Hands-Free Mouse Control System for Handicapped Operators

Alexey Karpov¹ and Alexandre Cadiou²

(1) Speech Informatics Group, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS), Russia

karpov@iias.spb.su

(2) Advanced Institute of Electronics of Paris (ISEP), France

alexandre.cadiou@isep.fr

Abstract

The paper describes an evolution of the multimodal system ICANDO (Intellectual Computer Assistant for Disabled Operators) intended for assistance to persons without hands or with disabilities of their hands or arms for human-computer interaction. The evolution of this device relates to the head movements tracking system. The system ICANDO combines the modules for automatic speech recognition and head tracking in one multimodal system. The architecture of the system, methods for head detection and tracking, experiments with hands-free mouse control are described in the paper. The obtained results described at the end of the paper demonstrate an acquisition of increased reliability and faster operation of the ICANDO system.

1. Introduction

The first version of the ICANDO system was presented and demonstrated in 2005 [1]. This paper gives the description of system’s improvements which were made with the system during last year. The following sections of the paper present the evolution of the head tracking system with the problems encountered in the initial version of ICANDO and the solutions brought to solve these problems as well as the results of experiments with the last version.

In ICANDO system two natural input modalities are used: speech and head movements. As both modalities are active ones [2], their input must be controlled continuously (non-stop) by the system. Figure 1 shows the common of ICANDO multimodal system.

![Figure 1. Architecture of the assistive system ICANDO](image)

The system processes human’s speech and head movements in parallel and then combines both informational streams in joint multimodal command, which is used for work with GUI of a computer. Each of the modalities transmits own semantic information: head (nose) position indicates the coordinates of some marker (cursor) in a current time moment, and speech signal transmits the information about meaning of the action, which must be performed with an object selected by the cursor (or irrespectively to the cursor position). Russian speech recognition is applied in the system, for this purpose the SIRIUS engine (SPIIRAS Interface for Recognition and Integral Understanding of Speech) is used [3]. The list of voice commands includes 40 commands for control of GUI of MS Windows (for instance, “left button”, “right button”, “print”, “open”, etc.).

The synchronization of two information streams is made by the speech recognition module, which gives the special signals for storing of the mouse cursor coordinates calculated by the head tracking module, and for multimodal fusion. The methods of information fusion and synchronization are described [4]. The current paper presents details of the approach for hands-free mouse control based on facial points tracking.

2. Hands-free mouse cursor control

This section describes the head tracking technologies intended for tracking the operator’s head motion instead of hand-controlling motions. It contains the evolution of the used head tracking approach namely a presentation of the problems encountered with the first version of ICANDO and the solutions brought to increase the robustness of the ICANDO system. These solutions are firstly the multiplication of the tracking points on the operator’s face, secondly a method of rectangles restoring the tracking points in their operational areas on the operator’s face to keep the system efficient, thirdly a method which adapts the velocity of mouse cursor according to the velocity of the operator’s head in order to increase the ergonomics of the system.

2.1. Hardware devices for mouse control without hands

The head tracking can be performed by two diverse ways: hardware and software-based methods. In the hardware techniques a user should wear some special devices on his head. At present there exist several hardware systems for head tracking in computer market (some examples are presented on Figure 2). For instance, the NaturalPoint company presented the SmartNAV device for hands-free mouse. This system consists of the special transmitter-receiver device working in the infrared mode and several reflective marks, which should...
be attached to the face of a user or to the special hat. The company InterSence produces professional trackers InterTrax for helmets of virtual reality or computer stereo glasses. Inside of this device there exists a gyroscope, which allows tracking the orientation of a head. Also hardware trackers can be applied based on special device with light emitting diodes and a video-camera [5].

![Hardware-based head tracking systems](image)

Figure 2. Hardware-based head tracking systems

However, these devices are very expensive and their cost varies from several hundred till thousand euros. It is one of the reasons why they are not popular in assistive systems for impaired users. The second reason is that users find inconveniently to wear special device on the head during work with a computer.

2.2. Developed system for hands-free mouse control

A software method for tracking operator’s head movements is applied in the system. It is based on the free available software library Intel OpenCV (Open Source Computer Vision Library) [6]. This library realizes many known algorithms for video processing.

For video processing USB web-camera Logitech QuickCam for Notebooks Pro with resolution 640x480 and 30 fps is applied. The usage of a professional digital camera (Sony DCR-PC1000E was tested in some experiments) provides better accuracy of tracking, but taking into account that the system should be available for most users, we apply camera of low-end class with the price under 50 euros.

The special approach in the first version of ICANDO device was developed for control of the mouse cursor, which is able to work in real-time mode. It includes two stages of functioning: calibration and tracking. At first short starting stage the position of face in the video is defined. It is realized by the software module which uses the Haar-based object detector to find rectangular regions in the given image that likely can contain face of a human [7]. This region should not be less than 150 per 150 points that allows accelerating video processing.

Then taking into account the standard proportions of a human’s face the approximate position of nose is marked by a green point on the image. During several seconds of calibration process a user should combine the tip of his nose with the position of this green point. Then this point is captured by the system and the tracking algorithm is started. In the first version of ICANDO, it was determined experimentally that the most suitable point on a face for tracking is the tip of nose. It is the center of the face and when an operator make any head gestures (turn to the right, left, up or down) the position of the tip of nose is moved to this direction and it can indicate the position of mouse cursor [8].

The main problem of this system is the lack of robustness. Indeed the algorithm uses the iterative Lucas and Kanade technique for optical flow [9], which is an apparent movement of image brightness and the algorithm often loses the position of human’s nose that is caused by the lack of light or very quick movements of user’s head. No method was designed previously to restore the tracking point on the tip of nose during the use of ICANDO device. The only solution to solve this problem was to introduce the special voice command “Calibration” in the system, which runs the process of calibration described above. As the consequence, the operator has to stop his work to calibrate his face as soon as the algorithm loses the position of human’s tip of nose and it is not a convenient use of the device.

2.2.1. The multiplication of tracking points and rectangles method

In order to increase the robustness of the head tracking system of the last version of ICANDO, the number of tracking points has been multiplied on the face. The different positions of the new tracking points are justified, namely a vertical line is constituted by choosing the upper lip (point 1), the tip of nose (point 2) and the middle of both eyes (point 3), an horizontal line is constituted with the point 3 and the eyes (point 4 for the left eye and point 5 for the right eye) as described on the Figure 3. The proportions of the face vary from one operator to another, that’s why it is possible to adjust the position of the tracking points on the operator’s face by modifying some parameters which describe the proportions of user’s face, in order to personalize the system. Moreover, it was found that for some users with light eyes, the tracking points on eyes are not stable for tracking, thus for such users two these points can be discarded.

![Positions of tracking points](image)

Figure 3. Positions of tracking points

To keep these tracking points operational, a method has been implemented in order to restore all of them at their operation positions. This method uses rectangles which define thanks to a couple of tracking points assumed fixed an operation area in which the third tracking point has to remain to stay operational. Experimentally, it was observed that the tracking points move slightly on the face and independently. It is unlikely to lose all tracking points at the same moment.
That’s why it is assumed that only one tracking point can move during that the others two tracking points which define the rectangle which restores this moving point remain fixed on the face. Under this condition only the rectangles method is efficient and the experiments confirm this assumption. For instance, if the point 2 has to be maintained in its operation area, a vertical rectangle is defined with the coordinates of the points 1 and 3. The vertical borders of this rectangle define the operational area in which the point 2 has to be maintained to remain operational for an efficient use of the device. For that, as soon as the values of the coordinates of the point 2 show that this point is moving out of the rectangle or is not roughly in the middle of the vertical segment defined by the points 1 and 3, the algorithm affects to the couple of variables (X-point 2, Y-point 2) defining the position of the point 2 two linear combinations of the X-coordinates and Y-coordinates of the points 1 and 3 to restore the point 2 in its operational position, as shown on the Figure 4.

One rectangle is defined to restore each tracking point on the face in its operational position, excepted for the point 3 which is common to the vertical line and the horizontal line as shown on the Figure 3.

If the point 3 moves out of the vertical rectangle defined by the point 2 and the point 1 or out of the horizontal rectangle defined by the points 4 and 5, a linear combination of the X-coordinates of the points 1 and 2 is affected to the X-variable of the point 3, and a linear combination of the Y-coordinates of the points 4 and 5 is affected to the Y-variable of the point 3 as shown on the Figure 5.

So, the restoration process of the point 3 is divided into the horizontal rectangle defined by the points 4 and 5 and the vertical rectangle defined by the points 2 and 1, then since the tracking points move on the face independently from each others and not all at the same time, the restoration process of the point 3 is the most robust of those of all tracking points on the face. Moreover, the middle of both eyes is the point of the face which moves with the lowest magnitude during the use of the ICANDO device and it is located in an area with different brightness. So this point is the most robust of all the tracking points and it defines a tracking point reference on the operator face. It is the key of the head tracking system of the new version of ICANDO because this tracking point is one of the two tracking points which define all the rectangles to restore all the others tracking point. Then it is assumed that as long as the restoration process will maintain the point 3 in its operational position, the head tracking system will remain stable and operational. On the whole, there are 6 rectangles.

Then, the averages of the (X, Y)-coordinates of the points 1, 2 and 3 whose movements indicate the position of the mouse cursor on a desktop are affected to the mouse cursor’s variables. The points 4 and 5 are not taken into account for the mouse cursor position because the coordinate system which is chosen is a Cartesian coordinate system and not a polar coordinate system. Indeed the points 4 and 5 being far from the middle of the face, when the operator rotates his head on the left or on the right the movement of eyes is not linear but rotary, consequently those points are not representative of the real movement of the operator’s head in a Cartesian coordinate system. They are used only for the restoration process of the head tracking system.

2.2.2. Method of velocity adaptation

All valid computer operators constantly adapt the velocity and the magnitude of the movements of the mouse according to their needs, and this without realizing it. Several scenarios of mouse’s use must be considered. For instance, when the mouse cursor is located in one of the corners of the screen and an operator wishes to move the mouse cursor to the opposite corner, the operator will move the mouse amply and quickly. In opposition, the user can keep the mouse cursor located in a reduced area of the screen namely for instance the up left corner of the screen where are usually located most of windows on the desktop, and carry out reduced and slow movements with the mouse to guide the mouse cursor from an item to another close item. Of course, a valid operator is not limited by the evolution capabilities of the mouse, it is very convenient to carry out fast and full movements with the mouse and conversely.

The current case for disabled people is totally different because they use the movements of their head to guide the mouse cursor so they haven’t the same freedom of movements than the valid operators. Indeed the operator’s head cannot evolve amply to guide the cursor from a corner of the screen to the opposite one. We have faced a problem of mouse cursor stability in the first version of ICANDO. Only one velocity was available for the mouse cursor. This velocity was a compromise between fast and slow to comply with all
possible situations faced with the users. But this velocity was not adapted to really slow movements of the mouse cursor required by a precise task as to draw a picture in a graphical editor because the mouse cursor moved too quickly and could not remain stable. Moreover when an operator wished to go through the screen quickly with the mouse cursor, the velocity of cursor was not high enough and obliged the operator’s head to move amply. Because the magnitude of the movements of head’s operators is limited, it was not convenient with a single velocity of the mouse cursor to traverse the whole screen.

So a method has been designed to adapt the velocity of the mouse cursor to the needs of disabled operators. This method is based on the velocity measurement of the movements of operator’s head.

To perform this method, the algorithm measures the number of pixels traversed by the point 2 tracked by the webcam which has a fixed resolution of 640x480 between two consecutive frames. If this number of pixels is higher than a certain threshold, it means that the operator’s head moves quickly namely he wishes to traverse a large distance on the screen with the mouse cursor, so the algorithm will increase the velocity of the mouse cursor to reduce the magnitude of the movements of operator’s head required for a such use of the mouse cursor. If this number of pixels is lower than a certain threshold, it means that the operator’s head moves slowly so the movements of the mouse cursor have to be precise and slow to keep the mouse cursor stable. The velocity of the mouse cursor is defined like a number of pixels of the screen it can traversed between two consecutive frames. The screen has an adjustable resolution. This number of pixels represents a different distance for different resolution of the screen. Therefore in the last version of ICANDO in order to comply with all needs of the disabled operators in all possible situations it has been defined three different velocities for the mouse cursor according to the velocity of the operator’s head and the resolution of the screen chosen. These three different velocities have been chosen during practiced tests to reproduce with the movements of disabled operators’ head the same ergonomics of the mouse of a valid user. The different velocities available for the mouse cursor are adjusted automatically according to the resolution of the screen chosen, namely on a low resolution a fast velocity is not required in opposition to a high resolution.

This method generates another problem to which we have to face. Indeed the operator’s head and the mouse cursor can be shifted because of the consecutive changes of the velocity of the operator’s head and the velocity of the mouse cursor. The ratio of the velocity of the operator’s head and the velocity of the mouse cursor is not constant; therefore it was observed that the position of the cursor on the screen is sometimes not significant of the orientation of the operator’s head especially when the mouse cursor moves a quarter of the screen to the opposite one at different velocities. To solve this problem, the screen has been divided into four quarters, and it is assumed that when the point 2 crosses one of the two lines dividing the screen in four quarters, the mouse cursor has to be on one of these two lines, therefore the area in which the mouse cursor can be shifted with the orientation of the operator’s head is reduced. After tests, it has been observed that this shift is not troublesome in a reduced area like a quarter of the screen.

3. Experimental results

As the hardware for hands-free computer control the miniature web-camera Logitech QuickCam for Notebooks Pro is used. This camera provides a video signal in 640x480x30fps.

The testing of the system was fulfilled by an independent tester. The system was tested on the task of drawing a picture with the MS Paint editor. It was connected with a compromise of precise work and use of the entire screen to emphasize the efficiency of the method of velocity adaptation and the gain of robustness of the last head tracking system described above. To show the improvements of the ICANDO device, due to these methods, a complex picture with both versions of ICANDO (see Figure 6) has been drawn.

Several colors and shapes were drawn to show that it is possible to select in the toolbox of MS Paint editor different items with an acceptable precision. The voice command “left” was used to click the left button of mouse and the command “calibration” for starting the recalibration process when a user find that tracking point were lost by the system.

![Figure 6. Sample of the picture for the system testing](image)

The results of experiments are presented and compared in the Table 1. Several parameters are taken into account namely: the duration of the experiment; the number of recalibration during the experiment; and the maximum resolution of the screen chosen as shown on the table 1.

Table 1. The comparison of two version of the system

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Old version</th>
<th>New version</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time to fulfill scenario (minutes)</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Number of recalibrations (times)</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Maximal resolution (pixels)</td>
<td>800x600</td>
<td>1280x1024</td>
</tr>
</tbody>
</table>

It has been observed that the draw of the picture is faster with the last version of ICANDO due to the fact that no tracking points were lost so the user doesn’t need to recalibrate his face during the experiment. The method of adaptation velocity provides to the user a better precision and flexibility in use. The last version supports a higher resolution of the screen due to the multitude of cursor velocities. So it could be use on diverse resolutions. Globally, even if the
same quality of drawing can be reached with the first version, the use of the new version is definitely more comfortable and faster than the first one.

Real work of the multimodal system for hands-free computer control based on speech recognition and head tracking was shown in the main Russian TV channel (“First channel”) in the news program (“Vremja”) on 6 September 2005. During the demonstration the impaired person successfully worked with a PC by ICANDO system (see “http://www.1tv.ru/owa/win/orb6_main?p_news_title_id=82825&p_news_rzdel_id=4”).

The additional video fragments of testing of ICANDO assistive multimodal system are available at the web site of the Speech Informatics Group of SPIIRAS (see “http://www.spioras.nw.ru/speech/demo/assistive.html”).

4. Conclusions

The presented assistive multimodal system ICANDO is aimed mainly for handicapped operators, which have problems using a computer keyboard and a mouse. The human-computer interaction is performed by voice and head movements. The system uses standard web-camera, which provides video and audio signal with acceptable quality. It simplifies the usage of the system, since no any additional hardware (like microphone or helmet) is required for a user. The experiments have shown that the last version of the head tracking system described above is more robust and permits disabled operators to work quickly and in a relaxed manner. Any task of GUI control can be accomplished with the last version of ICANDO as to draw a complex picture with MS Paint editor or to perform a task like Internet communications and work with text documents. Thus the obtained results allow concluding that the assistive multimodal system ICANDO can be successfully used for hands-free work with a personal computer for users with disabilities of their hands.

5. Acknowledgements

This research is supported and funded by the EC-FP6-IST Program in framework of SIMILAR Network of Excellence contract № 507609.

6. References


