GESTURE RECOGNITION FOR AUGMENTATIVE HUMAN COMPUTER INTERACTION

Richard Reilly, *member IEEE* and Andrea Hanson
Dept. of Electronic and Electrical Engineering, University College Dublin, Belfield, Dublin 4, Ireland.

I Abstract

Augmentative and alternative communication has always suffered from difficulties of equipment set-up and interaction. A three-dimensional input device is presented which is based on a non-contact movement sensor. In this way the movement of the head or limb can be tracked and converted into the movement of a PC mouse. The advantage of the non-contact approach is further increased by the incorporation of adaptive gesture recognition, thus allowing user defined movement interpretation and click strategies.

II Introduction

The majority of research attention to date in the area of gesture recognition, has focused on the use of three-dimensional tracking of the body or limb [1],[2],[3]. These methods involve the use of specific input devices [4],[5],[6],[7] which attempt to provide complex information on the motion of or movement made by an individual.

It is important for any motion detection system to be adaptive, preserving the natural posture of the user, and must be fully configurable for those with limited or low amplitude movement. Coherent light falling on a mobile surface produces a diffraction pattern or speckle [8]. Through analysis of this diffraction pattern, the displacement of the surface can be evaluated to high level of accuracy. A motion sensor employing this principle when coupled with movement interpretation and gesture recognition, was found to provide a very effective interface for augmentative communication.

III Non-Contact Movement Sensor

The input sensing device functions on the principle that a coherent beam (laser) aimed at a diffusing object creates a speckle pattern of light, the diffusing object's displacement causes a proportional displacement of the pattern, which can be measured using a multi-element light sensor.

A fully operational demonstrator has been realised, incorporating the use of a laser diode as emitter and a CCD linear array, with collection optics, as detector. In such as configuration the skin of the user acts as a diffuser which back-scatters a speckle pattern so that movements of the skin directly drive the mouse cursor in two dimensions. The device also provides a third dimension which is required for movement validation and gesture recognition. The reflected signals are processed and adaptively optimised to the sounding ambient light.

IV Movement Interpretation

Classification of human movement is a complex task. One element of this is to pre-process and interpret movement. In the interface developed, signals from the CCD detector, were adaptively processed and these movements combined into x, y co-ordinates of a PC mouse. This was accomplished in hardware and firmware, with smoothing filters being required to reduce tremor and noise. A simple moving average filter and a high order Chebychev filter have been incorporated. Moving average lowpass filters are typical of input devices [9][10]. The filter employed a low-pass with a cut-off of 4Hz, as it was noticed that tremor in the hand, was in the order of 5-7Hz.

Certain features have been included within the interface to ensure maximum usability. The ability to use a standard PC mouse as input device is possible with all commercial augmentative and alternative communication software systems. The ability to replace the standard PC mouse with the developed interface was felt to be a central feature of usability. Coupled with this fact, recent advances in the features provided by mouse driver software, were deemed also to be necessary to conserve for maximum usability. To accomplish this, a terminate-stay resident program (TSR) was written to monitor and provide communication between the application, mouse driver and input sensing device.

As mentioned above, posture is an important issue in the design of input devices. It can happen that the user may slump in their seating position while using interface devices, preventing access to certain areas of the PC application screen. One feature included allows the mouse cursor to be centred by accelerating off the screen. The movement validation system within the device, recentres the mouse accordingly. Ballistic motion is also included that can allow movement in a direction by moving rapidly in that direction.

V Gesture Recognition

To allow effective human-computer interaction, gesture analysis must be incorporated within the interface. To approach this end, a neural based recognition system was utilised. Once the device no longer senses movement, the TSR launches into recognition phase, which involves a

0-7803-2475-7/97 $10.00

movement/feedback window displayed on top of the application screen. The mouse cursor is relocated within this window and the user is instructed to gesture. The output from the gesture recognition phase, provides mouse click information for the PC, by directly accessing the mouse driver via the TSR. The mouse trail is drawn within the recognition window to provide user feedback, see figure 1.

An intuitive gesture recognition system, analysing predominately x co-ordinate movement (analogous to head nod) or predominately y-co-ordinate movement (analogous to head shake) has also been implemented. This system, together with relative amplitude variation, have been employed as click strategies, with the device.

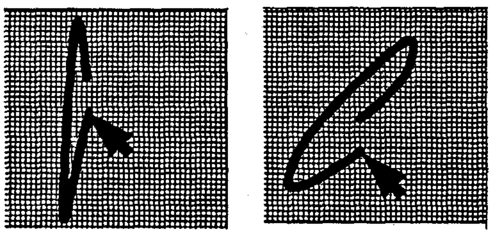

Figure 1: User Feedback within Gesture Recognition Window

A more complex recognition system, has been developed using a neural based recogniser. Although recurrent neural networks have been used successfully in gesture recognition for three-dimensional gestures [5], a three layer preceptron configuration, using an input feature vector generated by pre-processing the x & y co-ordinate data was employed. The pre-processing stage is composed of a normalisation and a segmentation phase, fragmenting the movement into a "gesture matrix". From this matrix a feature vector is generated with 25 elements.

The network was trained with training data using backpropagation techniques, classifying into four separate gestures. Training was carried out with twenty versions of the required gesture, nod or shake. The distinct gestures considered for the prototype system include those for the intuitive recognition system, together with a null set, i.e. no reflection of the laser and a saturated set, i.e. full reflection from the laser. Evaluation and testing has centred on improving the pre-processing stages thus allowing poorly performed gestures to be classified successfully.

To provide maximum usability an adaptive neural based recognition system has been investigated. This allows the user, with the aid of a Therapist, to train the device to specific needs. In this way the system will respond to the needs of users with limited upper body control.

VI Conclusion

An adaptive non-contact interface system, of interest to motor disabled persons suffering from quadriplegia and motor-cerebral injuries, has been developed. It can offer this group "mouse-like" access to computers and may prove beneficial in the field of education, in conjunction with the areas of communication, environmental control and occupational therapy.

VII Acknowledgement

This work was supported by the European Union under the TIDE programme. The authors acknowledge the work carried out by other members of the LAMP consortium.

VIII References

[1] Harwin W.S., Jackson R.D., *Computer Recognition of Head Gestures in Cerebral Palsy*", Proceedings of 13th Annual Conference of RESNA, pages 257-258, Washington, USA, 1990.

[2] Harwin W.S., Jackson R.D., *Analysis of Intentional Head Gestures to Assist Computer Access by Physically Disabled People*, Journal of Bio-Medical Engineering, Vol. 12, No.3, May 1990.

[3] Harwin W.S., Jackson R.D., *A Simple Mathematical Analysis of Head Movement*, Proceedings of ICAART, pages 386-387, Montreal, Canada, 1988

[4] Hanson P., Dobson D., Wanner J., "*Free Wheel : The Cordless New Headpointing Device*", Proceedings of ICAART, pages372-373, Montreal, Canada, 1988.

[5] Murakami K., Taguchi H., *Gesture Recognition using Recurrent Neural Networks*, Proceedings of CHI'91, New Orleans, Louisiana, USA, 27th April-2nd May 1991.

[6] Brickmann H., Boland F., Wright J., *A Gesture Input Interface Implemented Using Neural Networks*, Proceedings of the Irish DSP and Control Colloquium, Dublin, Ireland, July, 1994.

[7] Cairns A.Y., Newell A.F., *Towards Gesture Recognition for the Physically Impaired*, Proceedings of 17th Annual Conference of RESNA, pages 414-416, Nashville, USA, June 17-22, 1994.

[8] Dainty, J.C., *Laser Speckle and Related Phenomena*. Topics in Applied Optics. Springer-Verlag. 1975.

[9] Hanson P., Wanner J., "*Software Drivers for Pointers used by Persons with Disabilities*", Proceedings of 16th Annual Conference of RESNA, pages 443-444, June 12-17, 1993

[10] Mandler E., *Advanced Preprocessing Technique for On-line Recognition of Handprinted Symbols*, Computer Recognition and Human Production of Handwriting, R. Plamondon *et al* editors, Singapore World Scientific, 1989