Image processing used to harness blinking as a channel of communication and control for physically disabled people

D. Burke[1,2] T. Ward[1,3] A. de Paor[1,2]

[1]Rehabilitation Engineering, National Rehabilitation Hospital, Dun Laoghaire, Co. Dublin, Ireland
[2]National University of Ireland, Dublin, Ireland
[3]National University of Ireland, Maynooth, Ireland

Abstract—*Sophisticated mathematical algorithms (such as differencing, thresholding, aggregation and statistical analysis of skin colours) are used to compare successive frames of computer-captured images of the face. From these, changes in state of the eyes are determined and are used to detect blinks. A recognition performance of 83.74 ± 0.03% is achieved over five subjects with a low rate of false positives 2.71 ± 0.01%. A logical decision rule identifies purposeful blinks and applies them to control either a custom-designed communication package or an external device.*

Keywords—*Image processing, Blinking, Rehabilitation, Locked-in syndrome, Communication*

Med. Biol. Eng. Comput., 2001, **39**, 000–000

1 Introduction

FOR SOME people with very severe physical disabilities (e.g. locked-in syndrome), eye movements may represent the only feasible channel for communicating with others and for environmental control. There are effective non-contact systems available to detect eye gaze based on infrared, e.g. Eyegaze System, CLEVELAND *et al.* (1992). However, these systems require specialist equipment and suffer from the requirement that the user must calibrate the system, which necessitates some level of cognition to be first established with disabled patients. In contrast to marker-based systems (MIYAZAKI *et al.*, 2000), the philosophy adopted here was to develop an entirely non-contact, marker-free, and real-time solution, using an inexpensive digital camera ('a webcam'), a desktop PC (Pentium III, 500 MHz), and a judicious choice of sophisticated techniques from digital image processing. Digital image processing is a very wide field but for present purposes, can be regarded as a collection of mathematical algorithms applied to computer stored visual images to extract features, which can be used to accomplish some desirable end. By dispensing with the requirement for markers, the system presents itself as unobtrusive and simple to use to both therapist and patient alike. Based on a development of ideas described in CLARKE *et al.* (1998), the mathematical techniques are summarised in the next section. We present the results of a reliability measure of the system in the subsequent section, and the paper concludes with a brief description of two rehabilitation applications.

Correspondence should be addressed to D. Burke; email: david.burke@ucd.ie

First received 8 December 2000 and in final form 2 March 2001

MBEC online number: 20013579

2 Methods

The basis of our blink detection method invokes the classic Euclidean distance formula to measure the *colour difference* between successive frames of images. Each pixel has a colour associated with it represented by its discrete *tristimulus* values (red (R), green (G), and blue (B) parameters corresponding to intensities ranging from 0 to 255). The colour distance for a pixel (i, j) between frame n and $n-1$ is calculated from:

$$d_{ij}[n] = \sqrt{(R_{ij}[n]-R_{ij}[n-1])^2+(G_{ij}[n]-G_{ij}[n-1])^2+(B_{ij}[n]-B_{ij}[n-1])^2}$$

By applying a threshold to the above equation, we can extract only those pixels that have changed significantly in intensity between two successive frames, i.e. if $d_{ij}[n] > d_{threshold}$ then set pixel (i, j) colour to blue say, otherwise set to white. Fig. 1 illustrates the effect for a blink.

Next, we determine the two aggregations of pixels having the largest areas by using an algorithm that traces the perimeter of all groups of adjacent pixels in the frame and reports the respective areas. Eyes are classified as detected when the horizontal and vertical separations between the largest two aggregations are verified to be between obvious anatomical bounds (assuming an approximate vertical orientation of the head).

Although the techniques presented thus far can be used to detect eye blinks, they still require some refinement to remove false detections due to saccadic movements of the eyes, or in the case of one patient, to filter the effect of nystagmus. Nystagmus can also be problematic with infrared-based systems (CLEVELAND *et al.*, 1992). To differentiate eye movements from actual eye blinks, we make use of the fact that on average, pixels representing closed eyes will be predominately flesh coloured. On detection of the first blink (or a saccadic eye movement as the case might be), we sample a rectangular area

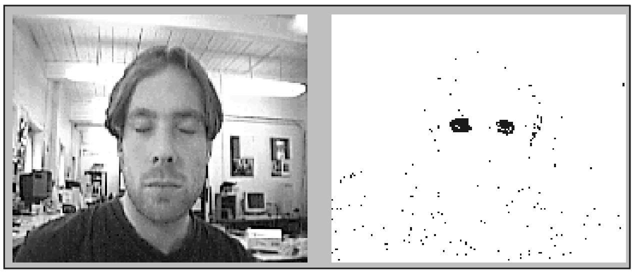

Fig. 1 *Detection of blinks by the colour distance method*

surrounding the eyes (proportional to the inter-eye separation) to generate a statistical histogram of flesh colours. To simplify computation, and remove brightness information, we normalise the tristimulus values, producing the *trichromatic coefficients* (GONALES and WINTZ, 1987):

$$r = \frac{R}{R+G+B}$$

$$g = \frac{G}{R+G+B}$$

$$b = \frac{B}{R+G+B}$$

Since $r+g+b=1$ we can now evidently describe *chromaticity* by just two values, e.g. r and g. Fig. 2 illustrates an example of a histogram of flesh colours sampled in an area surrounding the eyes.

The brightness has been normalised and we are now viewing just the chromatic (*hue* and *saturation*) information. Once the histogram has been generated, the trichromatic coefficients of pixels from subsequent frames are compared to the stored distribution and the relative frequency of each measured chromaticity is determined. Frequencies above a threshold are characterised as flesh colour. Fig. 3 illustrates the flesh colour filtering mechanism (flesh colours are shown in black).

For successive blinks, we compute the average probability that the detected pair of pixel aggregates is flesh coloured (i.e. eyelids are closed) and probabilities above a certain value are classed as true blinks.

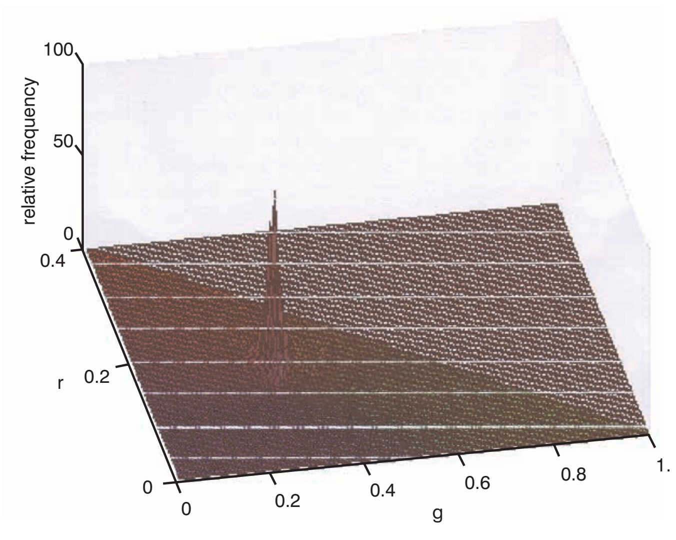

Fig. 2 *Distribution of the trichromatic coefficients of a flesh sample*

3 Results

Five volunteer subjects (two male, three female, aged between 21 and 25) each performed a set of five experimental blocks lasting two minutes each. This translates to 50 minutes ($5 \times 5 \times 2$) of analysis. During each experimental block, subjects were placed in a seated position 0.5 m from the digital camera and asked to look in the direction of the camera while remaining reasonably still. A second camera recorded the subject's face and its output was routed to a video/audio capture board in a separate PC. Each detected blink resulted in an audio track to toggle between play and pause and this audio was recorded simultaneously by the capture board. The captured audio and video were then manually analysed offline and the fraction of true positives and false positives determined for each experimental block for each subject (Table 1). Fig. 4 presents the results (mean and standard deviation) for each subject. Performance averaged over all subjects yields 83.74 ± 0.03% for true positives and 2.71 ± 0.01% for false positives.

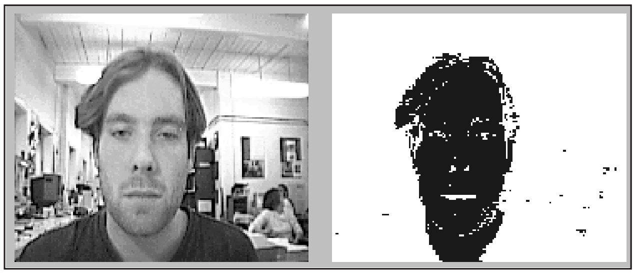

Fig. 3 *Detection of flesh colours*

Table 1 Results

	Trial 1			Trial 2			Trial 3			Trial 4			Trial 5		
	A	D	F	A	D	F	A	D	F	A	D	F	A	D	F
Subject 1	54	48	0	48	39	2	48	42	0	45	38	1	57	52	1
Subject 2	25	20	0	13	10	0	11	10	1	40	35	0	32	27	0
Subject 3	11	9	1	23	19	2	16	15	0	12	8	0	21	19	0
Subject 4	42	34	0	29	22	2	49	38	1	17	15	0	23	21	2
Subject 5	33	27	0	34	27	2	33	26	0	36	30	0	33	29	3

A: actual blinks; D: detected blinks; F: false positives.

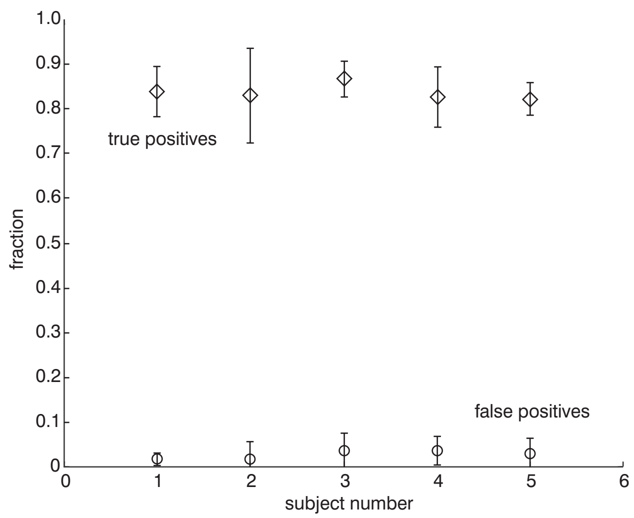

Fig. 4 *Fraction of true positives and false positives for each subject*

4 Conclusions

We have harnessed the above techniques in two successful rehabilitation applications to date. A component architecture has been developed in which the blink detection system runs in a separate process and interprocess communication (IPC) allows different applications to be effectively 'plugged-in'. The first application uses purposeful blinks to control the switching action of a CD player, ubiquitous in the average personal computer. The second application utilises blinks to control a custom-designed communication program, allowing users to build up sentences letter by letter, and produce synthesised speech. With each application, an adaptable logical decision rule operates allowing different blink rates to trigger the corresponding control action. Only the on-response is detected since this allows us to filter eye movements, which can sometimes give erroneous results.

We have presented preliminary results on the use of real-time digital image processing in the implementation of a non-contact human–computer interface. The results are promising, indicating that this type of technology is sufficiently robust to be used in rehabilitation applications. Where large head movements occur, the system does not attempt to detect blinks (since determining the position of the eyes becomes problematic) although this has not been an issue with the type of patients the system was built for. We are currently developing an improved set of algorithms to account for large head movements by tracking the position of the eyes (by first locating the face by the flesh colour filtering mechanism) and performing the difference operation on the pixels marked as representing the eyes.

Acknowledgments—The authors wish to thank the Occupational Therapy Department at the National Rehabilitation Hospital, Ireland for constructive criticism and fruitful discussions.

References

Clarke, L., Harper, P., and Reilly, R. (1998): 'Video based gesture recognition for augmentative communication'. TIDE Third Congress Papers, Helisinki

Cleveland, N. R., and Doyle, P. G. (1992): 'Eyegaze communication system, How does it work? Who can use it?'. Closing the gap – 10th Annual Conference, Minneapolis, Minnesota, Oct. 22–24, 1992

Gonales, R. C., and Wintz, P. (1987): 'Digital image processing', (Addison-Wesley), 2nd edn

Miyazaki, S., Ishida, A., and Komatsuzaki, A. (2000): 'A clinically oriented video-based system for quantification of eyelid movements', *IEEE Trans. Biomed. Eng.*, **47**, pp. 1088–1096

Authors' biographies

David Burke received his BE and MEngSc from the National University of Ireland (NUI), Dublin, in 1997 and 1998, respectively. He is currently working towards a PhD at the NUI, Dublin. His research interests are in the areas of dynamical systems analysis applied to biological modelling and Internet-based human-computer interaction.

Tomas Ward received his BE, MEngSc and PhD degrees from the National University of Ireland, Dublin, in 1994, 1996 and 1999, respectively. From 1999–2000 he taught at the Department of Computer Science at the National University of Ireland (NUI), Maynooth. Currently he teaches in the Department of Electronic Engineering at NUI Maynooth. His research lies in the areas of biological modelling and human-computer interaction.

Annraoi De Paor is Professor of Electrical Engineering at the National University of Ireland, Dublin. He founded and directs the Engineering Research Laboratory at the National Rehabilitation Hospital, Dún Laoghaire. He has published over 200 papers on control theory, biomedical engineering, renewable energy systems, engineering education, classical electrical engineering, and geomagnetism.