International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:7, No:6, 2013

Mouse Pointer Tracking with Eyes

H. Mhamdi, N. Hamrouni, A. Temimi, and M. Bouhlel

***Abstract*—**In this article, we expose our research work in Human-machine Interaction. The research consists in manipulating the workspace by eyes. We present some of our results, in particular the detection of eyes and the mouse actions recognition. Indeed, the handicaped user becomes able to interact with the machine in a more intuitive way in diverse applications and contexts. To test our application we have chooses to work in real time on videos captured by a camera placed in front of the user.

***Keywords*—**Computer vision, Face and Eyes Detection, Mouse pointer recognition.

I. Introduction

ONE of the characteristics of the human is to be able to acquire images, through the eye, as information, then to be able to interpret it with the brain. The challenge for computer vision, subject which we are going to treat handle, about which we will treat, is to allow a computer to "see". That is to say, as humans, to retrieve information via an image acquisition and exploit it. So, the machine will be then able to recognizing the eyes of a user to replace the device of interaction, the mouse, by the latter. We shall see that this analysis is far from being immediate and that the image has to undergo a preprocessing with the aim of simplifying at most the work of the computer

Eye motion is an indicator of system performance. So we can at every moment know how a person's experience structure, and how it represents.

In this paper, we focus on the comparison of two approaches in order to simplify the handling of the machine for the disabled.

The first approach is based on the global appearance by applying first of all the Haar descriptor to detect the face and the eyes without passing by the phase of segmentation. We have to use in the continuation the functions proposed in the library OpenCV to move the mouse cursor and it ends with the location of the left and right eye to recognize the actions provided by the eyes and processed by the computer.

A comparison between both approaches made afterward at the level of time of answer and the space of research.

For both approaches the stages to be followed are presented in the Fig. 1.

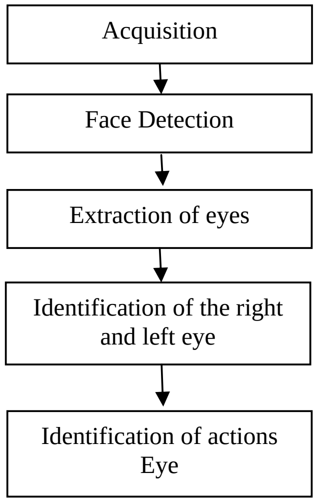

Fig. 1 Diagram of the motion detection algorithm

II. Approach Based On Global Appearance

In this approach which based on techniques such as the statistical analysis and the machine learning to find the appropriate characteristics of the images of face and the images of non-face, we apply the descriptors of haar-like.

Paul Viola and al. [1] proposed a quick and strong way for the detection of objects and especially the face. It based essentially on three contributions:

- Haar-like Descriptors: simple primitives, sensitive to the presence of edges, bars, and other simple image structure.
- The integral image: image representation which allows calculating descriptors Haar-like very quickly.
- The cascade structure: the structure of the global classification function, a combination of functions that allows a low classifications rapid rejection of non-face regions.

Based on Haar-like descriptors, this method joins the performance to the simplicity. Indeed it guarantees a rate of acceptable detection with a rate of relatively low error.

A. Face Detection and Eyes Localization

The detection of the face is at present implemented in the library OpenCV and relaying on methods of data classification.

The image is available in different size for each resolution;

H. Mhamdi is with Research Unit Sciences and Technologies of Image and Telecommunication, Sfax University, (e-mail: mhemdi.halima@gmail.com).

N. Hamrouni is with Research Unit Sciences and Technologies of Image and Telecommunication, Sfax University, (e-mail: neswebhamrouni@yahoo.fr).

A. Temimi is with Research Unit Sciences and Technologies of Image and Telecommunication, Sfax University, (e-mail: teay2006@yahoo.fr).

M. Salim Bouhlel is with the research Unit Sciences and Technologies of image and Telecommunications, University of Sfax, High Engineering School of Sfax (ENIS), (e-mail: medsalim.bouhlel@enis.rnu.tn).

window size common to each image is scanned at random

This window is going to be then used to cut one under image and to begin the analysis of human face. By a process of learning, certain number of "boxes" is defined to add pixels and define a typical value characterizing the human face. These critical areas have been made automatically by algorithms by submitting a large number of individuals.

Fig. 1 Face detection by the Haar descriptor

For the location of the user's eyes, we detected the position of the eye from the image provided by the camera. The system used is the early detecting system of the object based on a cascade of classifiers using simple Haar descriptors to detect eyes. This method was initially proposed by Paul Viola [2] and improved by Rainer Lienhart [3].

First of all, classifiers are built with thousands of positive and negative images (images of popular objects and images of not objects) by using simple descriptors [2].

There are a large number of descriptors in a sub-window of the image 24x24 pixels (117,941 descriptors) [3] that is a number much bigger than counts it of pixels.

The algorithm of the construction held is AdaBoost [1]: it selects small set descriptors allowing separating at best the positive and negative examples and creates classifiers based on selected descriptors.

After the training of the simple classifiers, cascade classifier is developed to improve the performance of detection while reducing radically calculation time.

The cascade classifier is constructed to reject a large number of sub-windows to detect negative and most positive sub-windows.

Fig. 2. Eye detection with Haar descriptor

In our application, we use the cascade classifier which is proposed by OpenCV and which was elaborated with 7000 positive samples of the eye. The result of this classifier is efficient for the detection, real-time eye from images provided by the camera. The result of eye detection is shown in Fig. 2.

By using this classifier to follow the real time eye, the treatment consisting in looking for the eye in all the successive images provided by the camera, lead in too important calculation time.

B. Determination of Eyes

By studying more attentively the box including of the detection of face, it was noticed that the size of the face was proportional in the size of the box. As the recognition works only of face, it is then possible to use the geometry of the face to deduct the position from it of eyes.

The detection of the left eye (Fig. 3) will be made in a box of width 1/3 of the image and the height 1/4 of the image thus the position is 1/6 of the width and 1/4 of the height. So, we can use the classifier to detect it easily.

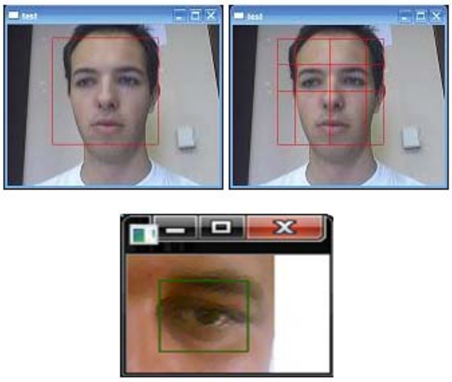

Fig. 3 Detection of the left eye

The right eye (Fig. 4) is determined in the same way that the left eye, except that the position of the including rectangle begins in the middle of the image.

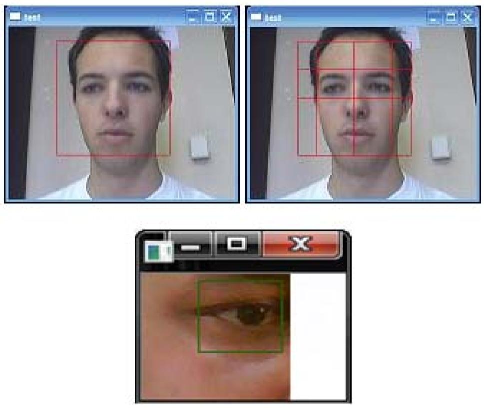

Fig. 4 Detection of the right eye

C. Move the Mouse Pointer and Automatic Eyes Actions Recognition

Having located the observer eyes; we reduce the space of analysis to the found position, to avoid studying the various pixels on the entire image every time. When the user makes a too big movement, what is translated by a big variation of pixel from an image to the other one, we reset the zone of analysis to the whole image perceived collected by the camera.

The vector of movement is initialized every time during a movement of eyes by making the difference between the values of the current pixels of eyes and the values of the pixels of the initial position. This vector can be reset in every new image.

The motion of the mouse pointer is thus established with regard to the eyes movement.

It allows, for example people with motor problems, such as quadriplegics to control a computer by blinking his eyes through a webcam.

Fig. 5 Motion vector initialization

The movement of the mouse by eyes is very important for the disabled peoples in particular for the quadriplegics persons, but it does not seem very appropriate without using the blinking of eye as means of communication

Thus the eye represents an interesting means by identifying its state (opened, closed) and to adapt it in function the answers of the machine or by closing for example a file when the eye right is opened and the left one is closed during a well determined time. As illustrates in Fig. 6.

Therefore, both eyes play the role of the right and left button of the mouse.

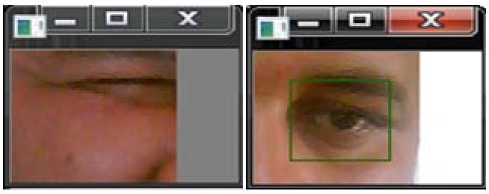

Fig. 6 Identification of the condition of the eyes to close a file

The new functions of eyes allow us to communicate better with the machine during diverse reactions let us quote:

- Copy a file.
- Open a file.
- Delete a file.
- Move a file.

III. The Approach Based On Invariable Characteristics

The main idea of this approach is to find the structural characteristics which exist even when the pose, the angle of view, or the condition of lighting change, then they use these invariable characteristics to localize faces.

A. Segmentation of the Skin Color

For the segmentation of the skin color, there are two techniques: the segmentation by growth of regions and thresholding. The latter is used for the detection of face.

The principle of thresholding is to compare the intensity of levels of grey of a pixel with a threshold to attribute it to one of the class of image.

In our case, the segmentation is done by scanning all the pixels of the image to search for those observing the desired threshold according to the color space used. Pixel belonging to the thresholding interval desired takes the value "1" otherwise it takes the value "0". This process produces a binary image highlighting segments with flesh color (white); the others having no skin color will take the black color [4], [5].

Detection techniques based on the analysis of skin color are fast and robust methods. They reduce the search space of the face region in the image. Several color spaces can be used to detect, in the image, the pixels having the color of the skin. The detection efficiency depends mainly on the selected color space. There are several color spaces for image; we will define the space Y CrCb.

We are going to make a conversion of the original image in space RVB towards the space Y CrCb. Fig. 7 represents the obtained results.

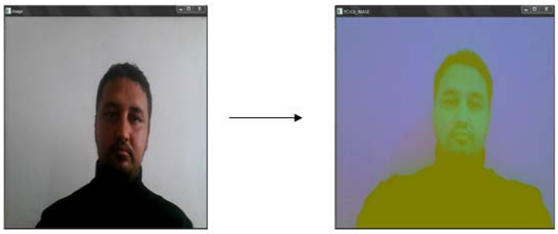

Fig. 7 Image converted to YCrCb space

After a threshold ($130 < Cr < 170$ and $70 < Cb < 125$) is applied to obtain an image segmented in black and white, where from the white color represents the skin color and the black color represents the rest. So, to improve the quality of the obtained image and to avoid the noise the median filter is used.

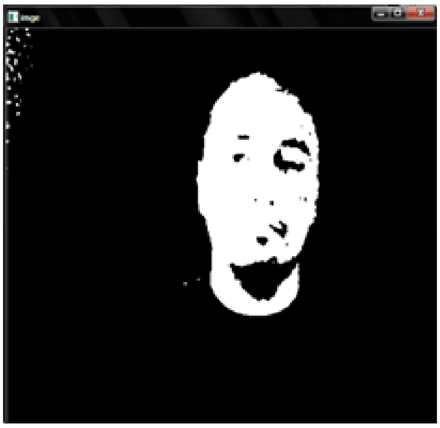

Fig. 8 Segmented Image

We notice that the skin color of the face is detected in spite of the movement of the face, and the conditions of lightings.

B. Face and Eyes Detection

To localize the face and the eyes in the segmented image, it is enough to detect the characteristic points having the shape of ellipse [6]. In this case, the face and the eyes are detected.

Fig. 9 Détection des ellipses

C. Mouse Pointer Motion and Action Eyes Recognition

The same principle of the approach 1 is applied to the second approach using the motion by a subtraction between the current pixel coordinates and pixel coordinates that are the initial position of the eyes.

The machine can not respond with operations allowing the user to better communicate with the machine. This problem is due to the complexity of eye detection.

IV. Comparison

The detection of the real time eyes postpones from an approach to another one.

The criterion for comparison is based on the segmentation, face and eyes detection, mouse pointer motion and the automatic eyes actions recognition in real time. We try to determine the best approach to achieve our goal. In our comparison we took the images from the webcam.

According to the test we did using the first approach, we find that the segmentation phase does not exist instead of the second approach where we apply the segmentation of skin color thresholding in YCrCb space.

We note, again, after the test the first approach applies Haar descriptor for the detection of face and eyes, which easily allows the detection of face and eyes while the second approach applies a method of detecting ellipses to locate the face that appears a bit difficult.

In addition, the first approach raises a problem at speed of mouse pointer that it can see a long but it is effective in real-time response operations provided by the user (open, close, copy, .. .) instead of the second approach where the movement is more dynamic but it is a problem at the sending requests to the machine.

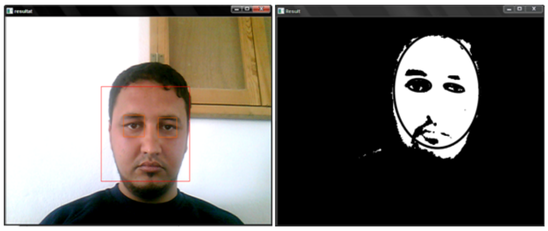

Fig. 10 Face and eyes detection

Visual comparison between the approach based on the global appearance and approach based on invariant characteristics shows that the first approach is more efficient, because the response time is shorter for the eyes actions recognition. In contrast, the second method gives a good result in reducing the search space to locate the face and eyes but poses problems in motion mouse pointer and eyes actions recognition by the machine. More in, we see that the effect of the lighting affects the performance of both approaches.

V. Conclusion

The human-computer interaction defines the methods and tools implemented, so that humans can control and communicate with a machine. Engineers in the field studying how humans interact with computers and each other using computers as well as how to design systems that are ergonomic, efficient, easy to use, or more generally adapted to their environment Usage.

The study of our work was to define a tool for a human to control and communicate with a machine. We then tried to replace the basic functionality of interaction device (open a folder, select, copy ...) through the eyes. To achieve our goal, we used an approach based on the global appearance specifically using Haar-Like method as well as the approach based on invariant features using the method of skin color. Similarly, we used two approaches for the task to move the mouse pointer with the eyes after identifying their condition and adapt based on the responses of the machine.

REFERENCES

[1] P. Viola and M. J. Jones, “Robust Real-Time Face Detection”, International Journal of Computer Vision, Vol. 57, No. 2, pp. 137-154, 2004
[2] P. Viola and M. Jones, “Rapid Object Detection Using a Boosted Cascade of Simple Features”, vol. 1, 2001
[3] R. Lienhart and J. Maydt, An Extended Set of Haar-like Features for Rapid Object Detection, vol. 2, no. 1, pp. 900-903, 2002.
[4] H.P. Graf, T. Chen, E. Petajan, E. Cosatto, "Locating Faces and Facial Parts". Proc. First Intl Workshop Automatic Face and Gesture Recognition, pp. 41-46, 1995.
[5] H.P. Graf, E. Cosatto, D. Gibbon, M. Kocheisen, et E. Petajan, "Multimodal System for Locating Heads and Faces", Proc. Second Intl Conf. Automatic Face and Gesture Recognition, pp. 88-93, 1996.
[6] K. Sobottka, I. Pitas, "Face Localization and Feature Extraction Based on Shape and Color Information", Proc. IEEE Int’l Conf. Image Processing, pp 483-486, 1996.