Gesture-based Computer Control System applied to the Interactive Whiteboard

Michal Lech, Bozena Kostek
Multimedia Systems Department
Gdansk University of Technology
Gdansk, Poland

Abstract — **In the paper the gesture-based computer control system coupled with the dedicated touchless interactive whiteboard is presented. The system engineered enables a user to control any top-most computer application by using one or both hands gestures. First, a review of gesture recognition applications with a focus on methods and algorithms applied is given. Hardware and software solution of the system consisting of a PC, camera, multimedia projector and screen for projected images is presented. Image processing methods used and the fuzzy rule-based gesture recognition concept are described. The application called Interactive Whiteboard, enabling to draw shapes on a virtual whiteboard developed for use with the system is then presented. The functionality of the Interactive Whiteboard and the way of managing the contents with hand gestures is shown. Results of testing the system are also contained.**

Keywords-gesture recognition; hands detection; fuzzy logic; human-computer interaction

I. Introduction

Among human-computer interfaces that recently have gained much interest, gesture recognition systems can be mentioned. Various solutions provide not only recognition of forearm and palm gestures but also recognition of lip [1], head or eye movements and body poses (e.g. Natal project). However, main attention is still devoted to hand gesture recognition. The best recognition accuracy can be achieved using motion sensors (e.g. Datagloves) but from the point of view of user's convenience the most promising solutions are vision-based. Majority of such solutions use a camera placed in front of a user.

In the paper a vision-based gesture recognition system which employs a camera and a projector, both placed behind the user, is first presented. Then the concept of using fuzzy logic for reliable gesture recognition is proposed (Section 3.2). Section 4 presents the Interactive Whiteboard application engineered to demonstrate possibilities of the developed interface. Conclusions are also contained.

II. Gesture Recognition Systems

Gesture recognition solutions can be divided with regard to the type of gesture used for controlling a computer. Gesture can be considered as a change of the object position (e.g. hand movement) in a particular time interval with a given velocity or as a change of the object shape (e.g. forming ellipse with the thumb and the index finger). Gestures that belong to the first group are typically called dynamic gestures while these from the second group are often referred to as static gestures [2]. Because the system presented bases on dynamic gestures, below some chosen methods used for recognition of these particular activities are given. Especially fuzzy logic-based recognition systems are of the main interest of the authors, since the fuzzy rule-based approach to gesture recognition is proposed in the paper.

A typical approach to dynamic gesture modelling uses probabilistic models based on conditional probabilities, most often with the discrete state value. As an example Hidden Markov Models (HMM) can be mentioned. For each dynamic gesture a separate HMM is created. Gesture recognition is based on estimating the probability of representing the observed sequence by a particular model [3].

On the other hand, Kalman filters are often employed for tracking object position changes in noisy conditions [2]. Filters enable to determine the most probable state (e.g. position, velocity) basing on observations of the tracked object. The assumption is made that the disturbance influencing motion of an object and measurement errors are of the unimodal Gaussian distribution. For multimodal distributions, the state of an object can be determined using particle filtering from a group of Monte-Carlo sequence methods.

Moreover, HMM and fuzzy logic used to recognize Korean sign language may be an example of another approach [3]. The linguistic variables are the speed and the change of speed. Created fuzzy sets are used to define five motion phases, i.e. stop, preparation, stroke, moving, end. Basing on these phases Finite State Automata (FSA) is created for the sentence segmentation. The discrete HMM is further used for classifying hand motions of the sign language. Another solution that uses fuzzy logic along with the Interval Mathematics to represent erroneous values obtained from the data glove sensors was proposed [4]. Each sensor corresponds to a linguistic variable (i.e. denoting sensor of each finger), which values are linguistic

Research funded within the project No. POIG.01.03.01-22-017/08, entitled "Elaboration of a series of multimodal interfaces and their implementation to educational, medical, security and industrial applications". The project is subsidized by the European regional development fund and by the Polish State budget.

terms representing typical angles of joints (i.e. STRAIGHT, CURVED, BENT) and separations (i.e. CLOSED, SEMI-OPEN, OPEN).

III. Gesture-Based Computer Control System

The gesture-based computer control system is the patent-pending interface based on subtracting the image displayed by a multimedia projector from the image acquired from a video stream, and recognizing gestures in the further processed output. A user stands between the multimedia projector coupled with a camera and a projection screen (Fig. 1). Controlling is based on the recognition of one or both hands dynamic gestures. There are basic dynamic gestures predefined in the system, i.e. moving one hand or both hands left/right and up/down, moving both hands further apart/close together, simultaneous movement of the left hand up and the right hand down and vice versa. Keeping one hand or both hands steady is also considered as two independent basic gestures. Each basic gesture is associated with a typical system action. For instance, moving left hand up and right hand down is interpreted as rotating the image right when browsing images. Additionally, mouse events handling is implemented. Hand motion is associated with the mouse cursor movement. A user can choose which hand has a priority over mouse cursor if two hands appear in the camera field of vision. All basic gestures are collected in a circle buffer and constitute complex gestures analyzed as FSA. A user can assign system actions to these gestures. For example, drawing an *X* in the air causes the application which is being gesture-controlled to be closed.

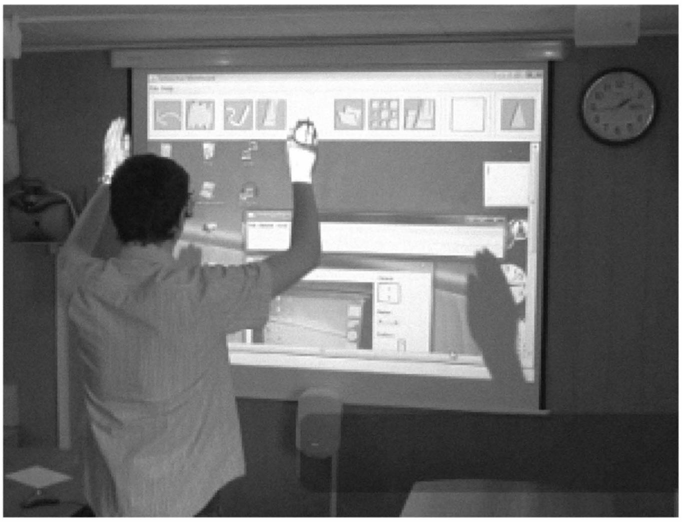

Fig. 1. The way of using Interactive Whiteboard (zooming the contents)

A. Preliminary image processing

First, the calibration process is performed. Five color images (red, green, blue, white, black) are displayed. Captured frames are subtracted from respective displayed images. Results of subtraction are stored in five arrays from which they are retrieved and used during further image processing. This enables to reduce distortions introduced by camera lens, especially the vignetting effect. The task of the calibration phase is also to determine the area of the displayed image in the camera frame. This process is based on indications of a user, who points corners of the image in the camera frame displayed in the graphical user interface (GUI). The cropped image is perspective corrected and the displayed image is scaled to ensure identical dimensions of both images. Each gesture recognition iteration begins with subtracting the frame processed from the scaled projected image. Subtraction is performed in RGB color space. To each output pixel an appropriate value from the arrays created during calibration is added. This value is chosen based on pixel color and intensity. The resulting image is binary thresholded and median filtered with a separable square mask.

B. Fuzzy rule-based recognition

Image processing described above enables to obtain image containing user's hands only. Therefore, the hand detection algorithm is based on a simple feature vector containing the shape area threshold and the pixel intensity threshold. Detected hand positions are analyzed using Cartesian coordinates. Changes of detected hand positions can be interpreted as gestures using fuzzy rule-based processing.

There are three linguistic variables proposed, i.e. velocity υ_t, change of velocity $\Delta\upsilon_t$ and direction. Input values for the fuzzy system associated with the first two linguistic variables are calculated using Eqs. (1) and (2), respectively. The value of direction is denoted as an angle between the positive *y* axis and a vector constructed for two consecutive detected hand positions. Linguistic terms for υ_t are {*very small*, *small*, *medium* and *large*}, for $\Delta\upsilon_t$ {*negative medium*, *negative small*, *no change*, *positive small* and *positive medium*}. For directions the terms are described as {*north*, *east*, *south*, *west*}.

$$\upsilon_t = \frac{\sqrt{(x_t - x_{t-1})^2 + (y_t - y_{t-1})^2}}{\Delta t} \tag{1}$$

where: x_t – x coordinate of hand position for time *t*,
x_{t-1} – x coordinate of hand position for time *t-1*,
y_t – y coordinate of hand position for time *t*,
y_{t-1} – y coordinate of hand position for time *t-1*,
Δt – change of time

$$\Delta\upsilon_t = \upsilon_t - \upsilon_{t-1} \tag{2}$$

where: υ_t – velocity in time *t*,
υ_{t-1} – velocity in time *t*-1.

Fuzzy sets for linguistic variables υ_t and direction are given in Fig. 2.

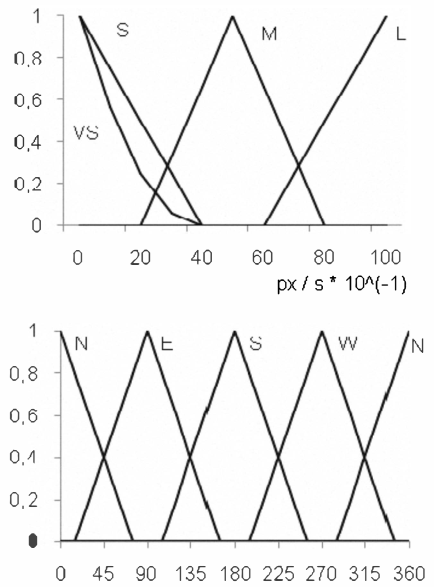

Fig. 2. Fuzzy sets for linguistic variables: velocity and direction

The membership function $\mu_{CON(S)}(v)$ for linguistic term *very small* is created using concentration operation of fuzzy set *small* (*S*) according to Eq. (3):

$$\forall v \in V: \quad \mu_{CON(S)}(v) = (\mu_S(v))^2 \tag{3}$$

While analyzing natural human hand motions one can observe that moving hand in a particular direction is preceded by a short preparation phase (anticipation) during which the movement in the opposite direction occurs. Taking this feature into account fuzzy rules are created with regard to directions for time t_i and time t_{i-1} and by analyzing change of speed. Thus, an example of a rule for a gesture from class g_1 denoting fast movement in the left direction with the anticipation phase can be described as:

if v_{t-1} is S and d_{t-1} is not $N \cup S$ and v_t is L and d_t is W then g is g1 (4)

According to forms of fuzzy rules used, fuzzy operators are of a type min-max. The inference zero-order Sugeno model is used with singletons denoting gesture classes. The output of the system is the maximum of all rule outputs.

IV. Interactive Whiteboard

The Interactive Whiteboard application has been developed to present functions provided by the engineered Gesture-based Computer Control System. The system besides the common functionality of writing and drawing shapes on the board provides features not possible to obtain with the classical school black- or whiteboards. Among functions directly associated with previously mentioned gestures recognized by the engineered gesture-based interface are rotating, zooming and scrolling the contents. Writing and drawing are achieved based on the function of the mouse cursor controlling and handling left button clicking event. All operations are performed from the distance, thus the whiteboard contents is not obscured by a user. To begin writing or drawing, the user places his hand over the whiteboard and keeps it stable. Based on the mechanism of the complex gesture recognition, mentioned in the previous section, actual pixel printing begins when a gesture labeled ‘long click’ is detected. To stop writing/drawing the same gesture is used. A time interval required to recognize gesture as a ‘long click’ can be adjusted according to the user’s preferences. The value can be chosen from the range of 340 – 2040 ms. Among other functions provided by the interface of the Interactive Whiteboard are undo operation, clearing the whole area of the board, choosing a chalk color and its size, choosing a sponge and opening files. The way a file is opened depends on its format. Graphical files are loaded directly into the whiteboard contents. Therefore, drawing on the picture is possible. Files of a format other than graphical are opened using the default application system assigned to the particular extension.

V. System Tests

There were two types of tests performed during the experiments on Gesture-based Computer Control System recognition effectiveness. The first type consisted in examining recognition effectiveness of gestures made on static fully-colored backgrounds. Tests were performed for 4 various colors (red, green, blue, white) of backgrounds displayed by the multimedia projector. For each background color, each gesture recognizable by the system was repeated 10 times. The second type of tests consisted in browsing series of 15 exemplary images provided in the operating system. Each test lasted 10 minutes. There were no constraints put on the user’s actions. The user decided which gesture from the set of gestures is chosen for image browsing.

Overall recognition effectiveness during tests with solid color backgrounds was above 95% and during tests with images viewing was above 81%. Improper recognition was mainly due to existence of a 1-second pause in recognition immediately after the particular gesture is detected. Such a pause in recognition was introduced as a temporary solution for lack of the perfect time synchronization between frames captured from the camera and the printed system screens. Improving synchronization which would produce no pause or remembering about this 1-second pause while using the system result in a higher reliability.

VI. Conclusions

In the paper, the Gesture-based Computer Control System with dedicated Interactive Whiteboard application has been presented. In comparison with other interactive whiteboards no motion sensors or infrared diodes and cameras were employed. The system uses versatile components, i.e. a PC, webcam, multimedia projector and optional projection screen. The system can be especially useful while giving lectures or presentations. Considering functions provided for image browsing, one can use the interface with image viewers, e.g. for contactless handling in photo laboratories. The system in its current form enables to write and handle the contents from the distance. Such a solution is useful due to the fact that the whole area of the screen is well visible to viewers sitting in front of a user. However, writing or drawing standing in a close proximity to the whiteboard may be a more convenient functionality. It would provide more similar way of contents handling in comparison to non-interactive white- or blackboards. Thus, better accuracy while drawing would also be achieved. Therefore, the work on the system will continue on applying methods enabling to extract area of hands separately from the rest of the user's body.

REFERENCES

[1] Dalka P., Czyzewski A., *Human-Computer Interface Based on Visual Lip Movement and Gesture Recognition,* International Journal of Computer Science and Applications, Technomathematics Research Foundation.

[2] Mitra S., Acharya T., *Gesture Recognition: A Survey*, Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 2007, vol. 37, no. 3, pp. 311-324.

[3] Jung-Bae Kim, Kwang-Hyun Park, Won-Chul Bang & Zenn Bien Z., *Continuous Gesture Recognition System for Korean Sign Language based on Fuzzy Logic and Hidden Markov Model*, IEEE, 0-7803-7280-8/02, 2002, pp. 1574-1579.

[4] Callejas Bedregal BR., Dimuro GP., Rocha Costa AC., *Interval Fuzzy Rule-Based Hand Gesture Recognition*, Scientific Computing, Computer Arithmetic and Validated Numerics, 2006. SCAN 2006. 12th GAMM - IMACS International Symposium on, pp. 12-12.