

Image recognition usage to control computer application

Abstract. The article deals with the possibility of control of selected applications running in Windows using images recorded in real time by a computer webcam. These images are simple tags, the more complicated the actual artwork hand and move your hand-gestures. It is possible to use two techniques. The first is placed on the hand of coloured markers. The second technique allows a more natural control - no markup. The article describes the basic problems affecting the image acquisitions. In the next part focuses on graphics libraries enable image recognition in real time. The fundamental errors of those libraries that hinder recognition were analysed and a number of solutions to correct these errors were proposed.

Streszczenie. W artykule zaprezentowano możliwość sterowania aplikacjami działającymi pod kontrolą systemu operacyjnego Windows, za pomocą obrazów zarejestrowanych przez kamerę internetowa. Obrazami tymi mogą być zarówno proste znaczniki, jak również znacznie bardziej skomplikowane obrazy ruchome czy też gesty dłoni. Zaprezentowano dwie techniki. Pierwsza z nich oparta jest o kolorowe znaczniki umieszczone na dłoni. Drugi sposób to technika bez markerów. Jest ona znacznie bardziej naturalna od pierwszej. W artykule wspomniano także o podstawowych problemach związanych z pozyskiwaniem obrazu w szczególności w kontekście późniejszego rozpoznawania. Przeanalizowano także elementarne błędy zawarte w bibliotekach utrudniające automatyczne rozpoznawanie obrazów. Zapropozowano kilka rozwiązań poprawiających te błędy. **(Wykorzystanie rozpoznawania obrazów w sterowaniu komputerowymi aplikacjami)**

Keywords: Image recognition, natural user interfaces, gesture recognition.

Słowa kluczowe: rozpoznawanie obrazów, naturalne interfejsy, rozpoznawanie gestów.

doi:10.12915/pe.2014.04.43

Introduction

The earlier versions of Microsoft Windows operating system allowed to communicate with the user mainly using the keyboard and computer mouse (Fig.1). Mouse control computer was implemented back in the 70's and the 80's improved. With the development of computers and specifically with the advent of portable computers: laptops, notebooks, netbooks it was necessary to control using the touchpad. The next evolutionary step was the creation of human-computer communication by NUI approach as well as research and improve the interfaces [2, 6].

Microsoft Windows 8 has been designed to enable and facilitate control by the touch screen.

This is a more novel form of application management than the obsolete computer mouse. Due to the fact that the vast majority of computers are equipped with a webcam there is a legitimate need testing and finding a new way of communicating one computer [1, 3, 5].

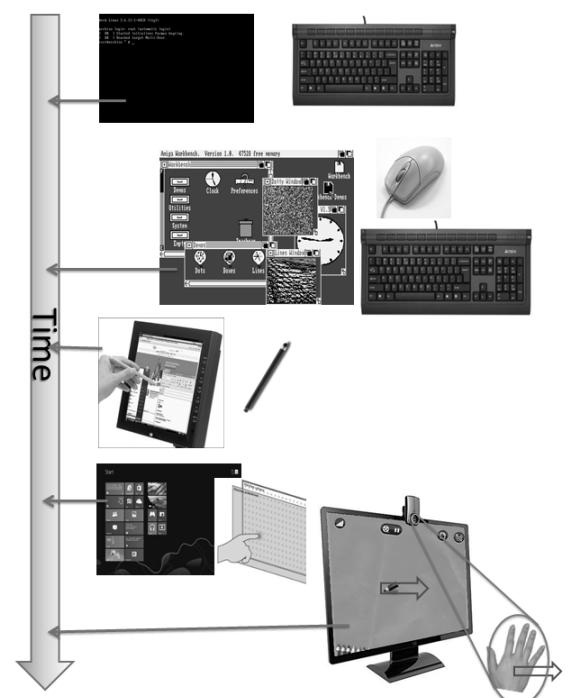


Fig.1. The evolution of interfaces.

Natural User Interface

The natural user interface (NUI) is a system for human-computer interaction that the user operates through intuitive actions related to natural, everyday human behaviour. The NUI may be operated in a number of different ways depending on the purpose and user requirements.

The NUI include: Gesture recognition, Speech recognition, Touch screen and Brain-machine interfaces

Gesture recognition systems track user motions or user hand motions and translate those movements to system instructions. They are used very frequently in entertainment especially in computer games. Nintendo Wii and PlayStation Move motion gaming systems work through controller-based accelerometers and gyroscopes to sense tilting, rotation and acceleration. A more intuitive type of NUI is outfitted with a camera and software in the device that recognizes specific gestures and translates them to actions. Microsoft's Kinect, for example, is a motion sensor for the Xbox 360 gaming console that allows users to interact through body motions, gestures. Kinect recognizes individual players' bodies. Gesture recognition can also be used to interact with computers.

Speech recognition allows users to interact with a system through spoken commands. The system identifies spoken words and phrases and converts them to a computer system actions, running embedded programs and applications. Speech recognition is common used also in hands-free computer and mobile phone operation.

Touch screen interfaces allow users interact with controls and applications more intuitively than a cursor-based interface because it is more direct then traditional mouse or touchpad operating.

Image recognition

Image recognition is a complex process. In the process of image acquisition, there are many problematic aspects. One is the optical distortion of the image, further reducing the CCD sensor noise. Methods of eliminating these distortions are described, inter alia, in several papers [4]. There is also a very serious problem of lighting [9]. Adverse lighting significantly hinders or even makes totally impossible to correctly recognize objects. Recognition of objects which reflect the light and the recognition of objects that cast a shadow is a process that requires additional calibration of the system.

In this article two concepts hand gesture recognition were presented. In the first algorithm on the hand markers were located. Each finger has a circular marker in a different colour (Fig.2). The first concept is based on markers using a two-step image recognition. In the first phase take place the detection of the characteristic colours of markers and then determining whether the colour of the marker belongs to the accidental element in the background. Later on the search using spherical filter elements were executed. The filter mask searches for objects with a given radius equal to the radius of the defined rate. The aim is to reduce the number of false detect "false marker" from the background image.



Fig.2. The hand with a different colour markers

This paper describes results obtained from an application using spherical markers of different colours. For the purpose of controlling the cursor and the mouse click and mouse double click marker placed on the fingers in green and blue colours.

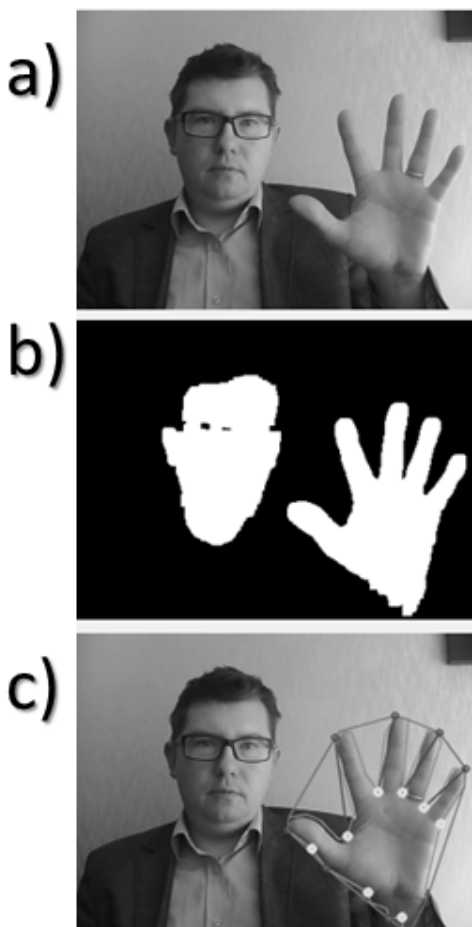


Fig.3. The skin detector: a) the real camera user view, b) computer vision –skin areas, c) founded and described hand area

The second idea is based on the detection of skin - the detection of an object of a particular colour from the defined range (Fig.3). From the image is separated portion which corresponds to a specific predefined colour - colour of human skin.

The methods limitation

Both presented methods are characterized by some limitations. In the case of detection of different colored markers placed on fingers, make sure that the background was not spherical elements like color saturation. The presence of such elements in the background can disrupt the successful detection markers. Due to the fact that the method searching for spherical elements defined color in the background (in the garment) that controls the user's system may be located colored elements even with similar color intensity.

In the second method, gesture recognition, it is based on detection of human skin. Just as in the first method, the range of the image to be recognized as the skin is defined by the appropriate range of colors in the calibration process. Very often, as shown in Figure 2 the method qualify as a skin area not only a hands but very frequently qualify the user's face. It is a natural phenomenon, as the face and hands are the same colour. Due to the fact that the webcam does not have a measure of depth that would allow immediate detection of objects in 3D space, it is not possible to determine which object is a hand (it is closer) and which is the face (which is usually referred to as away from the camera). This method uses, however, some simplification is that the hand cursor control surface is greater than the face. This is the right perspective: the hand is located closer to the camera than face of the user.

Cursor control technique

Two cursor controlling applications were created bases on two different methods. The applications uses webcams in Windows operation system. Created applications take advantage of a free library Emgu CV and were written in C #.

Finding a hand in the image is the basis of gesture recognition process and move the cursor. The cursor movement corresponds to a hand movement acquired by a camera image. Image captured by the webcam is a mirror image of the real. This allows man's hand move to the right in a natural way to move the cursor to the right. Control thus becomes completely intuitive. Because of different resolution image between a webcam (X_k, Y_k) and desktop MS Windows resolution (X_w, Y_w) it was necessary to rescaling what was shown in Formula 1:

$$(1) \quad X_w = AX_k, \quad Y_w = BY_k$$

where: X_w, Y_w - Windows positions on the screen, X_k, Y_k - position in web camera screen, A, B- coefficients

The coefficients A and B were chosen empirically. In case of recognizing the hand without markers take into consideration the calculated coordinates of the center of gravity of the figure obtained by surrounding palm convex hull. These coordinates are substituted into the formula 1 as (X_k, Y_k).

In the second case - to recognize the hand with the markers it is necessary to approximate coordinates of the center of gravity of the first marker.

Then, exactly the same in both cases, the position of the cursor is calculated. By converting (formula 1) it is proportional representation of the position of the cursor on the screen to the position your hand in the camera image.

Gesture recognition

Another issue is to recognize the gesture [6-8, 10]. One of the fundamental Windows functions is support to click on the object and support the double click. The double click, from the perspective of recognizing gestures is a double detection of single-click in a short period of time. On this functionality of the applications presented was focused. In the first application the gesture corresponding to a mouse click is hiding second marker in the still presence of the first marker. Showing and hiding again the second marker has the effect of double click. It was shown on Fig. 4.

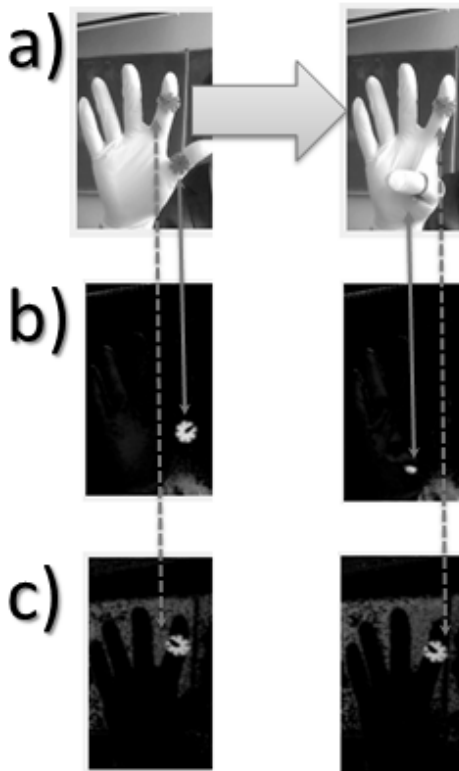


Fig.4. The first application, the characteristic gesture- mouse click: a) gesture-mouse click, b) computer vision –searching of first marker, c) computer vision –searching of second marker

In the second application: a system to detect hand with skin detector - clicking is a sudden change in the geometric center of gravity of the hand image. Established that the hand responsible for the single click consisted in the firmament of his fingers (Fig. 5). Such a move causes a rapid change of coordinates of the center of gravity of the figure.

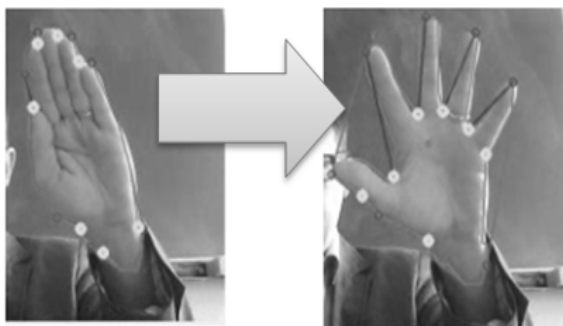


Fig.5. The second application, the characteristic gesture- mouse click

Calibration of the system

Before using the application, it is possible and recommended to make the calibration of the system. It applies the same camera settings such as contrast, brightness and saturation. This is the same web camera calibration in both applications. Mentioned calibration mechanisms are defined separately for each application. In the first of them, the system calibration is to determine what specific colour marker is and how tolerance is to be searched. The HSV (hue, saturation, value) colour space in addition allows you to modify the brightness - equivalent to the power of the present scene illumination. In addition, the filter parameters are modified tolerance radius search and classified as a marker circles.

In the second application, after the camera calibration process, takes place a skin detector calibration. The user is able to redefine the default colour range for the classification of the test object as the hand.

In the calibration process, in addition to technical camera calibration, it is also possible to set all other ranges.

First of all, is possible to set values of the coefficients A and B to allow full movement around the screen.

In the first user has the ability to set color ranges for all markers. It is possible to also view the other items in the background are classified as markers. In this way can change the color markers on the one that does not interfere with the background environment.

For the second applications, it is possible to set the color ranges for which the system recognizes as a skin area.

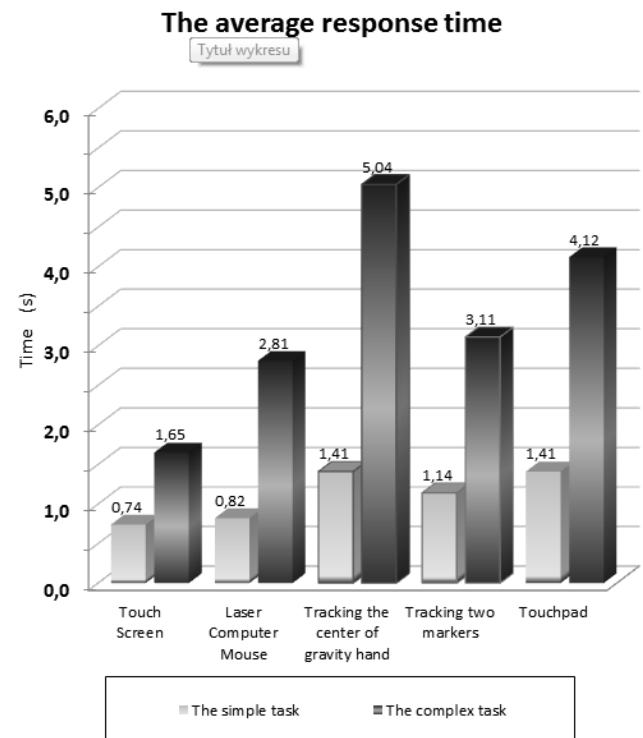


Fig.6. The average response time

The tests of applications

A computer with a conventional web camera processes the image in real time. On detection the hand algorithm operates the palm extract from an image, select it and then contour it and the center of gravity is measured as the resulting figure. It was observed that each type of gesture causes to change the center of gravity of the figure. By assigning the appropriate gestures vectors change the

center of gravity figures compiled program correctly responds to movement and gesture.

In order to test the validity and ergonomic working with these two applications are defined and then the test was performed. It consisted of two tasks named simple and complex. The simple task simulate clicking the two buttons spaced apart by about 80% of the screen.

The complex task was to choose the object SELECT particular, the penultimate value. SELECT field values and had 16 was located at the upper edge and the screen mats. The precise choice of simulated real use of more complex applications. Taken on 10 measurements for each of the steps and then compared them with the results of the same test activities performed by the same user using the keypad familiar Windows. Averaged results were shown in figure 6.

Summary

The paper presents a new way of human-computer communication. Gesture recognition is one of the Natural User Interfaces. Applications do not require any additional costly interfaces, it uses a built-in webcam. Both presented applications to work properly. The proposed gestures seem to be natural but there is a possibility to change the other, if proved to be too tiresome.

In order to confirm the usefulness gesture control, tested speed of execution common tasks in Windows. Proposed two simple test and then measured execution time with new and well known interfaces: touch screen, gesture recognition - tracking the center of gravity hand (skin detector), gesture recognition-tracking markers, computer mouse and touchpad.

The application was tested on a group of five professionals working as programmers. Not surprisingly, they prefer to work with a mouse (LED 1000 dpi) and its use in the most efficient way. The touchpad is also very usefully interface. However the size of the described test required the user shift after the finger twice. This had a significant impact on the execution time. One of the most natural forms of communication-touch screen did a great. But there are many situations which prevent use such technology. The main part of the study was to investigate controlled gesture. Control of "non-contact" with both hands and markers without it took the most time. This is mainly the result of the

need to scale cursor movement. It causes the characteristic tremor occurring cursor with minimal hand movement.

REFERENCES

- [1] Athitsos V., Wang H., Stefan A., A database-based framework for gesture recognition, *Personal and Ubiquitous Computing*, 14 (2010), No 6, 511-526
- [2] Borys M., Miłosz M., Plechawska-Wójcik M., Luján-Mora S., Synergy effect in GUI usability and accessibility education improvement., *Proceedings of 2013 IEEE Global Engineering Education Conference (EDUCON)*, (2013), 311-317
- [3] Chern-Sheng Lin; Hsiang-Wen Tseng; Yun-Long Lay; Chia-Tse Chen; Mau-Shiun Yeh, Design and reliability analysis of image type head gesture interface in distributed control system., *Indian Journal of Science & Technology*, 5 (2012), No 8, 3150-3156
- [4] Koziel G., Fourier transform based methods in sound steganography; *Actual Problems of Economics*, 120 (2011), No 6, 367-373
- [5] Lui Y.M., Human gesture recognition on product manifolds, *Journal of Machine Learning Research*, 13 (2012), 3297-3321
- [6] Malizia A., Bellucci A., The Artificiality of Natural User Interfaces, *Communications of the ACM*, 55 (2012), No 3, 36-38
- [7] Reilly R.B., Applications of face and gesture recognition for human-computer interaction, *Proceedings of the sixth ACM international conference on Multimedia*, (1998), 20-27
- [8] Song Y., Demirdjian D., Davis R., Continuous body and hand gesture recognition for natural human-computer interaction, *Transactions on Interactive Intelligent Systems (TiiS)*, 2 (2012), No 1, 1-28
- [9] Szymczyk T., Image recognition based on texture detection as the element of modern industrial automatic control systems; *Polish Journal of Environmental Studies*, 17 (2008), No 3B, 456-461
- [10] Wachs J.P., Kölsch M., Stern H., Edan Y., Vision-Based Hand-Gesture Applications, *Communications of the ACM*, 54 (2011), No 2, 60-71

Author:

dr inż. Tomasz Szymczyk, Lublin University of Technology, Electrical Engineering and Computer Science Faculty, Institute of Computer Science, Nadbystrzycka 36B, 20-618 Lublin, e-mail: t.szymczyk@pollub.pl