Scientific Bulletin of the „Petru Maior" University of Tîrgu Mureş
Vol. 11 (XXVIII) no. 1, 2014
ISSN-L 1841-9267 (Print), ISSN 2285-438X (Online), ISSN 2286-3184 (CD-ROM)

HAND GESTURES MOUSE CURSOR CONTROL

Marian-Avram VINCZE[1], Horațiu-Ştefan GRIF[2]
[1,2] "Petru Maior" University of Tîrgu Mureş
Nicolae Iorga Street, no.1, 540088, Tîrgu Mureş, Romania
[1]vincze.marian.avram@gmail.com
[2]horatiu.grif@ing.upm.ro

Abstract

The paper describes the implementation of a human-computer interface for controlling the mouse cursor. The test reveal the fact: a low-cost web camera some processing algorithms are quite enough to control the mouse cursor on computers. Even if the system is influenced by the illuminance level on the plane of the hand, the current study may represent a start point for some studies on the hand tracking and gesture recognition field.

Keywords: human-computer interface, mouse cursor, hand tracking, hand gesture, HSV color model

1. Introduction

In Human-Computer Interaction (HCI), traditional input devices such as keyboard, mouse and joystick become less suitable in the interaction between human and computer since it is more natural for Human-Human Communication (HHC) to communicate with his/her body. Especially, hands and fingers have a significant role in HHC. They used to accompany or even replace spoken communication in some cases, such as in sign languages, where hands completely replace spoken communication. Nowadays, computer vision gives a lot of potential for motion capturing without to put limits to the user and this offers the possibility to create more natural gestural interfaces.[14] The ability to recognize hand gestures is indispensable and important for successful interpersonal social interaction. Vision-based hand gesture recognition, enabling computers to understand hand gestures as humans do, is an important technology for intelligent HCI. Therefore, visual analysis of hand gestures has attracted much attention in the last decade, and much progress has been made. [13] Reference [8] introduces a gesture recognition system based on image blocking where the gestures are recognized using a suggested brightness factor matching algorithm. In [13] is presented the realization of a real-time hand tracking system in dynamic environments for a wheelchair using a mean shift embedded particle filter. In [9] the authors propose a method for the recognition of the gestures of two hands using the depth map of the frame. The depth map is calculated in hardware level by a VisionST stereo camera. In [11] is proposed and implemented an application for objects manipulating in virtual environment using the hand gestures. In [15] the authors present a gesture recognition system based on the accelerometer from a mobile device. The system able to recognize a collection of 10 different hand gestures finally is configured to control the actions of a wheeled robot in an indoor space. In [6] is described a motion tracking system based on integrated accelerometers which is tested to control the navigation in virtual spaces.

2. Color models

Gonzales and Woods (2001) define the color model (color space or color system) as a specification of a coordinate system and a subspace within that system where each color is represented by a single point. The RGB color model is the "*most commonly used and popular color space*". (Sarifuddin 2005, p. 1) In this model each color is represented using the three additive primaries colors: red, green and blue. [3], [7] Combined additively the spectral components of these colors will produce a resultant color. [3] According to the definitions mentioned above the RGB model uses as a coordinate system a Cartesian coordinate system and as a color subspace a cube (fig. 1) which has three corners colored in red, green and blue and the other corners are colored in cyan, magenta, yellow, black and white. Black is the origin of RGB system, and white is situated diagonally opposite the black color. Each color of this model is defined as a point on the surface of the RGB cube or as a point within the RGB cube. [7] [1] The straight line which join the black corner with white corner define the gray scale level [7] (each points from this line has equal values for the three components RGB).

According to "*this space presents some limitations: (i) the presence of a negative part in the spectra, which does not allow the representation of certain colors by a superposition of the three spectra,* *(ii) the difficulty to determine color features like the presence or the absence of a given color, and (iii) the inability of the Euclidean distance to correctly capture color differences in the RGB space.*"(Sarifuddin 2005, p. 1) the authors of the present paper have considered a perceptual color space [4]. More accurate, they chose the HSV color model (fig. 2) because the color components are relate to the human perception of color, saturation, and luminance [5].

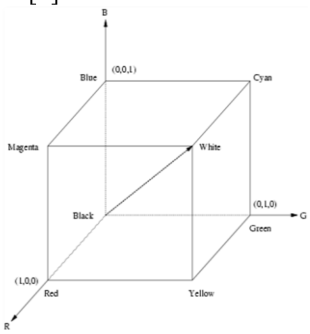

Fig. 1. RGB color model [7][1]

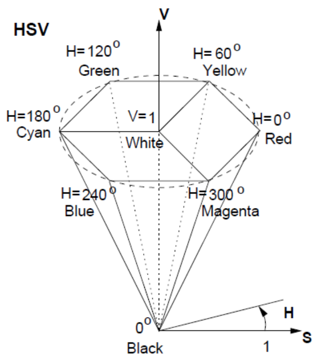

Fig. 2: HSV color model [4]

The V, S and H components of the HSV model are computed by [1], [5], [10]

$$V = max\ (R, G, B) \quad (1)$$

$$S = \begin{cases} 0, \ if \ V = 0 \\ \dfrac{V - X}{V}, \ otherwise \end{cases} \quad (2)$$

where

$$X = min\ (R, G, B) \quad (3)$$

$$H = \begin{cases} 0, \ if \ V = X \\ \left(60^o \times \dfrac{G - B}{V - X} + 360^o\right) \bmod 360^o, \ if \ V = R \\ 60^o \times \dfrac{B - R}{V - X} + 120^o, \ if \ V = G \\ 60^o \times \dfrac{R - G}{V - X} + 240^o, \ if \ V = B \end{cases} \quad (4)$$

where R, G and B are the components of the RGB models with values on the interval [0,1].

3. The system architecture

The system architecture is presented in fig.3 and is composed by three blocks:

- Image acquisition block;
- Hand tracking and gesture recognition block;
- Mouse cursor control.

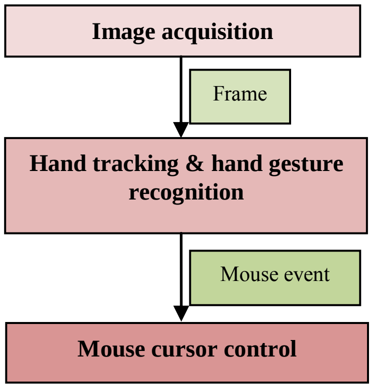

Fig. 3: System architecture

The image acquisition block consists of a web camera which captures images of the human hand. The hand tracking and hand gesture recognition block group the methods used to detect the human hand and interpret the hand gestures. The mouse cursor control block is responsible with the forwarding of the appropriate mouse event to the operating system.

The system overview is presented in fig. 4.

Fig. 4: The system overview

The web camera (A4Tech PK-635M), denoted by (1), captures the image of the human hand placed on a hand pad (3) and sends it on the computer (2). In computer, the image of the hand is processed and based on the gesture detected a command for the mouse cursor will be generated. The ring (4) is used for an easiest spatial detection of the hand on the plane of the hand pad. In fig. 5 is presented the picture of the hand and the ring attached to the hand during the system operation.

Fig. 5: Hand and the ring

4. The proposed procedure

To achieve the mouse cursor control the following general algorithm was applied:

Step 1. Acquire the current image of the hand pad area;

Step 2. Remove the noise from the acquired image;

Step 3. Detect the position of the hand and sets the mouse cursor (x,y) coordinates;

Step 4. Detect the hand pixels;

Step 5. Based on the detected hand gesture send the corresponding command to the mouse cursor;

Step 6. Jump to Step 1.

The noise removal from the Step 2 is done by applying a blurring filter with a 7x7 Gaussian mask.

The hand (x,y) position (detect on Step 3) on the hand pad plane is achieved by the detection of the ring (x,y) position. The ring has a blue color. To detect the pixels of the ring the authors converted the acquired image from RGB model in HSV model. After conversion the HSV image was threshold by the hue and saturation ranges of the ring area. After the threshold a morphological closing operation was applied. The morphological closing operation was achieve by applying first, a dilate operation and second, an erode operation [2]. The (x,y) coordinates of the ring are computed by

$$\begin{cases} x = \dfrac{1}{N_{ring}} \displaystyle\sum_{i=1}^{N_{ring}} x_i \\ y = \dfrac{1}{N_{ring}} \displaystyle\sum_{i=1}^{N_{ring}} y_i \end{cases} \qquad (1)$$

where N_{ring} is the total number of the pixels corresponding to the ring area.

The pixels corresponding to the hand area (Step 4) was achieved in the same way as the detections the pixels corresponding to the ring. The same HSV image was threshold by the hue and saturation ranges of the hand skin (Caucasian skin color) area. After the threshold, the previously mentioned morphological closing operation was applied.

The detection of the hand gesture was done by counting the pixels corresponding to the hand. The authors define three hand gestures (default, right click and left click) which are presented in fig. 6 – 8 and have different numbers of pixels.

Fig. 6: Default hand gesture

Fig. 7: Left click hand gesture

Fig. 8: Right click hand gesture

5. Experimental results

The source code of the software part of the system was written in C language using the OpenCV [10], library.

The system was tested for different levels of the daylight illuminance on the hand pad. For each considered level of illuminance the authors perform 20 hand gestures. In fig. 6 are presented the number of the detected hand gestures for each considered level of illuminance.

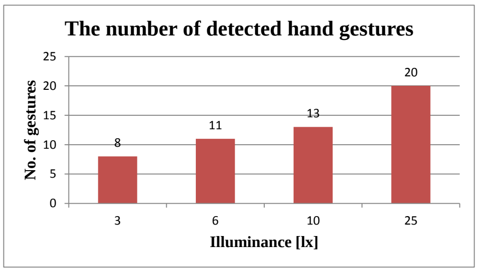

Fig. 6: The numbers of hand gestures detected for different levels of the daylight illuminance on the hand pad plane

6. Conclusions

As a conclusion to this paper, we proved that with few hand gestures and image processing algorithms, we achieved full control over the mouse cursor eliminating the mouse completely.

From the tests we've made, the conclusion is that the application is best to be used in the highest daylight or in a room with a high artificial light level.

With every research we came up, there are some benefits and disadvantages which pop out during the research process.

The benefits from our research are as follow:

- ability to adapt for various skin colors;
- ability to adapt after different hand gestures based on the number of pixels of each gesture made.

The disadvantage of this research is that this application can't be used every minute of every day because daylight varies from an illuminance value to another in few seconds based on the weather outside.

For example, if a cloud is passing by and blocks the sun light, the illuminance is changed drastically making the detection process more inaccurate and there for, less hand gestures recognized.

As we presented initially, this research can be moved on to a higher ideology and that is to fully gain control over a PC or gadget, not just by hand gestures, but with full body language (eyes, face expressions, mouth movements, arms gestures). All you need is a low-cost web camera and other image processing algorithms.

Motto: That's one small click for man, one giant gesture for mankind.

References

[1] Bardoşi F. I., Grif H. S. and Răţoi O. (2009), Mouse Controlled by Hand Gestures Using Digital Image Processing, *International Conference Interdisciplinarity in Engineering INTER-ENG 2009*, Tg. Mureş, Romania, pp. 258-263.

[2] Bradski G. and A. Kaehler A. (2008), *Learning OpenCV*, O'Reilly, US.

[3] Crane R. (1997), *A simplified approach to image processing: classical and modern techniques in C*, Prentice-Hall.

[4] De Campos T. E. (2006), *3D Visual Tracking of Articulated Objects and Hands. PhD. Thesis*, University of Oxford.

[5] El-Hafeez T. A. (2010), A New System for Extracting and Detecting Skin Color Regions from PDF Documents, *International Journal on Computer Science and Engineering*, Vol. 02, No. 09, pp. 2838-2846.

[6] Farella E., Benini L., Riccò B. and Acquaviva A. (2007), MOCA: A Low-Power, Low-Cost Motion Capture System Based on Integrated Accelerometers, Advances in Multimedia, vol. 2007, Article ID 82638, pp. 1-11. doi:10.1155/2007/82638. [Online]. Available: http://www.hindawi.com/journals/am/2007/082638/abs/

[7] Gonzalez R. C. and Woods R. E. (2001), *Digital Image Processing*, Prentice Hall.

[8] Hasan M. M. and K. Misra P. (2011), Brightness factor matching for gesture recognition system using scaled normalization, *International Journal of Computer Science & Information Technology*, Vol 3, No 2, pp. 35-46.

[9] Igorevich R. R., Park P., Choi J. and Min D. (2013), Two Hand Gesture Recognition Using Stereo Camera, *International Journal of Computer and Electrical Engineering*, Vol. 5, No. 1, pp. 69-72.

[10] OpenCV documentation (2013). [Online]. Available: http://docs.opencv.org/index.html

[11] Rautaray S. S., Agrawal A. (2012), Real time hand gesture recognition system for dynamic applications, *International Journal of UbiComp*, Vol.3, No.1, pp. 21-31.

[12] Sarifuddin M. and Missaoui R. (2005), A New Uniform Color Space with Associated Color Similarity Measure for Content-Based Image and Video Retrieval, *Multimedia Information Retrieval Workshop, 28th annual ACM SIGIR conference on Information Retrieval*, Salvador, Brasil, pp.1-8. [Online]. Available: http://mmis.doc.ic.ac.uk/mmir2005/CameraReadyMissaoui.pdf

[13] Shan C., Tan T. and Wei Y. (2007), Real-time hand tracking using a mean shift embedded particle filter, *Pattern Recognition*, 40, pp. 1958–1970.

[14] Tsagaris A. and Manitsaris S. (2013), Colour space comparison for skin detection in finger gesture recognition, *International Journal of Advances in Engineering & Technology*, Vol. 6, Issue 4, pp. 1431-1441.

[15] Wang X., Tarrío P., Bernardos A. M., Metola E. and Casar J. R. (2012), User-independent accelerometer-based gesture recognition for mobile devices, *Advances in Distributed Computing and Artificial Intelligence Jornual*, Vol.1, No. 3, pp. 11-25. [Online]. Available: http://oa.upm.es/16199/1/INVE_MEM_2012_132745.pdf