MULTI-MODAL HANDS-FREE HUMAN COMPUTER INTERACTION: A PROTOTYPE SYSTEM

Frangiskos Frangeskides and Andreas Lanitis
School of Computer Science and Engineering, Cyprus College,
P.O. Box 22006, Nicosia, Cyprus.

Keywords: Contact-less HCI, HCI for Disabled Users, Multi-Modal HCI.

Abstract: Conventional Human Computer Interaction requires the use of hands for moving the mouse and pressing keys on the keyboard. As a result paraplegics are not able to use computer systems unless they acquire special Human Computer Interaction (HCI) equipment. In this paper we describe a prototype system that aims to provide paraplegics the opportunity to use computer systems without the need for additional invasive hardware. The proposed system is a multi-modal system combining both visual and speech input. Visual input is provided through a standard web camera used for capturing face images showing the user of the computer. Image processing techniques are used for tracking head movements, making it possible to use head motion in order to interact with a computer. Speech input is used for activating commonly used tasks that are normally activated using the mouse or the keyboard. Speech input improves the speed and ease of executing various HCI tasks in a hands free fashion. The performance of the proposed system was evaluated using a number of specially designed test applications. According to the quantitative results, it is possible to perform most HCI tasks with the same ease and accuracy as in the case that a touch pad of a portable computer is used. Currently our system is being used by a number of paraplegics.

1 INTRODUCTION

Conventional Human Computer Interaction (HCI) relies on the use of hands for controlling the mouse and keyboard thus effective HCI is difficult (and in some cases impossible) for paraplegics. With our work we aim to design a system that will enable paraplegics to use a computer system. The proposed system is a multi-modal system that combines both visual input and speech input in order to allow the user to achieve full control of a computer system, in a hands-free fashion.

Visual input is provided through a standard web camera attached on the monitor of the computer. Images showing the user of the system are analysed in order to track his/her head movements. The face tracker activates cursor movements consistent with the detected head motion allowing the user to control cursor movements using head motion. Visual input can also be used for activating mouse clicks and entering text using a virtual keyboard. Figure 1 shows users using a computer system based on the system developed in this project.

Figure 1: Hands-Free HCI based on the proposed system. (In this case a microphone is attached on the camera).

Speech input is provided through a standard microphone attached to the system. In the proposed system speech input can be utilized in two different modes of operation: The Sound Click Mode and the Voice Command Mode. When using the Sound Click Mode, speech input is used only as a means for activating a mouse click. In this case the user only needs to generate a sound in order to activate a click. In the Voice Command Mode we use speech recognition so that the user can verbally request the execution of predefined tasks. Verbal commands handled by the system have been carefully selected in order to minimize the possibility of speech recognition errors and at the same time allow the user to carry out usual HCI tasks efficiently.

Frangeskides F. and Lanitis A. (2006).
MULTI-MODAL HANDS-FREE HUMAN COMPUTER INTERACTION: A PROTOTYPE SYSTEM.
In *Proceedings of the Eighth International Conference on Enterprise Information Systems - HCI*, pages 19-26
DOI: 10.5220/0002443900190026

The algorithms developed as part of the project, formed the basis for developing a prototype hands-free HCI software package. The package contains a program that enables the user to control his/her computer using visual and speech input. The package also includes training and test applications that enable users to become familiar with the system before they use it in real applications. Test applications enable the quantitative assessment of the performance of users when using our system. A number of volunteers tested our system and provided feedback related to the performance of the system. Both the feedback received and quantitative results prove the potential of using our system in real applications.

The remainder of the paper is organized as follows: In section 2 we present a brief overview of the relevant bibliography and in section 3 we describe the proposed system. In section 4 we describe the functionality offered by the proposed system and in section 5 we present the test applications developed for testing the performance of the system. In sections 6 and 7 we present experimental results and concluding comments.

2 LITERATURE REVIEW

Toyama (Toyama1998) describes a face-tracking algorithm that uses Incremental Focus of Attention. In this approach they perform tracking incrementally starting with a layer that just detects skin color and through an incremental approach they introduce more capabilities into the tracker. Motion information, facial geometrical constraints and information related to the appearance of specific facial features are eventually used in the tracking process. Based on this approach they achieve real time robust tracking of facial features and also determine the facial pose in each frame. Information related to the face position and pose is used for moving the cursor on the screen.

Gorodnichy and Roth (Gorodnichy2004) describe a template matching based method for tracking the nose tip in image sequences captured by a web camera. Because the intensities around the nose tip are invariant to changes in facial pose they argue that the nose tip provides a suitable target for face tracking algorithms. In the final implementation cursor movements are controlled by nose movements, thus the user is able to perform mouse operations using nose movements. Gorodnichy and Roth have used the system for several applications like drawing and gaming but they do not provide a quantitative evaluation of the proposed system.

Several commercial head movement-based HCI systems are available (Assistive2005). In most cases head tracking relies on special hardware such as infrared detectors and reflectors (HandsFreeMouse2005) or special helmets (EyeTech2005, Origin2005). Hands free non-invasive systems are also available in the market (CameraMouse2005, MouseVision2005).

Human Computer interaction based on speech has received considerable interest (O'Shaughnessy2003) since it provides a natural way to interact with a machine. However, under some circumstances speech-based HCI can be impractical since it requires quit environments. In several occasions (Potamianos2003) speech recognition algorithms are combined with automatic lip-reading in order to increase the efficiency of speech HCI and make it more robust to speech recognition errors. A number of researchers describe multi-modal HCI systems that combine gesture input and speech. Such systems usually target specific applications involving control of large displays (Kettebekov2001, Krahnstoeve2002). With our system we aim to provide a generic speech-based HCI method rather than supporting a unique application.

3 MULTI-MODAL HCI

We describe herein the face tracking algorithm and the speech processing techniques used as part of the multi modal HCI system.

3.1 Face Tracking

We have developed a face-tracking algorithm based on integral projections. An integral projection (Mateos2003) is a one-dimensional pattern, whose elements are defined as the average of a given set of pixels along a specific direction. Integral projections represent two-dimensional structures in image regions using only two one-dimensional patterns, providing in that way a compact method for region representation. Since during the calculation of integral projections an averaging process takes place, spurious responses in the original image data are eliminated, resulting in a noise free region representation.

In order to perform tracking based on this methodology, we calculate the horizontal and vertical integral projections of the image region to be tracked. Given a new image frame we find the best match between the reference projections and the ones representing image regions located within a predefined search area. The centre of the region where the best match is obtained, defines the location of the region to be tracked in the current frame. This procedure is repeated on each new frame in an image sequence.

The method described above formed the basis of the face-tracking algorithm employed in our system. The face tracker developed, tracks two facial regions – the eye region and the nose region. The nose region and eye region are primarily used for estimating the vertical and horizontal face movement respectively. During the tracker initialisation process the vertical projection of the nose region and the horizontal projection of the eye region are calculated and used as the reference projections during tracking. Once the position of the two regions in an image frame is established, the exact location of the eyes is determined by performing local search in the eye region.

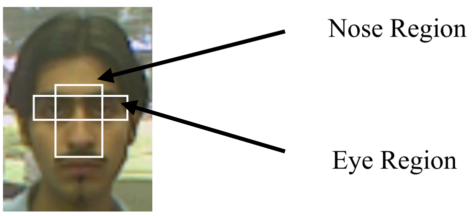


Figure 2: The Nose and Eye Regions.

In order to improve the robustness of the face tracker to variation in lighting, we employ intensity normalization so that global intensity differences between integral projections derived from successive frames are removed. Robustness to face rotation is achieved by estimating the rotation angle of a face in a frame so that the eye and nose regions are rotated prior to the calculation of the integral projections. Constraints related to the relative position of the nose and eye regions are employed in an attempt to improve robustness to occlusion and excessive 3D rotation. In this context deviations of the relative positions of the two regions that violate the statistical constraints pertaining to their relative positioning, are not allowed.

The results of a rigorous experimental evaluation proved that the face tracking algorithm is capable of locating the eyes of subjects in image sequences with less than a pixel mean accuracy, despite the introduction of various destructors such as excessive rotation, occlusion, changes in lighting and changes in expression. Even in the cases that the tracker fails to locate the eyes correctly, the system usually recovers and re-assumes accurate eye-tracking.

3.2 Speech Processing

Instead of implementing our own speech recognition algorithms, we have employed the speech processing functionality offered by the Microsoft Speech Software Development Kit (MicrosoftSpeech2005) that contains the Win32 Speech API (SAPI). SAPI provides libraries with dedicated functions for recording, synthesizing and recognizing speech signals. Our work in this area focuses on the development and testing of a suitable protocol to be used in conjunction with the head-based HCI system in order to allow computer users to achieve efficient hands-free control over a computer system.

We have implemented two methods for using speech input: The Sound Click and the Voice Command Mode.

3.2.1 Sound Click

When the Sound Click mode is active, users activate mouse clicks just by creating a sound. In this mode we continuously record speech input and in the case that an input signal stronger than the background noise is detected, a click is triggered. In this mode any sound of higher intensity than the background is enough to trigger a mouse click, hence this mode is not appropriate for noisy environments. The main advantage of the Sound Click mode is the fast reaction time to user-initiated sounds enabling in that way real-time mouse click activation. Also when using the Sound-click it is not necessary to perform person-specific speech training.

3.2.2 Voice Command

We have utilized speech recognition algorithms available in the SAPI in relation with an appropriate HCI protocol in order to add in our proposed hands-free HCI system, the ability to activate certain tasks by sound. Our ultimate aim is to improve the speed of activating frequently used HCI tasks. Our work in this area focuses on the specification of a suitable set of verbal instructions that can be recognized with high accuracy by the speech recognition algorithm.
All verbal commands supported, have been separated into five groups according to the type of action they refer to. In order to activate a specific command the user has to provide two keywords: The first keyword is used for specifying the group and the second one is used for specifying the exact command he/she wishes to activate. Both the groups and the commands in each group are user configurable – in table 1 we present the default selection of voice commands specified in the system.

Table 1: Voice Commands used in the system.

Group	Command	Description
Mouse	Click	Perform left click
	Right Click	Perform right click
	Drag	Hold left button down
	Drop	Release left button
	Scroll Up	Scroll active window up
	Scroll Down	Scroll active window down
	Stop	Stop face tracker
	Begin	Start face tracker
Move Cursor	Top	Move cursor to the screen position specified
	Top right	
	Top left	
	Bottom	
	Bottom right	
	Bottom left	
	Centre	
Computer	Copy	Copy selected item
	Paste	Paste
	Enter	Press enter
	Close	Close active window
	Shut down	Shut down computer
	Sound	Enable Sound-Click mode
Open	Windows Explorer	Run corresponding application
	Media Player	
	Internet Explorer	
	Keyboard	Run the "On-Screen Keyboard" application

While the Voice Command mode is active the system continuously records sounds. Once the system detects a sound with intensity higher than the background, it attempts to classify the sound to one of the group keywords. If none of group keywords matches the sound, the system rejects the sound. In the case that a sound is recognized as a group-keyword, the system is expecting to receive a second sound corresponding to a sub-command of the activated group. Sounds recorded after a group keyword, are tested against the commands belonging to the corresponding group and if a match is detected the appropriate action is activated. In the case that a match is not detected, the input is rejected.

The main reason for separating the commands in groups is to maximize the robustness of the speech recogniser by reducing the number of candidates to be recognized. Based on the proposed scheme a recorded word is classified only among the five keywords corresponding to each group. Once a correct group keyword is recognized the second word is classified based on the sub-commands for each group, instead of dealing with all system commands. In this way the probability of misclassifications is minimized and at the same time the tolerance of the system to background noise and microphone quality is maximized.

4 SYSTEM DESCRIPTION

In this section we describe how various functions are implemented in the proposed non-invasive human computer interaction system. Those actions refer to system initialisation, system training and simulation of click operations.

4.1 System Initialisation

The first time that a user uses the system he/she is required to go through a training procedure so that the system learns about the visual and speech characteristics of the user. Although it is possible to use the system based on a generic training procedure, the overall system performance is enhanced when person-specific training is adopted. In order to train the face tracker a dedicated tool is used, where the user is requested to keep his/her face still and perform blink actions. Based on a frame-differencing algorithm the positions of the eyes and nose regions are determined and integral projections for those areas are computed. Once the projections are computed the face tracker is activated. The tracker initialisation process requires approximately 10 seconds to be completed. A screen shot of the initialisation tool is shown in figure 3.

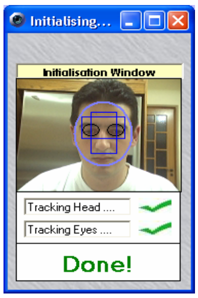


Figure 3: System Initialisation Window.

When using voice input the user is advised to configure the microphone using a dedicated tool provided by the Microsoft Speech Software Development Kit (MicrosoftSpeech2005). Once the microphone is configured it is possible to detect input signals of higher intensity than the background noise. The microphone configuration process requires approximately 30 seconds to be completed.

In the case that the system is used only in the "Sound Click" mode it is not required to perform any person-specific training.

When using the Voice Command mode the user is requested to read a sample text so that the system collects the necessary information required for speech recognition. The training of the speech recogniser is carried out using the Microsoft Speech Recognition Training Wizard, which is available in the Microsoft Speech Software Development Kit (MicrosoftSpeech2005). The speech recognition training process requires approximately 20 minutes to be completed.

It is important to note that system initialisation can be done in a hands free fashion (provided that the camera and microphone are already installed on the system). The tool used for visual initialisation is activated during start up and once the face tracker is in operation the user can use head movements in order to initiate and complete the training for speech processing or activate his/her speech profile in the case of a returning user.

4.2 Activating Mouse Actions

In this section we describe how mouse operations are implemented in our system.

Moving the cursor: The divergence of the face location from the initial location is translated in cursor movement speed, towards the direction of the movement. Based on this approach only minor face movements are required for initiating substantial cursor movement. The sensitivity of the cursor movement can be customized according to the abilities of different users.

When the Voice Command mode is active, it is also possible to move the cursor to predetermined positions by recalling commands from the group "Move Cursor" (see table 1). The use of speech commands is useful for fast initial cursor positioning; usually the cursor position is refined using head movements.

Mouse Click actions: Three different methods for activating mouse click actions are provided. Based on the first method, clicks are activated by the stabilization of the cursor to a certain location for a time period longer than a pre-selected threshold (usually around one second). In this mode, users select the required click action to be activated when the cursor is stabilized. The predefined options include: left click, right click, double click, drag and drop and scroll.

According to the second method, click actions are performed using an external switch attached to the system. In this mode the user directs the mouse to the required location and the appropriate click action is activated based on the external switch. The switch can be activated either by hand, foot or voice (when the Sound Click mode is enabled).

The third method is based on the voice commands available in the group "Mouse" (see table 1).

Text Entry: Text entry is carried out by using the "On-Screen Keyboard" - a utility provided by the Microsoft Windows Operating System (see figure 4). Once the On-Screen Keyboard is activated it allows the user to move the cursor on any of the keys of the keyboard and by clicking actions activate any key. As a result it is possible to use head movements and speech in order to write text or trigger any operation that is usually triggered from the keyboard.

Figure 4: Screen-Shot of the "On Screen Keyboard".

5 HANDS-FREE APPLICATIONS

Although the proposed multi-modal HCI system can be used for any task where the mouse and/or keyboard is currently used, we have developed dedicated computer applications that can be used by prospective users of the system for familiarization and system evaluation purposes. In this section we briefly describe the familiarization and test applications.

5.1 Familiarization Applications

We have developed three familiarization applications: A paint-tool application, a car driving game and a virtual piano. Familiarization applications aim to train users how to move the cursor in a controlled way and how to activate click actions. Screen shots of the familiarization applications are shown in figure 5.

5.2 Test Applications

Test applications are used as a test bench for obtaining quantitative measurements related to the performance of the users of the system. The following test applications have been developed:

Click Test: The user is presented with four squares on the screen. At any time one of those squares is blinking and the user should direct the cursor and click on the blinking square. This process is repeated several times and at the end of the experiment the average time required to direct the cursor and click on a correct square is quoted.

Paint-tool Car racing Virtual piano

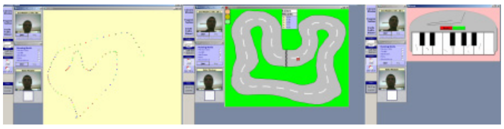


Figure 5: Screen shots of the familiarization applications.

Draw Test: The user is presented with different shapes on the screen (square, triangle and circle) and he/she is asked to move the cursor on the periphery of each shape. The divergence between the actual shape periphery and the periphery drawn by the user is quoted and used for assessing the ability of the user to move the cursor on a predefined trajectory.

Typing Test: The user is presented with a word and he/she is asked to type in the word presented. This procedure is repeated for a number of different randomly selected words. The average time required for typing a correct character is quoted and used for assessing the ability of the user to type text.

Screen shots of the test applications are shown in figure 6.

6 SYSTEM EVALUATION

The test applications presented in the previous section were used for assessing the usefulness of the proposed system. In this section we describe the experiments carried out and present the results.

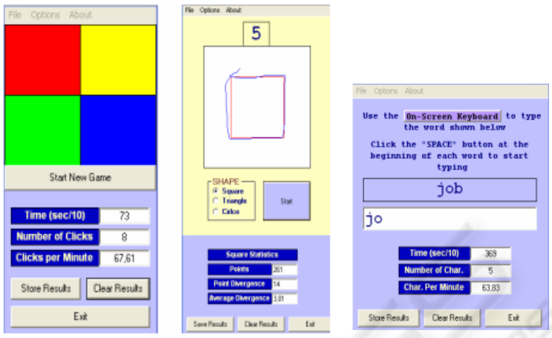
6.1 Experimental Procedure

Twenty volunteers tested our system in order to obtain quantitative results related to the performance of the system. The test procedure for each subject is as follows:

Familiarization stage: The subject is instructed how to use the hands-free computing system and he/she is allowed to get familiar with the system by using the familiarization applications. On average the duration of the familiarization stage was about 15 minutes.

Benchmark performance: The benchmark performance for each volunteer is obtained by allowing the user to complete the test applications using a conventional mouse and a typical touch pad of a portable PC. The performance of the user is assessed on the following tests:

Click Test: The average time required for five clicks is recorded.



Click test Draw test Typing test

Figure 6: Screen shots of the test applications.

Draw Test: The subject is asked to draw a square, a triangle and a circle and the average discrepancy between the actual and the drawn shape is quoted.

Type Test: The user is asked to type five randomly selected 3-letter words and the average time for typing a correct letter is recorded (In this test text input was carried out by using the “On Screen Keyboard”).

Visual test: The user is asked to repeat the procedure used for obtaining the benchmark performance, but in this case he/she runs the test programs using the hands-free computing system based on visual input only.

Visual with an external switch test: The test procedure is repeated, but in this case the user is allowed to use the Hands Free system based on visual input, in conjunction with an external hand-operated switch for activating mouse clicks.

Sound-Click test: For this test clicks are activated based on the Sound-click mode of operation. Cursor movements are carried out based on head motion.

Voice Command test: In this case head motion is used for moving the cursor, but mouse operations are activated using the appropriate commands from the “Mouse” group (see table 1).

The 20 volunteers who tested the system were separated into two groups according to their prior expertise in using the hands-free computing system. Group A contains subjects with more than five hours prior experience in using the hands free system. Subjects from group B used the system only as part of the familiarization stage (for about 15 minutes). All tests were carried out in standard office environments - no precautions for setting up lighting conditions or for minimizing background noise were enforced.

6.2 Results – Discussion

The results of the tests are summarized in the table 2. Based on the results the following conclusions are derived:

Table 2: the results of the quantitative evaluation.

Test	Method	Group A	Group B
Click test (Units: seconds/click)	Mouse	0.76	0.86
	Touch Pad	1.45	2.18
	Visual	3.58	4.84
	Visual + switch	1.41	2.5
	Sound-click	0.98	1.68
	Voice Command	2.10	2.38
Draw Test (Units: Divergence in Pixels)	Mouse	2.62	2.05
	Touch Pad	3.01	4.01
	Visual	3.07	5.93
Typing Test (Units: seconds/click)	Mouse	0.89	0.86
	Touch Pad	1.73	3.41
	Visual	4.39	5.99
	Visual + switch	2.37	3.70
	Sound-click	2.33	3.84
	Voice Command	3.19	5.53

Click test: In all occasions the results obtained by using a conventional mouse are better. When the hands free system is combined with a switch for performing click actions the performance of the system is comparable with the performance achieved when using a touch pad. In the case that the Sound-Click mode is used, the performance of the users compares well with the performance achieved when using a mouse. When the hands free system is not used with an external switch (or Sound-clicks), the performance of the users decreases. The additional delay introduced in this case is mainly due to the requirement for stabilizing the cursor for some time (1 second according to the default setting) in order to activate a click action.

Draw Test: For experienced users of the system (Group A) the performance achieved using the free hand mouse is comparable with the performance achieved when using a touch pad. Subjects from group B (inexperienced users) produced an inferior performance when using the hands free system. The main reason is the reduced ability to control precisely cursor movements due to the limited prior exposure to the system.

Typing Test: In this test the use of mouse or touch pad for typing text is significantly superior to the performance of users using the hands-free system, indicating that the proposed system is not the best alternative for typing applications. However, the performance obtained when using the hands-free system in conjunction with the external switch or Sound-click, is once again comparable to the performance obtained with the touch pad. The main reason for the inferior performance obtained when using the hands free system, is the small size of the keys on the "On Screen Keyboard" that requires precise and well-controlled cursor movements. The ability to precisely move the cursor requires extensive training. Instead of using the "On Screen Keyboard", provided by the Windows operating system, it is possible to use dedicated virtual keyboards with large buttons in order to improve the typing performance achieved when using the hands-free computing system.

User Expertise: The abilities of users to use the hands free system increase significantly with increased practice. Based on the results we can conclude that subjects with increased prior experience in using the hands-free system (from group A) can perform all usual HCI tasks efficiently. It is expected that with increased exposure to the system, users will be able to achieve even better performance.

External Switch: The introduction of an external switch that can be activated either by foot or hand or voice enhances significantly the performance of the system.

Voice Command: When using the Voice Command mode, additional delays are introduced, due to the processing time required for performing the speech recognition task. However, when the task we wish to perform is contained in the Speech Command menu (see table 1), then a speed up in task completion time can be achieved. For example the time required to verbally activate an application among the ones listed in the "Open" group menu (see table 1), is far less than in the case of using the rest of the methods.

7 CONCLUSIONS

We presented a prototype multi-modal hands-free HCI system that relies on head movements and speech input. The proposed system caters for common HCI tasks such as mouse movement, click actions and text entry (in conjunction with the "On Screen Keyboard"). Based on the quantitative results presented, head based HCI cannot be regarded as a substitute for the conventional mouse, since the speed and accuracy of performing most HCI tasks is below the standards achieved when using a conventional mouse. However, in most cases the performance of the proposed system is comparable to the performance obtained when using touch pads of portable computer systems. Even though the accuracy and speed of accomplishing various HCI tasks with a touch pad is less than in the case of using a conventional mouse, a significant number of computer users use regularly touch pads. We are convinced that computer users will also find the proposed hands free computing approach useful.

The proposed system does not require person-specific training, since the system adapts and learns the visual characteristics of the features to be tracked, during the initialisation phase. The only case that person-specific training is required is when the "Voice Command" mode is used. The training procedure in those cases requires about 20 minutes to be completed.

The proposed system is particularly useful for paraplegics with limited (or without) hand mobility. Such users are able to use a computer system based only on head movements and speech input. During the system development phase we have provided the system to members of the Cyprus Paraplegics Organization, who tested the system and provided valuable feedback related to the overall system operation and performance. Currently a number of paraplegic computer users are using the hands-free system described in this paper.

An important feature of the system is the provision of alternative methods for performing a task, so that at any time the user can choose the most appropriate way to perform an action. For example if the user wishes to run the Internet Explorer, he/she has the ability to perform the action using only head movements or by using speech commands or by using a combination of the two input media (i.e move the cursor to the appropriate icon using head movements and run the application by using sound clicks).

In the future we plan to upgrade the Voice Command mode in order to allow text entering based on speech input. Also we plan to stage a large-scale evaluation test is order to obtain concrete conclusions related to the performance of the proposed system. Since the hands-free system is primarily directed towards paraplegics, we plan to include evaluation results from paraplegics in our quantitative evaluation results.

ACKNOWLEDGEMENTS

The work described in this paper was supported by the Cyprus Research Promotion Foundation. We are grateful to members of the Cyprus Paraplegics Organization for their valuable feedback and suggestions.

REFERENCES

Assistive Technology Solutions. Retrieved October 4, 2005, from http://www.abilityhub.com/mouse/

CameraMouse: Hands Free Computer Mouse. Retrieved October 4, 2005, from http://cameramouse.com/

EyeTech Digital Systems-Eye Tracking Device. Retrieved October 4, 2005, from http://www.eyetechds.com/

Hands Free Mouse – Assistive Technology Solution. Retrieved October 4, 2005, from http://eyecontrol.com/smartnav/

Gorodnichy D.O. and Roth G, 2004. Nouse 'Use Your Nose as a Mouse' – Perceptual Vision Technology for Hands-Free Games and Interfaces, Image and Vision Computing, Vol. 22, No 12, pp 931-942.

Kettebekov S, and Sharma R , 2001. Toward Natural Gesture/Speech Control of a Large Display, M. Reed Little and L. Nigay (Eds.): EHCI 2001, LNCS 2254, pp. 221–234, Springer-Verlag Berlin Heidelberg.

Krahnstoever N, Kettebekov S, Yeasin M, Sharma R. 2002. A Real-Time Framework for Natural Multimodal Interaction with Large Screen Displays. Proc. of Fourth Intl. Conference on Multimodal Interfaces (ICMI 2002).

Mateos G.G. 2003. Refining Face Tracking With Integral Projections. Procs. Of the 4th International Conference on Audio and Video-Based Biometric Person Identification, Lecture Notes in Computer Science, Vol 2688, pp 360-368.

Microsoft Speech – Speech SDK5.1 For Windows Applications. Retrieved July 10, 2005, from http://www.microsoft.com/speech/download/sdk51/

Mouse Vision Assistive Technologies, Retrieved October 4, 2005, from http://mousevision.com/

Origin Instruments Corporation. Retrieved October 4, 2005, from http://orin.com/index

O'Shaughnessy, D., 2003. Interacting with computers by voice: automatic speech recognition and synthesis. 1272- 1305.

Potamianos G., Neti C., Gravier G., Garg A., Senior A.W. 2003. Recent advances in the automatic recognition of audiovisual speech. IEEE Proceedings, Vol 91, Issue 9, pp 1306- 1326.

Toyama K. 1998. Look, Ma – No Hands! – Hands Free Cursor Control with Real Time 3D Face Tracking, Procs. Of Workshop on Perceptual User Interfaces, pp. 49-54.