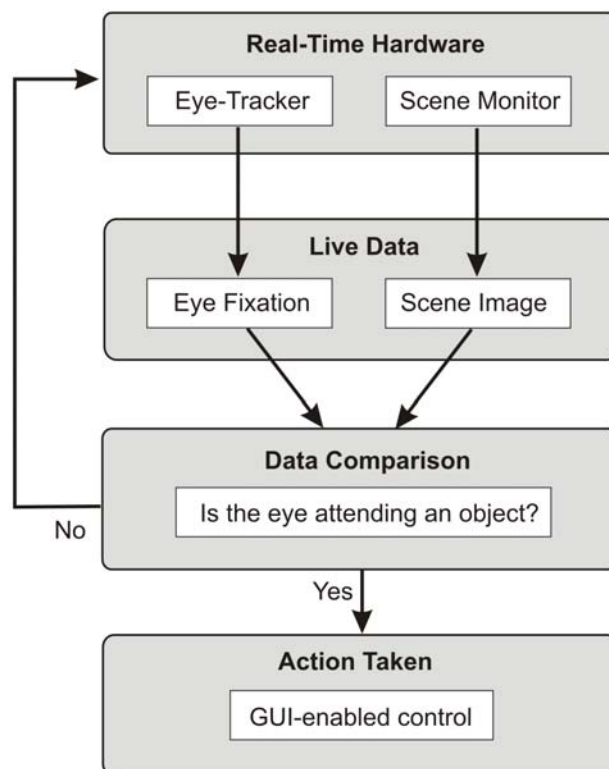
System integration
A laboratory-based prototype system has been set up to carry out the ICT device control using eye point of gaze. It consists of four main components:

 An eye tracker – to record users' eye movements and monitor eye fixations on objects in the environment
 An object monitor – to observe the user's environment with a view to identifying any object within the user's field of attention
 A user-configurable panel – to provide a GUI for the user to confirm his/her attention and initiate control of an object
 A controller – to enable the actual control of an ICT device upon any PC-based command.

Figure 1 System work process illustration

The integrated system runs under our specially developed software, which receives raw data from the first two units and issues commands to the last two units after performing extensive

computational process. The work flow is illustrated in Figure 1.

*Mini cameras for eye tracking and object monitoring*

The system development starts with a commercially available head-mounted eye tracking unit (ASL 501 system). This contains two mini compact video cameras. One is the eye camera, which records the eye pupil and corneal reflection and outputs the eye's line of gaze with respect to the head mounted system. The other is the scene camera which has a wide field of view lens. This is mounted facing the environment in front of the user and monitors the frontal scene.

These two cameras are linked together after being calibrated. The user's point of gaze can then be directly mapped to its corresponding position in the scene image.

System calibration

The eye monitoring system needs to be calibrated for each user. This process only takes a short time. Calibration entails the user sequentially looking at a matrix of nine points as shown in the left image of Figure 2. The eye fixation data for each point are recorded with reference to the eye camera system as illustrated at the bottom middle frame of Figure 2. At the same time, the image coordinates of the nine points with reference to the scene camera system are extracted by our image processing algorithm, which are known as target points and highlighted as crosses at the upper middle image of Figure 2. Through comparing these two sets of coordinates, a point of gaze recorded by the eye camera can be directed to the falling point on the scene image simultaneously.

Fixations can then be traced. Whether they fall on any target object is dependent on whether the scene image contains any recognisable object of interest on that point.
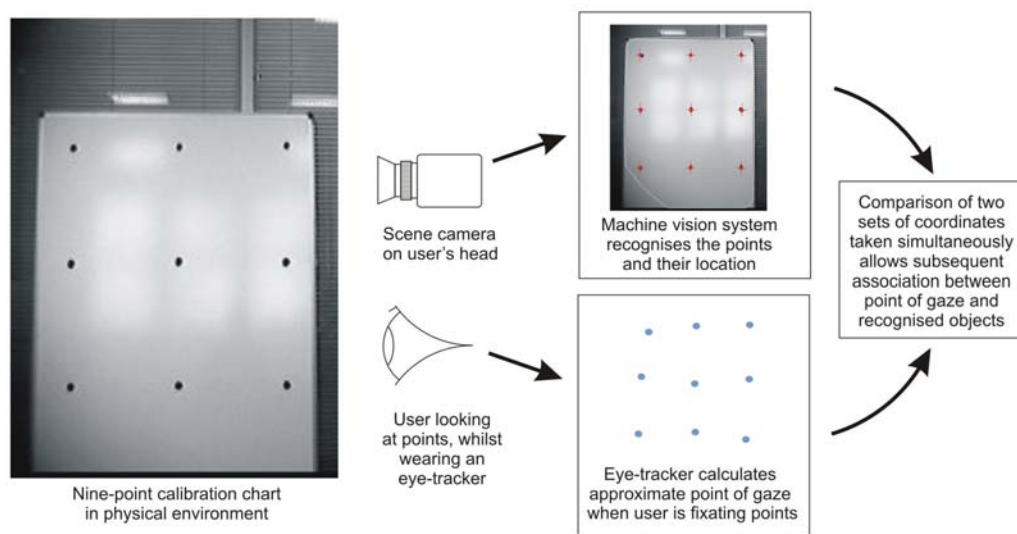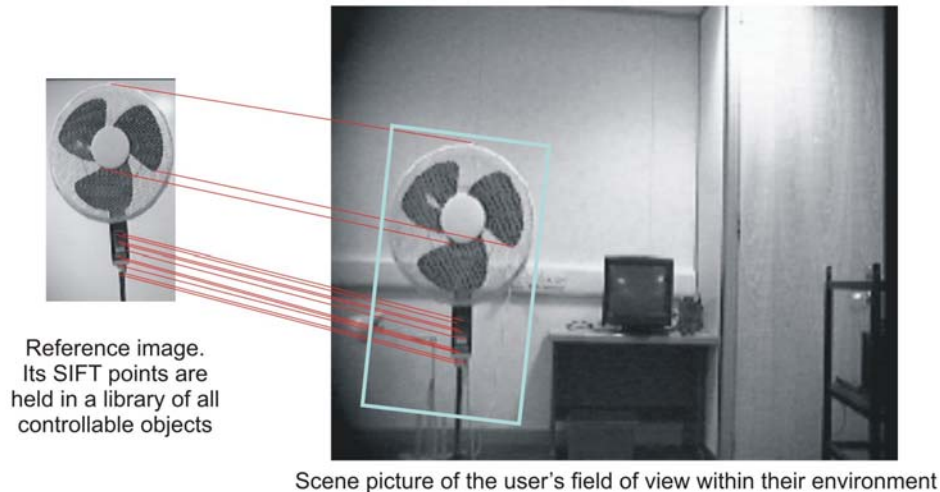

Figure 2 System calibration process

Object identification

The above approach makes the complexity of 3D object recognition and location reduced to 2D object recognition only. Algorithms can then be developed and applied to the scene camera output to try to recognise objects in the scene. An efficient and reliable object identification method is under development in the research project. This performs image feature matching between a scene image and an image database that collects images of target objects. The image feature detection algorithm is based on the SIFT- Scale Invariant Feature Transformation, approach

proposed by Lowe (2004). SIFT features are adopted because they have advantages over other existing feature detection methods in that their local features provide robustness to change of scale and rotation and partial change of distortion and illumination. An example showing the SIFT matching result for identifying an electric fan in a scene is given in Figure 3. At the right of Figure 3 is the image, which contains an electric fan, taken by the scene camera. The image of the electric fan to the left of the figure is one of the reference images in the database. The lines between the two images indicate the points of matching between the reference and the real object images and shows how an environmental object is recognised.

Reference image. Its SIFT points are held in a library of all controllable objects

Scene picture of the user's field of view within their environment

Figure 3 An example image showing the SIFT matching result

GUI-enabled ICT control
Having recognised an object in the scene, then the system needs to allow the user to choose whether to operate it. Assume a user keeps looking at a target object for a certain period of time, e.g. 0.5 seconds, and then a GUI will be enabled. The concept is that the interface will 'ask' the user to confirm whether or not it is his/her intention to operate the object. To conduct the control of the devices, for example, to switch on/off a TV, our system employs the X10 control protocol popularly used in the home automation community. A control command is issued from the PC via an X10 adaptor. Any object to be controlled is connected to the normal mains electrical supply by plugging into an X10 module first. There are no wires required for connecting between the objects and the PC - interaction is through wireless communication or so-called X10 signals.

Conclusions and future work

A system is described which enables a user to select and control ICT objects by using their eye gaze behaviour. The system entails using a head mounted eye movement recording device. Currently, the overall framework of the ART system has been achieved. ICT objects can be identified in the user's environment by the algorithms designed and built into the ART system. Additionally we have interfaced the eye tracking system to the object monitoring and identification system. At present the ART system components work separately and current research effort is focussed on integrating the separate modules into a fully cohesive ART system which will work in real time.

References

Gale A.G., 2005, Attention responsive technology and ergonomics. In Bust P.D. & McCabe P.T. (Eds.) Contemporary Ergonomics, London, Taylor and Francis, 273-276

Istance, H. and Howarth P, 1994, Keeping an eye on your interface: the potential for eye-gaze control of graphical user interfaces, Proceedings of HCI'94, 195-209

Lowe D.G. 2004, Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*. **2**, 91-110

Shell J.S., Vertegaal R. and Skaburskis A.W. EyePliances: Attention-Seeking Devices that Respond to Visual Attention, CHI 2003: New Horizons, 771-772

Ward D.J. and MacKay D.J.C., 2002, Fast hands free writing by gaze direction. *Nature*, **418**

Wood, S, Cox, R and Cheng, PCH (2006) Attention Design: Eight issues to consider. Computers in Human Behavior. Special issue on Attention-aware systems