

Adaptive Noncontact Gesture-Based System for Augmentative Communication

Richard B. Reilly, *Member, IEEE*, and Mark J. O'Malley, *Senior Member, IEEE*

Abstract—An adaptive noncontact gesture-based system for augmentative communication is described. The system detects movement of any anatomical site through the analysis of reflected speckle. This movement is converted into two-dimensional (2-D) cursor coordinates and an adaptive software interface provides click actions and decision strategies. The system requires no accessory to be placed on the user. The system was developed in conjunction with user groups, who participated in the evaluation of the system. The usability results obtained illustrate the utility of the system. The system also compared favorably with other interface solutions.

I. INTRODUCTION

AUGMENTATIVE and alternative communication (AAC) systems can be described as methods which provide enhanced communication possibilities. A major goal of AAC is to provide access to technology for those without the fine motor control necessary to drive the “standard” interfaces such as keyboard and mouse. Solutions can be divided into contact, coupled and noncontact procedures. Contact methods can be defined as those which require physical contact for communication, such as head pointers and pointing sticks. Coupled procedures can be defined as those which sense a biophysical change, e.g., electroencephalographic and event related methodologies. Noncontact methods include motion detection systems and may or may not be cordless.

When assessing or prescribing an AAC system, it has been understood by clinicians for some time that not one but rather a combination of device solutions is deemed best practice. For such a multimodal approach to function, the AAC system must be customisable to the individual, thus requiring each of the distinct elements of the system to be highly configurable to meet the individual's needs. It is generally accepted that users can “adapt” their response to suit the interface device. A more appropriate solution would include the ability for the system to adapt to the user. The inability of some AAC devices to address these issues results in a reduced number of available options for flexible human-computer interaction [1, pp. 141–200].

Comments on the design and subsequent evaluation of such interface systems with respect to pointing systems, such as the mouse interface, have been reported [1, pp. 311–374]. Functional requirements include the ability of the user to

access all areas of the application screen, the size and spacing of the icons, the ability to execute a mouse click and adequacy of the sensory feedback for the task in question. These requirements can be difficult to achieve using a noncontact device, as the movement sensing complexity poses a number of design restrictions on the quality of interaction possible. Several motion and movement tracking devices are available, some placing the sensing element at an anatomical site [1]–[6]. Some systems employ transmitters worn by the user.¹ Others use reflective markers which echo the sensing signal back to a fixed reference transceiver module.^{2,3} Others are based on changes in electromagnetic properties, utilising sensing coils and resolving their movement within a three-dimensional (3-D) fixed space. Some systems make use of eye movements where the centre of the iris is tracked [7], while in others image sequences from charge-coupled devices (CCD's) are analyzed and the contrast of the hair and face employed to recognize face orientation [8]. Such tracking systems provide mouse movement but to fully access and control standard application software, the user requires control of both the cursor and select/click action.

The main objective of this paper is to describe an adaptive noncontact gesture-based system for augmentative communication. It consists of a motion sensing device, a data processing unit and an adaptable control interface for select/click action. The developed system requires no accessory to be worn by the user and is independent of the software application used.

II. USER SPECIFICATION

A survey of users and an analysis of their requirements for a noncontact AAC system was carried out. The aim of the survey was to specify the human factors to be addressed and to define a corresponding functional specification list for the system. This included a description of the user's reference position with respect to the computer, a description of the functional movement available and a definition of the comfortable computer monitor reading position. It was also important to verify if a noncontact AAC system would be acceptable.

Eleven users were questioned, five women and six men. They ranged in age from 4 to 51 years, five children/adolescents and six adults. Their pathologies ranged from

Manuscript received March 29, 1996; revised June 2, 1998 and March 15, 1999. This work was supported in part by the European Commission under the TIDE program.

The authors are with the Department of Electronic and Electrical Engineering, University College of Dublin, Belfield, Dublin 4, Ireland.

Publisher Item Identifier S 1063-6528(99)04475-4.

¹HeadMaster System: Prentke Romich Co., Wooster, OH.

²HeadMouse, Origin Instruments Corp, 854 Greenview Drive, Grand Prairie, TX 75050.

³Tracker, Madenta Communications Inc. 9411A-20 Ave. Edmonton, Alta. 6N 1E5, Canada.

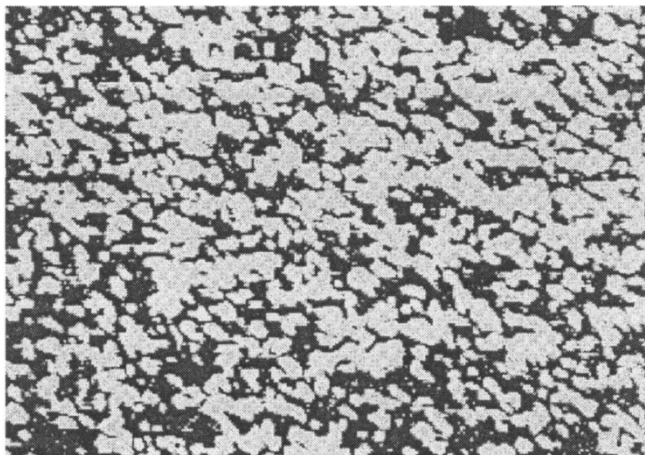


Fig. 1. The reflected speckle generated from skin.

cerebral palsy (2), Friedrich's ataxia (1), spinal cord injuries (2), multiple sclerosis (1), Werdnig Hoffmann (3), muscular dystrophy (1) and Lockin-syndrome (1). Seven regularly use a computer (all children/adolescents, except the four year old). Two intended to purchase a computer and one no longer wished to use one. The input interface solutions used ranged from mechanical contact switches, keyboards with keyguards, joysticks, trackballs, eye-gaze and movement tracking systems [HeadMouse²].

Following detailed questioning several observations were made. For the AAC system to be fully augmentative head, facial, feet or hand movements must be capable of being detected. Limited or low amplitude controlled movement, associated with traumatic quadriplegia or neurological quadriplegia must also be detectable. The AAC system should allow the motion detector component to be remote and independent from the monitor, such as being set up on an adjustable support, in order to facilitate reception of hand, feet, or facial movements. User-borne accessories are often considered a sign, label or advertisement for an individual's disability. Thus, the system should have no accessory. The tracking reception distance should be adaptable within a 20–80 cm range. Software application independence was deemed a minimum requirement by users, who all too often are restricted in the software they can control as a result of their disability and the interface system they use.

III. MOTION SENSING DEVICE DESIGN

When a beam of laser light is incident on a scattering object, it produces a speckle pattern, due to the object's surface roughness. The phenomenon of speckle patterns is illustrated in Fig. 1 where the reflection object is skin. Should the reflecting surface move, the speckle pattern will move proportionally. Therefore, the motion of a surface can be estimated. This motion estimation is similar to other techniques, where laser-based optical sensors are employed in a number of varied noncontact applications from range finding [9] to displacement sensors [10] to skin blood flow analysis [11], [12]. A movement sensor was developed based on this principle, incorporating the use of two laser diodes as emitters,

TABLE I
CHARACTERISTICS OF THE SENSING HEAD OF THE NONCONTACT SENSOR, WHICH INCLUDES TWO-LASER DIODES AND TWO-LINEAR CCD ARRAY AS DETECTORS

Transmitter	
Laser Diode	Sharp LT011MCO
Number of Lasers	2 (x and y)
Laser peak power	3mW
Wavelength	840nm
Mean power per Laser Spot	200 μ W
Receiver	
CCD Sensor	Thomson-CSF TH17803A
Number of CCD's	2 (x and y)
Number of pixels	1780
Pixel Clock	2 MHz
Sensor	
Working distance	20 - 40 cm
Maximum angular speed	10° per second

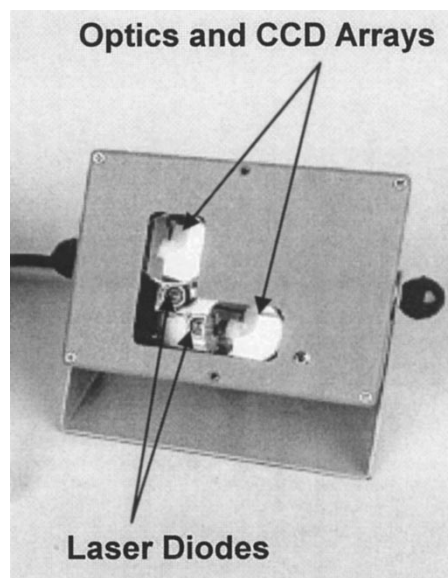


Fig. 2. Noncontact sensor. The two laser sources, one for each axis, can be seen in the lower left-hand side, with lens in front of each of the CCD arrays.

one for each axis and correspondingly two linear CCD arrays with collection optics as detectors. The characteristics of the laser diodes and CCD arrays, together with the resulting sensor properties are listed in Table I. A photograph of the prototype sensor is shown in Fig. 2, which gives the position of the laser diodes and CCD arrays, one per axis.

An optical bandpass filter at the wavelength of the laser diode, attached to the photosensitive areas of both CCD arrays, reduces saturation due to the external ambient light. An illumination diode was also incorporated in close proximity to the CCD arrays to bias the sensors with a low continuous level of light and thus offset lag effects of low-level signals [13].

An overview of the signal processing tasks required of such a sensor can be seen from Fig. 3. Signal conditioning entails analog amplification ($A_v = 600$) and bandpass filtering (passband: 20 kHz–150 kHz) prior to one-bit analog-to-digital conversion. Motion estimation can be achieved by correlating two consecutive frame signals. A frame is loaded into a first-in first-out (FIFO) store and is then cross-correlated with the following frame, by shifting to the left and to the right about

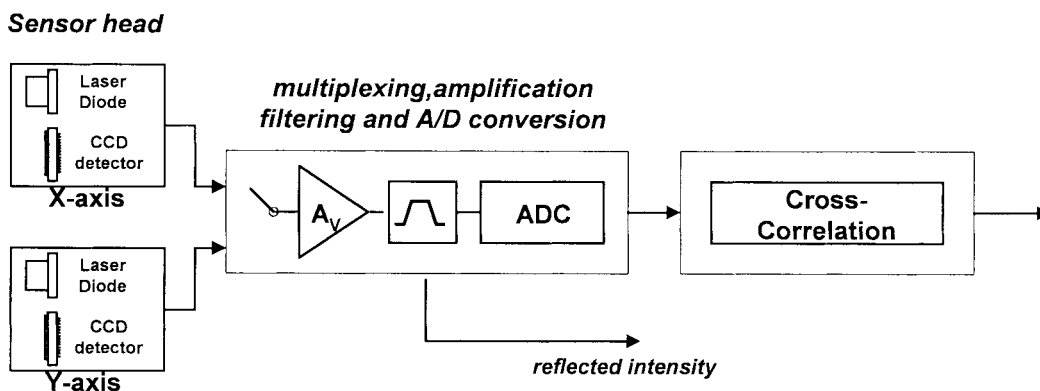


Fig. 3. Signal processing of reflected speckle signal. The sensor information is cross-correlated to provide directional information on the diffusing surface.

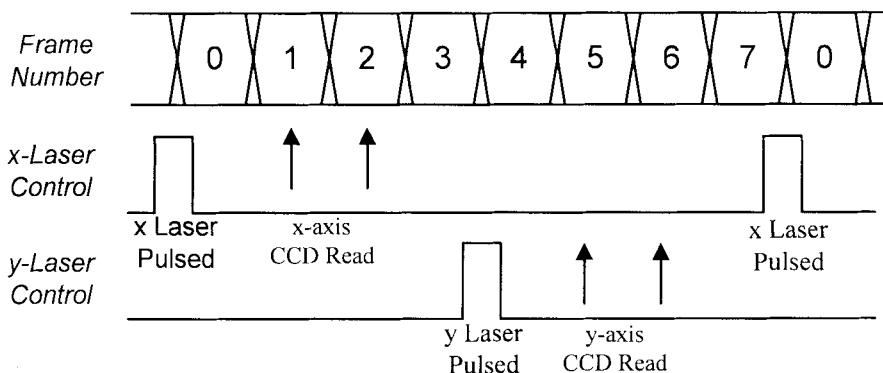


Fig. 4. Timing information for signal processing. The sensor information consists of output from the X-axis and Y-axis CCD arrays. Each array is cross-correlated with the previous frame to provide directional information on the diffusing surface.

the central pixel. This provides a movement and direction estimate of the reflecting surface. The correlated outputs from each axis are combined to generate a two-dimensional (2-D) Cartesian coordinate of the reflection object.

As the sensor is for use in human-computer interaction, it is necessary to ensure that mean laser light power is below the continuous laser light eye safety limit ($230 \mu\text{W}$ at 840 nm [14]). This was achieved by pulsing the laser diodes for half of one frame, and sensing the subsequent two frames, as shown in Fig. 4. With a 2-MHz pixel frequency, the duration of each frame is $880 \mu\text{s}$. Correlated frames are frame 1 and frame 2 for the x -axis and frame 5 and frame 6 for the y -axis. Although the light pulse total duration is half a frame, only one pulse is used for consecutively correlated frames. These pulses are synchronised so that the amount of light collected for each frame is the same (i.e., one pulse overlaps two consecutive frames: frame -1 and 0 for the x axis, frame 3 and 4 for the y axis). Thus, new movement information is produced at a rate of 142 Hz .

In the configuration shown in Fig. 5, the skin of the user acts as a reflector generating a speckle pattern from which 2-D movements can be identified. A third dimension can be resolved from this sensing configuration by calculating the reflected intensity. With the configuration described above the maximum angular speed was measured at 10° per second.

One of the main drawbacks with using this sensing principle is the diffusion properties of human skin. The skin acts

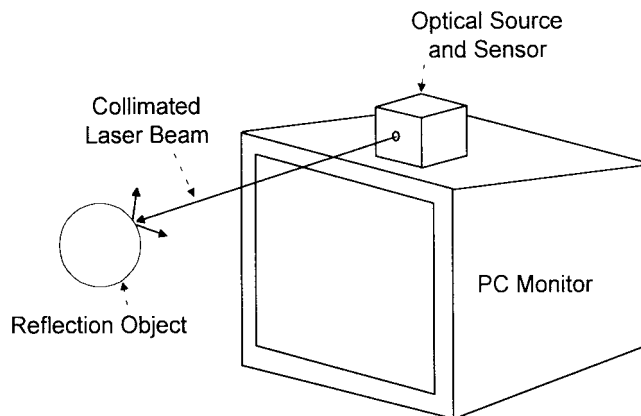


Fig. 5. System set-up for human-computer interaction using an optical light source as emitter and collection optics as receiver. The 3-D movement of the reflection object is transformed into 2-D cursor movement.

as a depth diffuser, as opposed to a surface diffuser such as an opaque screen, which can result in the detection of underlying blood flow. Moreover, biological tissue such as skin is continually subjected to small movements which in turn leads to fast movement of the speckle grains [12]. Both of these effects lead to a reduction in the signal to noise ratio. This interference is termed sensor interference and is in range from 20 to 100 Hz .

The 2-D coordinate signal may be observed to possess some residual sensor interference but also tremor associated with the

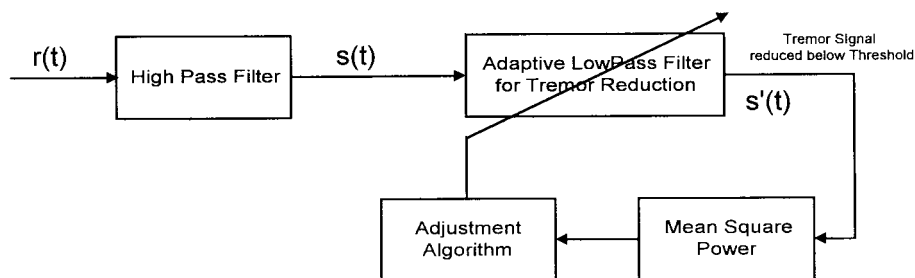


Fig. 6. Off-line training of adaptive filter.

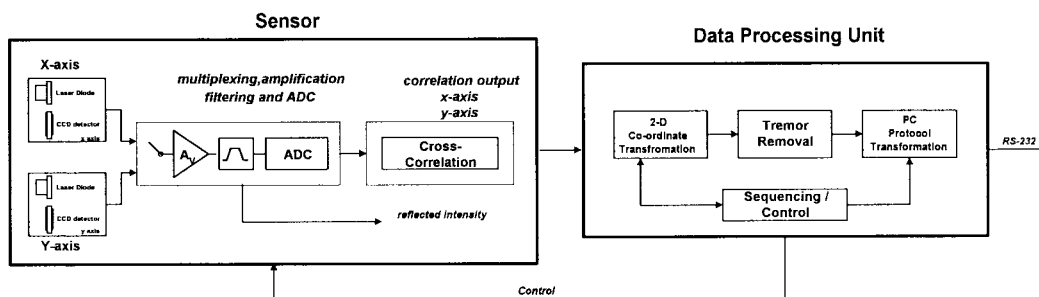


Fig. 7. Information flow from the sensor through the data processing unit to the RS-232 interface.

individual's pathology. Tremor is defined as the low frequency (1–15 Hz), rhythmic, purposeless, quivering movements resulting from the involuntary alternating contraction and relaxation of opposing groups of skeletal muscles [15]. A movement smoothing system was designed to remove low frequency tremor noise and higher frequency sensor interference. As the motion tracking system is continuously sensing movement, an adaptive smoothing system was employed to remove low frequency tremor noise while remaining highly responsive to intended voluntary movement.

An off-line adaptive filtering strategy is represented graphically in Fig. 6. The raw data $r(t)$ sampled at 100 Hz was passed through a third-order Chebychev high-pass filter, with a cutoff frequency of 0.1 Hz, followed by an analysis of the energy components contained in the resulting series, $s(t)$. This sequence was passed through a second-order Butterworth low-pass filter whose cutoff frequency was initially set at half the sampling frequency. The mean square power is calculated for the filtered residual sequence. The cutoff frequency of this filter was iteratively adjusted until the mean square power was less than a threshold value. This adaptive threshold for the mean square power was set to 25% of its initial value. This assumes that the energy associated with higher frequency components forms the majority of the residual energy. The information gained for each user, from this off-line adaptive procedure, was incorporated into the design of a fixed frequency lowpass suppression filter.

To develop a noncontact AAC system, the smoothed positional information is formatted to the typical protocol of a PC mouse (Microsoft© compatible). This allows a stream of 2-D Cartesian coordinate information from the noncontact sensor to be reproduced as smoothed cursor movement on a software application screen, via an RS-232 interface. The control of the sensor, transformation into 2-D coordinates,

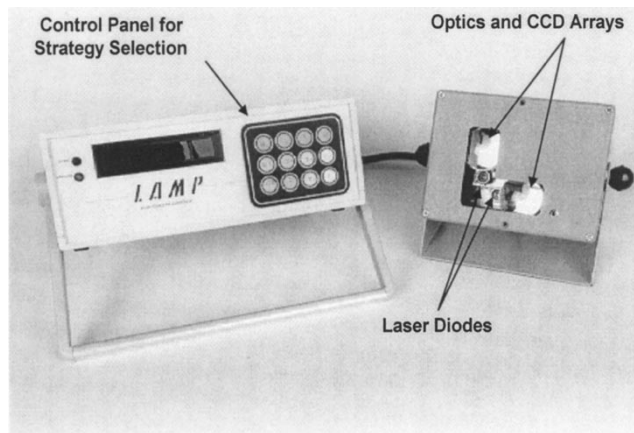


Fig. 8. Noncontact augmentative and alternative communication system. The sensor is shown on the right, while the data processing unit is shown on the left-hand side of the photograph. The control panel allows easy selection of click actions and decision strategies and other parameters.

filtering and generation of PC protocols is carried out by an 8-bit microprocessor (Intel 8051) within a data processor unit (DPU). The full tracking procedure is represented in Fig. 7. A photograph of the prototype sensor together with the DPU is shown in Fig. 8, where the front panel switches on the DPU allow adjustment of the sensor sensitivity.

IV. CLICK ACTIONS AND DECISION STRATEGIES

To allow augmentative interaction, systems must be capable of not only providing control of the mouse cursor movements, but full control of standard commercially available application software. This includes provision for item selection tasks such as the mouse click, double click and drag/drop features. For the system to be highly configurable it should require no internal PC hardware setup but use the standard I/O ports.

A click action can be defined as a single action performed by the user to generate a mouse click. A decision strategy is a procedure where the system makes assumptions on the type of tasks that are to be undertaken and requests a confirmation of this decision from the user through the use of a pop-up visual prompt or audio cue. Thus a click strategy allows for more complex item selection tasks. A number of click actions and decision strategies were developed for use with the noncontact sensor. These strategies are all generated within the DPU for maximum configurability. Commonly used methods of click generation include intensity variation and dwell time [1], see ^{1, 2}). Both of these methods were incorporated as click generation actions.

A. Click Actions

The variation of average intensity of the reflected speckle signal was employed as a preliminary strategy for the generation of a single click. This action relies on relative changes in the average reflected optical intensity and is not dependent on a predefined intensity threshold for trigger. Thus, moving a certain distance to or from the sensor generates a click action. It is believed that this type of movement is synonymous with "positive intention".

A dwell time strategy, also known as the "Acceptance Time technique" [1, p. 342] was also implemented as a click action generator, with audio cues. This click action is by far the most widely available in AAC systems. Due to the sensitivity associated with the noncontact sensor, a movement threshold was incorporated within this strategy design, allowing the user to remain close to a specified location for a set period of time.

Both of these techniques are generated within the DPU and are independent of the mouse driver and software application.

B. Decision Strategies

The limiting factor with these popular click actions for some users is that only one action is possible, e.g., single item selection or single click. The use of a more complex interaction strategy was developed to allow for the selection of more than one item, thus increasing user flexibility in controlling software applications. These more advanced decision strategies are an extension to the dwell time strategy and are based on the use of interactive dialogs or graphics.

These strategies are again generated by the DPU but unlike the previous ones, in conjunction with a Terminate-Stay-Resident program (TSR) on the computer. Once the DPU senses no movement above a user defined sensitivity threshold, the TSR is activated using a software command issued from the DPU. This command is embedded within the standard PC protocol signals from the DPU. The TSR hooks this command, notes the cursor coordinates, deactivates the current software application and causes a visual pop-up dialog box to appear on top of the application screen.

The frame dialog decision strategy is shown in Fig. 9 and provides the user with a visual cue that the interface phase has been entered. Cursor movement is now restricted to this dialog and the user has a specified time within which to execute a movement. If on the completion of the time duration the

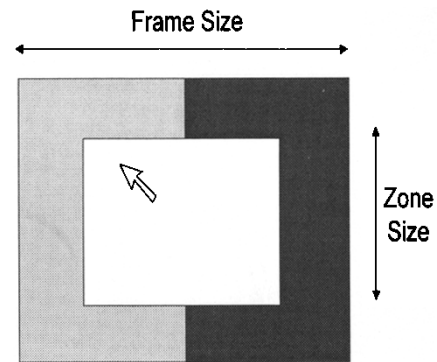


Fig. 9. Frame dialog decision strategy graphic, displayed to user for decision confirmation. By moving into the shaded/dark regions a click or double click action occurs. This dialog is initiated by the user remaining stationary for a defined time period. By remaining within the white inner area no click action occurs.

user has moved within one of the shaded zones, a click or double-click action is generated by the DPU. However, on remaining within the center region no click action occurs. The cursor is then mapped back to its position prior to the strategy and the TSR relinquishes control to the software application. Parameters such as the size of the dialog and duration of the gesture interface phase are adjustable from the control panel on the DPU (Fig. 8).

A gesture matrix decision strategy was also developed, which recognized specific movement, such as predominately y -coordinate movement, arising from head nodding. As with the frame dialog decision strategy, the DPU on sensing no movement forces the TSR to launch, popping-up the graphic and restricting cursor movement within this region. Fig. 10 illustrates the gesture matrix dialog graphic, where the user's movement was sensed to be vertical, from the centre to the top of the matrix. This decision confirmation strategy employs template matching to estimate the directionality of movement. Employing a series of different movements, each associated with specific actions, it is possible with the template matching procedure to structure a multidecision process. The decision process is based on the following criteria: the number of consecutive matrix elements traversed in the horizontal direction, the vertical direction and the inclined direction. The ability to alter the stored templates allows individual customization, however adaptability was found to be coarse with such a template matching procedure.

V. USABILITY EVALUATION

The assessment and evaluation of AAC systems is an important process and can be considered from both a technical and usability perspective [1], [16], [17]. Human factor testing was performed during the development phases of the system, culminating in usability tests. Three principal aims were considered in testing the usability: a functional test of the system as a human computer interface, through its function as a mouse pointer (cursor control, click, and decision strategies), its usability as a human computer interface for physically restricted users and a comparison of the system with other solutions that already exist for AAC.

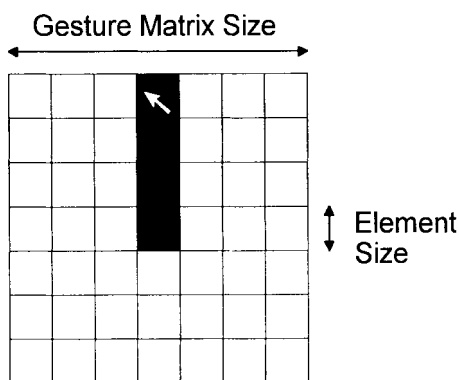


Fig. 10. Gesture matrix decision strategy graphic displayed to user for decision confirmation. The user is prompted by the display of the graphic to execute a gesture. In this example, the user's movement was sensed to be vertical, from the center to the top of the dialog graphic.

1) *Functional Test*: A functional test was initially carried out by twenty able bodied evaluators who are technically minded and aware of AAC interface issues. This test was to ensure the stability and reliability of the system. The testing concentrated on the filtered movement of the cursor, position of the user relative to the screen, start-up of the decision strategies and issues encountered by use of the device in multiple user situations. Testing of the interface dialogs was carried out on a simulated system, employing a standard mouse for cursor movement but using the click action and decision strategies for selection. The results provided useful information on the structuring of usability tests for the target population and on the formation of a detailed test plan, including the tasks assigned and the performance assessment methods. Detailed results obtained from the test were analyzed and all worthwhile recommendations and modifications to the user dialog were implemented.

2) *Usability Tasks, Test Population, and Procedures*: In the assessment of the system's usability as a human computer interface for physically restricted users, the main objectives were the verification that the interface dialog was valid for the intended user group, the verification that the dialog was both logical to novice users and sufficiently fast for experienced ones, the verification that the dialog did not distract the user from the task in hand and finally verification that the adaptive features of the interface were functional and of actual benefit to the intended user group.

The usability evaluation test on the complete noncontact AAC system was carried out in two phases. Phase one was carried out by fifteen able-bodied evaluators, with a full evaluation session of the order of 3–4 h duration. This phase defined typical parameters for the sensitivity of the movement tracking system and for the various click actions. Phase two was carried out by the physically restricted evaluators, who assessed the system using these parameter ranges. Due to the fatigue experienced by the physically restricted evaluators, clinical and work sessions are typically short in duration. Thus, the usability tests were reduced to 30 min by choosing only certain tasks.

The evaluation group consisted of six physically restricted users, three suffering from Traumatic Quadriplegia and three

from Neurological Quadriplegia. These were chosen with the aid of the Speech and Language Therapy Department of the National Rehabilitation Hospital (NRH) in Dublin, and were current or recent patients who had used PC-based communication devices. They typically had good head or limb control and were experienced users of a computer interface system and PC-based software. In this way, the test was not influenced by the evaluator being unaware of the PC, software or issues concerned with PC interfaces.

A set of tasks was designed in a format with which the physically restricted users were familiar and were not unduly difficult, but required comprehensive movements of the cursor and use of all of click actions and decision strategies. The tasks assigned were all typical of interfacing with commercially available software applications: select a specific item from a pull-down menu bar, page down to a specific screen of information using the scroll bar, select specific words from a sixteen word menu arranged in two rows of eight. Each of these tasks was performed with each of the following actions and strategies: dwell time click action, frame dialog decision strategy and gesture matrix decision strategy. Although the noncontact sensor can be used to track movement from any part of the body, it was located on top of the computer monitor for all usability tests, a distance from 30 to 35 cm from the head. This was to allow comparison with existing head tracking devices.

On performing the assigned tasks the physically restricted evaluators were asked for their comments on the tasks, which are summarized in Table II. The physically restricted evaluators all had success in carrying out the tasks assigned, mentioning that they were typically those required of interfacing with commercial software applications. The physically restricted evaluators were also asked for their specific comments on the click action and decision strategies, which are summarized in Table III.

3) *Comparison Tests*: Two comparison tests were carried out comparing the device with the HeadMaster system.¹ The first consisted of two expert able-bodied evaluators carrying out identical tasks using both systems. The task consisted of entering text from a typed page using a software keyboard emulator⁴ with the word prediction setting and the internal dwell time click strategy active. The results showed that both systems were comparable, as the same amount of text was entered by both evaluators in the same time period. A comparison of both systems with two physically restricted evaluators was also carried out with the same assigned task and click strategy. Both physically restricted evaluators are expert users of the HeadMaster system and are deemed by the Speech and Language therapy staff of NRH to be suitable for this device. The results indicated that both systems were again comparable as approximately the same amount of text was entered by both evaluators in the time allowed.

A comparison was made between a commercial, connected word voice recognition system⁵ and the noncontact system,

⁴WiVik System: Hugh MacMillan Rehabilitation Centre, 350 Rumsey Road, Toronto, Ont. M4G 1R8, Canada.

⁵Dragon Dictate: Dragon Systems, Inc., 320 Nevada Street, Newton, MA 02160.

TABLE II
COMMENTS FROM THE PHYSICALLY RESTRICTED EVALUATORS ON THE TASKS ASSIGNED TO THEM FOR THE USABILITY TEST

Physically Restricted Evaluator	Task Assigned		
	Menu Bar Selection	Word Selection	Scroll Bar Control
A	Good representative task. Using an icon to switch application would be useful.	Good representative task. More contrast required between text and box.	Good representative task. Non-linear movement feature interesting.
B	Difficulty with gesture matrix decision strategy.	Good representative task. Good idea of sensitivity threshold. Dialog graphic should disappear as soon as you click.	Need an inactive border around the screen. Variable width of scroll required. Variable size of text required.
C	Frame-dialog decision strategy requires practice. Dwell time click action easiest to use.	Word font is too large - more words in smaller fonts would be more typical. Dwell time click action easiest.	Good representative task - simple to use.
D	Good representative task. Hard to use gesture matrix decision strategy in corner of screen.	Colour/size should be variable. More contrast.	Difficulty with gesture matrix decision strategy. Good representative task.
E	Good. Bigger dialog graphics required.	Primary colours better (red, green, blue, black, white) Time delay before printing caused confusing.	No problem in accomplishing task.
F	Frame dialog decision strategy preferred option	Not enough words on screen. Frame dialog decision strategy preferred option.	Good representative task. Hold facility good. Identify zone in frame for double click, hold etc.

with a software keyboard emulator.⁴ Results showed that there was a difference in speed between the two systems when it came to entering text, dictation being the faster, more efficient and natural method when the speech recognition system was fully trained to the physically restricted user's voice and rate of speech production. However, it was found that the noncontact system offered a more intuitive method of cursor control than the voice command system.

A comparison was made between a mouth stick with keyboard and this noncontact system with a software keyboard emulator.⁴ The results showed that the noncontact system proved to be more appropriate for the physically restricted user, who experienced less fatigue and entered more text in a set period of time. This was due to the reduced range of motion required and word prediction facility that was provided with the keyboard emulation software.

VI. DISCUSSION

The movement tracking procedure using reflected speckle provides an unobtrusive, noncontact method of human computer interaction, which approaches a more socially acceptable solution. Indeed the absence of any accessory makes the system easier to configure and thus more user friendly in a clinical environment. It can be controlled by any reflective surface of the body, head, trunk or limb, by users of any age and any level of cognition. The system can be used with an unlimited number of software applications, principally due to the development of the click actions and decision strategies

around an already existing commercial mouse driver. The use of commercial software drivers providing further control over responsiveness and sensitivity, as their internal features, such as nonlinear cursor movement, are preserved. This ensures maximum compatibility with existing commercial software, one of the main specifications for the system.

The selection of the click actions and decision strategies, together with parameters such as stationary threshold, dwell time, gesture duration and frame dialog size are adjustable from the control panel of the signal processing unit. The storage of an individual's parameters is also possible, allowing the system to be setup quickly by therapists. This is advantageous in multiple-user or multiple application settings, such as schools and clinics. The DPU also allows standard switches, such as rocker switches, to be interfaced allowing the individual to by-pass the click action and decision strategies and use instead these standard contact switches. The simplicity of configuration and wide range of parameter settings all add to the utility of the device. The control panel can be replaced with a software-based pop-up dialog or pull-down menu. This advantageous feature allows a user to alter the click and decision strategy settings on-line.

Most of the physically restricted evaluators accepted the system immediately, as an alternative, original interactive system. They were further motivated by the speed at which they progressed with the system. The mouse functionality that was offered also provided further motivation. The fact that the device offered a noncontact form of computer interaction

TABLE III
COMMENTS FROM THE PHYSICALLY RESTRICTED EVALUATORS
ON THE CLICK ACTIONS AND DECISION STRATEGIES EMPLOYED
WITHIN THE TASKS ASSIGNED FOR THE USABILITY TEST

Physically Restricted Evaluator	Click and Decision Strategies		
	Dwell Click Action	Frame Dialog Decision Strategy	Gesture Matrix Decision Strategy
A	Good. Visual and audio feedback option good for training. Option to update parameters required.	Place a small keyboard within frame area. Visual aid to find mouse would be useful. Visual and audio feedback to start.	Nice access to function with hook movement. Needed to change parameters for accurate control.
B	Could be faster. Inactive region in part of application screen to disable selection.	More difficult than dwell time. Easier to use with smaller frame, less movement required. Perhaps not replacing cursor to original starting point would be useful.	Graphic on the screen too long - remove after gesture phase. Larger box more effective.
C	Easy to use. Dwell time at three seconds perfect for user. Does not have to think, is an automatic reflex.	Preferred smaller frame. More control - less movement. Would take more training.	Dependent on users range of motion. Slower movement - more control.
D	Fine. No problem to use.	Fine. A problem at times remembering operation.	Easier with bigger graphic.
E	Easy to use. Three seconds preferred speed.	Accuracy required. Tend to stick with one direction of selection Required more training.	More time needed before gesture phase. Hard to control. Easy to double click by accident.
F	Did not like it. No control, selection every time it stops.	Prefer to see screen - transparent graphic better. Easy, logical preferred option.	Difficult.

made it psychologically acceptable, leaving the user free to relax and speak.

Users initially experienced some problems in controlling the cursor about the application screen. However once the relationship between head (trunk or limb, etc.) movement, speed of movement and the resulting cursor movement was learnt, the user's progress with the system steadily increased. The level of success and accuracy achieved was dependent on the amount of training and on the degree of head control demonstrated by the user. Competent control was typically observed after a period from 10 to 15 min. Time was required to find the optimum sensor position for the user. In some cases this was complicated by the bulk of a wheelchair, which prevented the user from positioning his/her head sufficiently close to the monitor and sensor. In such cases different optical reflection points for the sensor were considered when appropriate. Hand or shoulder were employed as reflection objects and this resulted in increased control. However an improvement of the optical sensor range is a suggested solution to this general problem. The sensitivity and accuracy were considered sufficient, but in some cases the sensitivity was judged to be too high, particularly during sneezing and leg spasm.

The dwell time click action required the least amount of training time to be successful. The main comment on the dwell time action was that the cursor was always active, i.e., always ready to select. The ability of some users to maintain the cursor at a fixed position for a set duration of seconds was found to be arduous. This required experimentation with the motion tracking sensitivity and the subsequent introduction of a movement threshold. Dwell duration and sensitivity

threshold were made adjustable from the front panel of the DPU and were found to be user dependent.

The short learning curve experienced with the frame dialog strategy reduced the learning time required for the gesture matrix decision strategy. The physically restricted evaluators believed that the frame dialog strategy would be most satisfactory to use as a selection interface after a short period of training. This strategy was deemed by both the physically restricted evaluators and their Speech and Language Therapists to be very intuitive as movement into one of the shaded areas results in a click or double-click action. Three sizes of frame dialog were developed: small, medium and large. The medium frame dialog was the size preferred.

Most of the physically restricted evaluators when using the gesture matrix decision strategy had more success with the larger dialog. They felt that the time delay between positioning the cursor and entering the gesture stage was too short and that the time during the gesture analysis phase was too long.

The level of memorization of the strategies associated with the system was found not to be demanding. The audio and visual prompts provided by the system were seen as essential during the training phase. It was commented by some of the physically restricted evaluators that with practice and training, neither an audio nor visual prompt would be required, as selection of the different strategies/mouse functions would become an automatic or reflex action.

The responsiveness of the device proved to be good for each of the click actions and decision strategies. The range of functionality offered, such dwell time, stationary tolerances, audio and visual prompts, provided flexibility in meeting the needs of the majority of users.

The physically restricted evaluators found the system efficient, with all of them reporting success in carrying out all the defined tasks (Tables II and III). The level of efficiency obtained by the physically restricted evaluators was dependent on their degree of head control and level of fatigue. Fewest errors were made when using the frame dialog decision strategy. Lack of control in the dwell time click action caused undesired selection. Errors were also evident when using the gesture matrix decision strategy but after practice these were reduced considerably. It must be noted that without good head control, cursor movement errors were made irrespective of the click action or decision strategy used.

Factors such as fatigue and perceived exertion were deemed by the physically restricted evaluation group to be less than those experienced with similar augmentative communication devices, and considerably less than systems requiring mouth sticks and rocker switches. With training and practice, any sensation of discomfort and fatigue reduced and in most cases disappeared. However, a high level of concentration was required initially during training and some time was required to obtain the optimum sensor reflection position for each user. The possibility to be a noncontact method allowing the use of any anatomical site is considered by the physically restricted evaluators to be a major factor in making the system comfortable to use.

The motion sensing device proved comparable with the current commercially available HeadMaster device.¹ The Head-

Master was found to respond better to fast head motion. However the advantage of having no user accessory together with the click actions and decision strategies, makes the non contact gesture-based system a useful new solution for augmentative and alternative communication.

VII. CONCLUSION

A noncontact movement detection system has been developed. The system requires no accessory to be worn by the user and can operate standard application software. It consists of a motion sensing device, a data processing unit and an adaptable control interface for select/click action. The dwell time click action and the frame dialog and gesture matrix decision strategies all proved to be highly feasible interaction methods and can be used not only with this noncontact movement sensor, but with other AAC tracking systems. Current research is concentrated on increasing the optical range of the sensor and the development of on-line tremor suppression algorithms. A hidden Markov model-based pattern recognizer is currently being developed to allow recognition of user specific gesture [18]. This would allow the user, with the aid of a Therapist, to train the gesture recognition system to an individual's needs. In this way, the system would respond to the needs of users with limited control. This system provides the possibility to assess human-computer interaction and allow experimentation to improve the quality of augmentative and alternative communication.

ACKNOWLEDGMENT

The authors wish to acknowledge the support of the members of the LAMP consortium and also to A. Hanson who help conduct the user trails. The authors also wish to thank the reviewers whose helpful comments have led to an improved paper.

REFERENCES

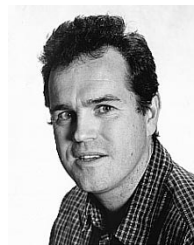
- [1] A. M. Cook and S. M. Hussey, *Assistive Technologies: Principles and Practice*. St. Louis, MO: Mosby-Year Book, 1995.
- [2] W. S. Harwin and R. D. Jackson, "Computer recognition of head gestures in cerebral palsy," in *Proc. 13th Annu. Conf. RESNA*, 1990, pp. 257-258.
- [3] W. S. Harwin and R. D. Jackson, "Analysis of intentional head gestures to assist computer access by physically disabled people," *J. Bio-Med. Eng.*, vol. 12, no. 3, May 1990.
- [4] W. S. Harwin and R. D. Jackson, "A simple mathematical analysis of head movement," in *Proc. ICAART*, Montreal, P.Q., Canada, 1988, pp. 386-387.
- [5] P. K. Hansen and J. F. Wanner, "Hawkeye, a "Penlight" head pointer," in *Proc. 16th Annu. Conf. RESNA*, Las Vegas, NV, 1993, pp. 440-442.
- [6] P. Hansen, D. Dobson, and J. Wanner, "Free wheel: The cordless new headpointing device," in *Proc. ICAART*, Montreal, P.Q., Canada, 1988, pp. 372-373.
- [7] X. Xie, R. Sudhakar, and H. Zhuang, "Real-time eye feature tracking from a video image sequence using Kalman filter," *IEEE Trans. Syst., Man and Cybern.*, vol. 25, no. 12, 1995.

- [8] K. Mase, Y. Watanabe, and Y. Yasuhito, "Headreader: Real-time motion detection of human head from image sequences," *Syst. Comput. Japan*, vol. 23, no. 7, 1992.
- [9] H. Turk, D. C. Leepa, R. F. Snyder, J. I. Soos, and S. Bronstein, "Beam deflection and scanning technologies," in *Proc. SPIE*, 1991, vol. 1454, pp. 344-352.
- [10] S. Huang and Y. Lin, "Prototype system of three-dimensional noncontact measurement," *Int. J. Advanced Manufacturing Technol.*, vol. 11, no. 5, pp. 336-342, 1996.
- [11] B. Ruth, "Measuring the steady-state value and the dynamics of the skin blood flow using the noncontact laser speckle method," *Med. Eng. Phys.*, vol. 16, no. 2, pp. 105-111, 1994.
- [12] S. Kashima, "Non-contact laser tissue blood flow measurement using polarization to reduce the specular reflection artefact," *Opt. Laser Technol.*, vol. 26, no. 3, pp. 169-175, 1994.
- [13] J. C. Dainty, "Laser speckle and related phenomena," in *Topics in Applied Physics*. New York: Springer-Verlag, vol. 9, 1975.
- [14] IRPA, "Guidelines on limits of exposures to laser radiation of wavelength between 180 nm and 1 mm," *Health Phys.*, vol. 49, no. 2, pp. 341-359, 1985.
- [15] Mosby's Medical, *Nursing and Allied Health Dictionary, 4th Edition*. St. Louis, MO, 1994.
- [16] D. Shafer and J. Hammel, "A comparison of computer access devices for persons with high level spinal cord injuries," in *Proc. 17th Annu. Conf. RESNA*, Nashville, TN, 1994, pp. 394-396.
- [17] J. Neilsen, *Usability Engineering*. New York: Academic, 1993.
- [18] S. McInerney and R. B. Reilly, "Hybrid multiplier/cordic unit for online handwriting recognition," in *Proc. 1999 IEEE Int. Conf. Acoust., Speech and Signal Processing*, Mar. 1999.



Richard B. Reilly (M'91) received the B.E., M.Eng.Sc., and Ph.D. degrees in electronic engineering from the National University of Ireland in 1987, 1989 and 1992, respectively.

In 1988, he joined Space Technology Ireland, and the Department de Recherche Spatiale, part of the CNRS group in Paris, France, where he developed a DSP-based spectrum analyzer as part of an on-board experiment for the NASA satellite, WIND. In 1990, he joined the National Rehabilitation Hospital as Research Engineer. In 1992, he became a Postdoctoral Research Fellow at University College Dublin, Dublin, Ireland, where his research interests included nonlinear system identification and modeling, together with alternative and augmentative human-computer interaction particularly focused on speech enhancement and gesture recognition. He is currently on the academic staff as a College Lecturer in the Department of Electronic and Electrical Engineering at University College, Dublin, where his research interests include biomedical and multimedia signal processing and human-computer interaction.



Mark J. O'Malley (M'88-SM'96) received the B.E. and Ph.D. degrees in electrical engineering from the National University of Ireland in 1983 and 1987, respectively.

After one year working in industry on European Space Agency projects, he joined the academic staff as a Lecturer in the Department of Electronic and Electrical Engineering at University College Dublin, Dublin, Ireland. He is joint head of the Rehabilitation Engineering Laboratory at the National Rehabilitation Hospital, (Dun Laoghaire, Ireland). His research interests are modeling and control with applications in biomedical engineering and power systems. In 1994, he was awarded a Fulbright Fellowship and subsequently spent seven months as a Research Fellow in the Department of Orthopaedic Surgery at the University of Virginia, Charlottesville.