International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 1, January 2016)

Gesture Recognition System for User Interface Control

Milena Lazarova
Assoc. Prof., Dept. Computer Systems, Technical University of Sofia, Sofia, Bulgaria

***Abstract*— The paper presents a gesture recognition system for human-computer communication. The system is based on static gesture detection using a hybrid approach for skin detection, seed markers watershed segmentation and morphological operations, palm centre allocation by a combination of minimal covering circle centre and palm convexity defect points and classification of gestures using decision tree based on geometric skeleton and the palm circle. The gesture recognised is associated to a particular user interface command. The paper presents experimental evaluation of the suggested approach in terms of recognition accuracy under various lighting conditions and usage of different skin detection approaches.**

***Keywords*— gesture recognition, human-computer communication, palm detection, skin detection, user interface, watershed segmentation.**

I. INTRODUCTION

Recognition and interpretation of an image content is deceptively easy for the human vision but is a complex problem when a computerized solution should be suggested and implemented. There are numerous practical applications that take advantage of identifying objects, people, buildings, places, logos in an image such as security, quality control, medicine, education, entertainment.

Electronic devices today are an integral part of everyday life. Different software applications are used for a variety of tasks in support of our life, work, study, entertainment. Whether the applications are running on a computer, a mobile device or a bank terminal, the user interface is of great importance for the effective communication between the human and the machine as well as the satisfaction of people using the software applications.

One of the main challenges in building a user interface is related to the achievement of suitability for the target user group and compliance with the computing power of the hardware platform. In the demand to provide user satisfaction, the classical methods for interaction with the devices such as the use of a mouse or a keyboard are gradually complemented and in some cases are shifted to the usage of innovative technologies such as touch screen, voice and gesture recognition.

Recognition of gestures have many applications in the field of remote control of devices and equipment as well as in entertainment and computer games.

The paper presents a system for user interface based on gesture defined commands. The gesture recognition algorithm suggested utilizes segmentation of an image, identifying the palm region, detecting and evaluating features based on the number and position of the fingers and recognition of a set of gestures that start specific user command of the operating system.

II. GESTURE RECOGNITION

The main task of a system for gesture based user interface concerns detection and identification of a hand gesture in an input image or a video stream. Once recognized the gesture can be associated to a particular command of the user interface.

The main advantages of gesture based user interface system are:

- the human-computer interaction does not require physical contact with the periphery devices making the system more compact and eliminating some points of failure;
- the user control is based on human body gestures, the most common being hand gestures, that people are accustomed to in their daily lives [2];
- the user interface may not require any specific accessories or might require some that people are accustomed to in their daily lives (as gloves, finger marks, etc.);
- the system control stimulates physical activity of the user, including gestures that might require active physical exercises for a variety of training and educational applications [3].

Some disadvantages of gesture based system control are:

- illumination dependence – the light sources in the scene have critical impact on the sensor's colour acquisition and therefore seriously influences the accuracy of the input image recognition;
- sensor dependence – the sensor's noise may confuse the proper interpretation of the input image;
- gesture dependence:

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 1, January 2016)

- the definition of gesture categories that the system can recognize have to provide adequate recognition accuracy while eliminating the uncertainty and is usually a main problem in any pattern recognition system [3]. Despite various gesture representation approaches exist high recognition accuracy is still one of the major challenges especially for a large number of gestures;
- gesture start and end should be carefully detected especially in real-time gesture based control systems which pose a critical requirement for minimal time between gesture perception and system command issue. It questions the moment you have to start recognizing gestures and timing in which recognition is complete. Smooth blending between gestures is natural for people thus causing difficulties in delineating and separating the gestures and [4].

- user dependence – some studies suggest that increased involvement and physical activity of the user in interaction with the system is sometimes inappropriate and leads to an excessive physical stress for longer periods of interaction [5].

The basic components of a gesture recognition system is presented on fig.1.

The pre-processor accepts input image by a sensor such as a camera and generates a digital presentation of the image in the system. In modern digital cameras usually this step is done in the camera itself, but it is possible the format used by the camera is not appropriate for the purposes of the system and a conversion might be required.

The sensors used for image acquisition may operate in different light spectrum channels. The infrared spectrum is preferred in the automotive industry for control of various devices as well as in some game consoles and devices for example MS Kinect sensor [17, 18].

In the case of gesture recognition system for user interface it is not relevant to count on and require a specialized and expensive camera sensors. Most often the users employ camera that operate in the visible portion of the spectrum operation with electromagnetic waves in the range from 390 to 700nm. Such devices are commercial video cameras, cameras of smartphones, webcams and more. Usually, these cameras produce images with triple channel digital RGB colour organization.

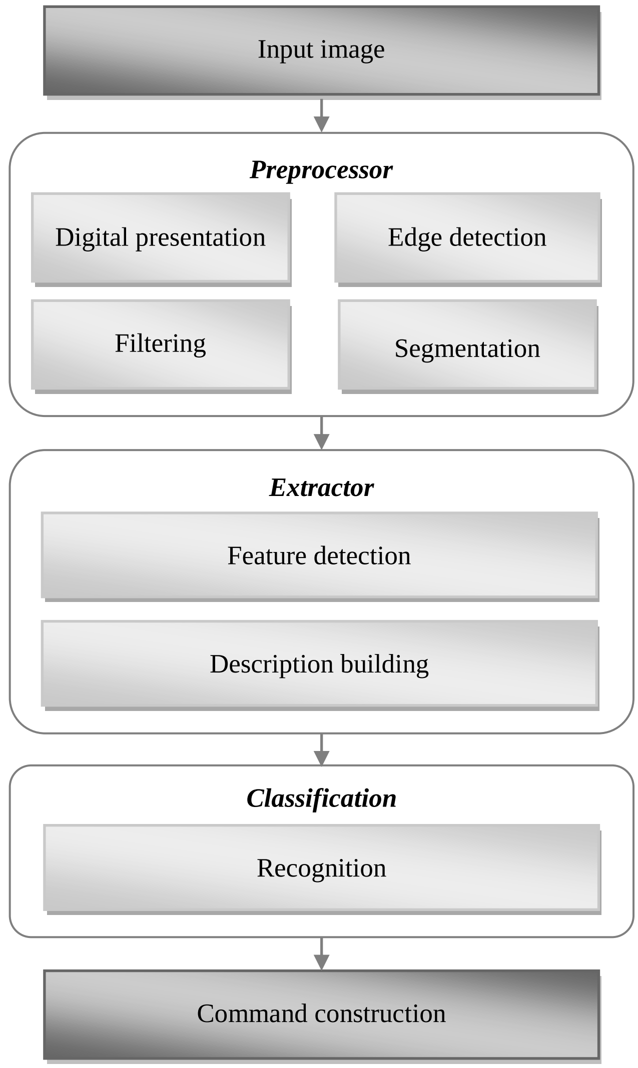


Fig.1. A basic components and stages of gesture recognition system

Since RGB colour space is very close to the biological model of human colour perception it is most commonly used in the image sensing devices. However colour based image analysis in the RGB space is inappropriate especially for skin colour segmentation mainly due to strong dependence on the light sources [6, 19]. To overcome the light conditions dependence the input RGB image is converted to presentation in another colour space.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 1, January 2016)

In the skin detection case the preferred representation uses HSV space with three components being hue, saturation and intensity or YCrCb colour space, where the brightness data are represented in Y channel and the colour tone is stored as Cr and Cb channels.

The influence of ambient light on the registered colour values decreased by ignoring the intensity component V [9, 10] or the luminance component Y [6, 7, 8].

In addition to the conversion of the input image to HSV or YCrCb colour space an image filtering is usually required at the pre-processing stage in order to remove noise and defects caused by either the technical specifications of the sensor or the acquisition conditions. The most common noise removal approach uses either linear convolution based filtering such as Gaussian filter or nonlinear rank-ordered filtering such as median or mode filter.

The next pre-processing step is aimed at object edge and boundary detection. The detected contour pixels are used for image segmentation as well as for extraction of some specific features such as the fingertips.

The segmentation is aimed at determining feature points and regions that are essential for the next stages of analyses and recognition of the image. The goal of the segmentation step is to locate specific feature points such pixels with particular colour properties and to outline region containing certain objects such as hand palm and fingers.

The main objective of the extractor stage of gesture recognition system is to detect and describe features in the input image that will support the particular image analyses goal in this case gesture detection and recognition.

While the pre-processing stage mainly consists of operations that are position invariant and context independent, the feature extraction and description stage requires intensive use of specific knowledge and constraints relevant to the given problem. In the case of gesture recognition, the features extracted are some hand palm features (range, centre) and finger features (peaks, length, angle between fingers). The choice if the set of particular features depends on the classification algorithm employed for gestures recognition. The goal of the extractor stage is detection of the selected features and building of relevant description as required by the classification stage.

Classification is the process of mapping an object in the image to a predefined object category [3]. The classifier uses the extracted descriptions in order to determine the objects category. An important aspect of the classification is the definition of object classes.

In general an object is described by a feature vector and the classification is based on evaluating the correspondence between the object's feature vector and the feature vectors associated with the predefined categories. Various classification approaches may be used for feature vector comparison such as minimum distance to the average class vector, minimum distance to random feature vector represented in the system, structural relationships between the elements of the vector, a decision tree, a probabilistic classification, a neural network model, etc.

III. Gesture Recognition System for User Interface Control

The paper presents an algorithm to detect static gestures that are integrated into the user interface control system. The system is designed as a C ++ application for Windows OS and uses OpenCV library [20].

The gesture recognition algorithm has the following characteristics:

- recognition of a set of static hand gestures that are used to control a system user interface;
- recognition based on human skin detection but a specific mode is provided that allows usage of certain accessories such as hand gloves;
- easy and convenient description of the hand gestures, straightforward adding of new hand gestures;
- independent on the particular user interface control system, allows application in any system through attachment of a hand gesture to particular system command;
- fast processing that allows further application in embedded systems.

When applied to the graphical user interface of the Windows operating system the suggested gesture recognition approach allows the user to control the graphical user interface remotely by using the defined hand gestures, working in real time, without specific requirements for sophisticated equipment and technologies

A. *Detection of potential pixels from gesture regions*

In order to achieve independence of the user interface system and to allow it to operate in a variety of lighting conditions and using variety of input devices a Gaussian filter of size 7×7 is applied at the pre-processing stage. As an additional measure to reduce the impact of ambient light filtered image is converted to YCrCb colour space.

Regions are sets of pixels that are joined together when common similarity criterion is satisfied. In the gesture regions case the similarity criterion used is based on evaluation of the maximum likelihood for the pixels to belong to a region that describes the hand gesture. The system supports both modes of recognition using regions of either hand skin or a glove of specific colour. The detection approach in both cases is the same and is based on an inverse projection of a predefined histogram of the gesture regions.

In order to predefine the gesture region histogram and to allow system independence of the skin or glove colour a systems requires training stage to be completed: the hand palm have to be positioned in front of the camera the camera so that a region of size at least 24×24 pixels containing only hand palm pixels to be detected and retrieved. A two-dimensional histogram for Cr and Cb channels of that region is calculated and is used at the hand region detection step. The histogram resolution is 100×100 pixels thus converting the [0, 255] interval of the Cr and Cb values to the range [0, 99].

An approach for object detection in an image based on histogram backprojection is used gesture region detection [12]. The method applied compares the colour histogram of the object to be detected to the histogram of the predefined hand region. First the two colour histograms are calculated: the histogram of predefined hand region is denoted by M, and the histogram of a region in the input image is denoted by I. A ratio of the two object histograms R is considered as probability that the pixels in the input image belong to a gesture region:

$$R = M / I \qquad (1)$$

To find the backprojection of R a matrix B is calculated in which each element with coordinates (x,y) denotes the probability the corresponding pixel to belong to the targeted object region:

$$B(x,y) = R[h_1(x,y), h_2(x,y), ..., h_n(x,y)] \qquad (2)$$

Where R is an n-dimensional histogram for n-band image – $h_i(x,y)$ is the histogram division to which the pixel belongs in the i^{th} dimension of histograms. In the system described in the paper only Cr and Cb channels are used from the representation of the input image in YCrCb colour space and in this case (2) has the following form:

$$B(x,y) = R[h_{Cr}(x,y), h_{Cb}(x,y)] \qquad (3)$$

In addition the following condition is applied:

$$B(x,y) = \min[B(x,y), 1] \qquad (4)$$

A convolution with a disk shape structural element D is applied on B:

$$B = D * B \qquad (5)$$

If a single pixel have to be detected then the element with a maximum value in B is selected. When a region have to be detected then a threshold value is applied to determine which pixels are joined in the region.

In the described system for gesture recognition the above approach is used to detect pixels with similarity specified by the histogram of the predefined input hand region.

An alternative and very simple approach for detection of regions of human skin applies a threshold for Cr and Cb colour channel values. The threshold values used are as described in [7]:

$$80 \le Cb \le 120 \qquad (6)$$
$$133 \le Cr \le 173$$

All pixels with values for Cr and Cb channels that satisfy (6) are considered as belonging to gesture region.

Another alternative approach is the histogram backpropagation to be applied for the searched object in which case the object comprises the pixels that are known to belong to human skin region.

A serious drawback of both methods for skin region detection is a critical dependence on the lighting conditions at the time of histogram calculation. To minimize this drawback the following hybrid approach is suggested:

- two channel Cr and Cb histogram hinput of the predefined skin region is calculated;
- the pixels with skin colour are detected for each frame of the video stream based on the static skin model (6);
- two channel Cr and Cb histogram hstatic for the detected pixels based on the static skin model is calculated;
- a new histogram hsum is calculated as a weighted sum of the previous two:

$$h_{sum} = 0.3h_{input} + 0.7h_{static} \qquad (7)$$

- a backprojection of the histogram h_{sum} to the input image is evaluated.

The suggested hybrid approach is based on the solution described in [21] for the HSV colour space. Some additional statistical calculation of the results for the aggregate histogram is also applied.

B. Segmentation of regions

A watershed algorithm is used for image segmentation into regions [13].

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 1, January 2016)

The input image is considered as a topographic surface with a height determined by the image function, i.e. the intensity of the corresponding pixel. Watershed segmentation is based on the assumption that each object in the image has a homogeneous texture and low gradient values. Therefore the objects correspond to minima of the morphological gradient and the object contours are the watershed of the gradient.

If all detected gradient minima are used in the watershed segmentation this may lead to an over segmentation [13] – too many regions but insignificant for the given problem. One solution to avoid this disadvantage is to extend the algorithm by introduction of primary markers for segmentation initialisation [14]. An important issue of this approach is the method for selecting the markers. The system described in the paper adopts a marker selection approach based on iterative conditional erosion and dilation:

- A morphological erosion with a square structuring element S of size 3×3 pixels is applied to the binary image B with the probable gesture pixels. The result B_{foregr} is assumed to consist of pixels that belong to gesture region. The purpose of the erosion is to eliminate probable gesture pixels that could be part of the neighbour background area:
- A morphological dilation with the same structuring element S is applied to the binary image B. The resulting image B_{backgr} consists of pixels that does not belong to a gesture region:

$$B_{backgr} = B \oplus S \qquad (9)$$

The result of selected gesture and non-gesture pixels are used as primary markers for the watershed segmentation.

Each region defined after the segmentation has to be identified by assigning a unique label. In the described system the region labelling employs a union-find algorithm [3]. The algorithm is a modification of the classical connected components detection [22]. Two scans of the segmented image are made and temporary labels are assigned to the pixels at the first scan according to their local similarity thus forming a set of sub-regions that are merged at the second scan. A hierarchical tree structure is used for the labels assigned at the first scan. The region union at the second scan selects the top parent label and assigns it to each pixel thus providing the relevant merge of the sub-regions.

C. Feature extraction

The segmentation stage provides regions that are considered as possible gesture regions. A set of features have to be extracted for each region and used for recognition of the gesture in that region. The main features that are used as gesture recognition features are region contour, convex hull, region centre and fingertip points.

To determine the contours of the region an algorithm for topological analysis by border tracking is used [23]. The algorithm builds hierarchical data structure of the contours in a binary image. The convex hull of a region is detected using the three coins algorithm [24]. A region contour and a region convex hull for the open hand gesture are shown in fig. 2.

Fig. 2. Open hand gesture region contour and convex hull

The location of the palm and the fingers in the gesture region are determined based on the convex hull. The space located between a closed set of points on the region contour and the corresponding line segment of the convex hull is called convexity defect (fig. 3). The convexity defects can be used for region feature extraction [9, 15].

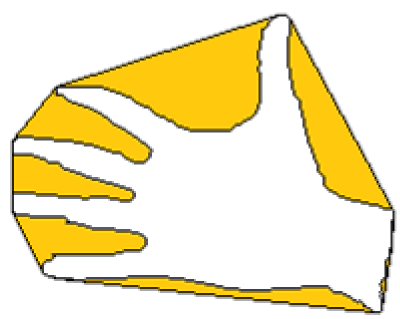


Fig. 3. Gesture region convexity defects

The convexity defects that are furthest from the convex hull represent the vertices of the palm region. A threshold is applied for the defects distance in order to filter the insignificant points and to delineate the palm region vertices (fig. 4).

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 1, January 2016)

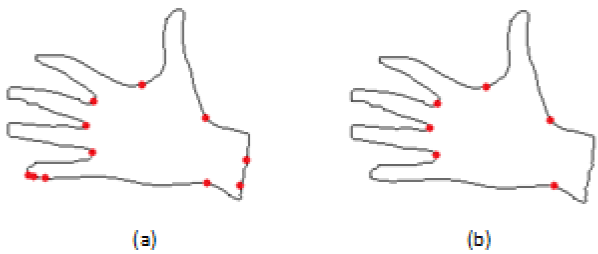


Fig. 4. Gesture region convexity defects: (a) no filtering; (b) distance threshold of 10 pixels

A combination of two approaches is used to locate the palm centre [9]:

- using a minimum cover circle centre P_{encl} for all convexity defects points in the palm [25];
- using the average coordinates of the convexity defects points P_{avg} (fig. 5).

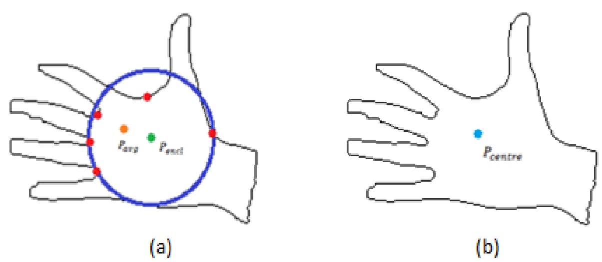


Fig. 5. Palm centre location: (a) minimum cover circle centre and averaged palm centre; (b) average coordinates of the convexity defects

Both approaches have drawbacks. The average coordinates of the convexity defects do not provide satisfactory results in the case of an open hand gesture when most of these points are concentrated at one end of the palm and the estimated centre is shifted to their position (fig.5a). On the other hand the minimum cover circle position may be displaced form the palm location in case one or several of the convexity defect points are located not on the palm but on the arm contour. In order to reduce the disadvantages of both approaches the described system uses a combination of the two approaches and the palm centre P_{centre} is located in as an average of the minimum cover circle centre and the average coordinates of the convexity defects (fig. 5b).

The gestures classification requires finger number and positions to be located. An algorithm for fingertip detection based on maximum acute angles on the hand contour is suggested in [16]. The angle of each point on the contour is calculated in order to find the maximum acute angles which are considered as peaks corresponding to fingertips. Instead of the angle value the maximum local cosine is determined as follows:

$$\cos\Theta=((X-X_{prevp})\cdot(X-X_{nextp}))/|X - X_{prevp}|\cdot|X - X_{nextp}| \qquad (10)$$

Where X is the current point, X_{prevp} and X_{nextp} are respectively the preceding and next points on the contour equidistant located from X (fig.6).

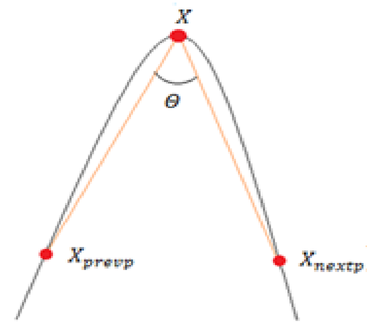


Fig. 6. Potential peak of a fingertip

In the described algorithm the minimum distance between points is selected to be 20 pixels. Local maximum is considered when the cosine is larger than 0.6. The potential fingertip peaks are further filtered using the convexity defects in order to remove any peaks between the fingers: the peaks that close to a defect are rejected. In order to remove several potential peaks corresponding to one and the same finger the peaks are filtered based on the distance between them.

The gesture features that are used for gesture classification and recognition are detected based on the geometric skeleton and the definition circle of the gesture region. The geometric skeleton comprises the line segments between the palm centre and each of the fingertips.

The definition circle has a centre that coincides with the palm centre and a radius equal to 75% of the maximum distance between the centre and a fingertip (fig. 7).

The geometric skeleton allows to find out the relative position of the fingers in a gesture. The definition circle is used to define the gestures based on the intersection points of the circle and the geometric skeleton. The definitional circle is traced in clockwise direction and for each finger the following features are identified:

- an intersection point P_{skel} of the circle with the line segment of the geometric skeleton corresponding to the finger;
- a point corresponding to the fingertip P_{tip};
- a central finger point P_{mid} located at the central axis of the finger and determined as a mid-point of the line segment between the two intersections of the finger contour and a circle with centre P_{tip} and radius P_{tip} P_{skel};

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 1, January 2016)

- a distance D_{rad} between the two P_{mid} points of the current finger and the next finger in a clockwise direction.

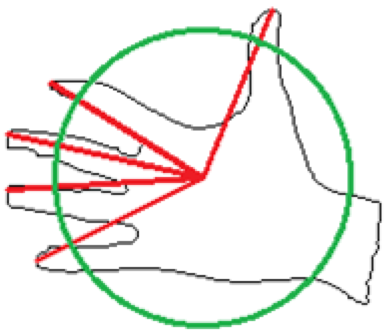


Fig. 7. Geometric skeleton and definition circle of the gesture region

D. Classification of gestures

The classifier used in the presented system is based on a decision tree. The classification follows comparison of the detected features for predefined criteria at several levels. The gesture recognition criteria are based on the number fingers, the angle between skeletal line segments, the angle between the fingers. At the first decision level only the number of the fingers is used. The angle between the skeletal line segments is used at the second level of the decision tree. At the third level the angle between the fingers is used calculated as a line segment between the central finger points and the fingertips. The angles between the skeletal line segments discriminate several gestures at the forth decision level.

Some of the twelve gestures that the system currently can recognise are shown on fig. 8.

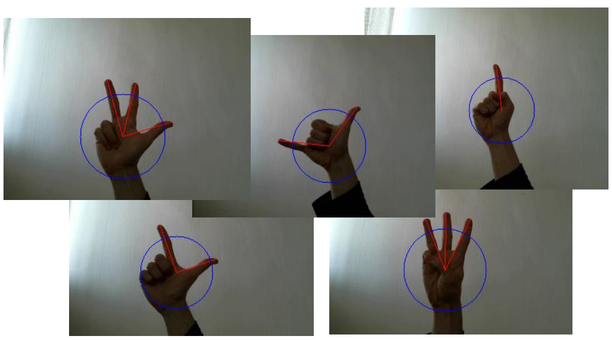


Fig. 8. Some of the gestures recognised by the system

A system limitation is its inability to recognise a gesture without fingers. For single finger gestures the utilized approach based on the geometrical skeleton and the definition circle does not allow detection of which exactly is the finger.

The suggested solution is not influenced by the use of left or right hand and also does not count on the position of the palm as face or back since the classification algorithm considers only the number of the fingers and the angular relations between them. Furthermore the gesture recognition based on such classifier is invariant to the palm rotation in the image.

After a particular gesture is recognise the system runs a user interface command associated with that gesture for example file open or close, start or terminate an application, etc.

IV. Experimental Results

The gesture recognition system for user interface control has two general concerns:

- detection of gesture regions;
- internal representation of gestures in the system.

The first is directly connected with suggested algorithms for detection of specific pixels, for example skin colour, and the second determines the classification logic and the gesture feature required for the gesture recognition. The system performance and requirements for work in real time are the reasons for selection of an algorithm for gesture region detection based on the pixel values. The algorithms based on contour detection, description and recognition are more computationally intensive especially in case of a complex background. The approach based on threshold operation is simple and fast and does not require maintenance of complex data structures in order to store and analyse a hierarchy of contours.

On the other hand the main disadvantage of that approach is the proper definition of criteria for gesture region and elimination of the lighting conditions impact. The criterion for detection of gesture region uses a combination of static colour model of the human skin and a histogram of an input region of gesture pixels. In order to minimize the influence of lighting conditions and to increase to recognition accuracy in pure lighting and low contrast images the systems allows adjustment and provide a skin or glove modes.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 6, Issue 1, January 2016)

An example comparison of the results for detecting skin regions using static model in both YcrCb and HSV colour spaces is shown on fig. 9. The results of application of histogram backprojection in YCrCb colour space for the input image on fig. 9a is shown on fig. 10.

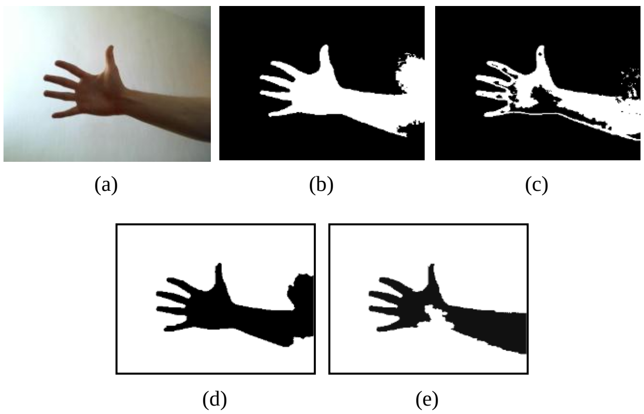


Fig. 9. Gesture pixels detection and gesture region segmentation using a static skin colour model: (a) input image; (b) skin pixels using YCbCr; (c) skin pixels using HSV; (d) region segmentation using YCbCr; (e) region segmentation using HSV

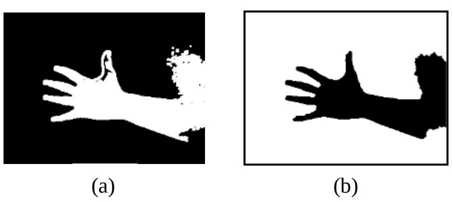


Fig. 10. Gesture pixels detection and region segmentation using histogram backprojection: (a) skin pixels; (b) region segmentation

The histogram approach applied for the human skin might give distorted results for the gesture region due to incorrect detection of the convexity defects, displacement of the palm centre or even incorrect detection of the number and position of the fingers. The hybrid approach suggested in the paper involves the static skin model in order to correct the possible errors and displacement of the histogram approach.

Results of the gesture region detection and segmentation using the hybrid approach are shown in fig.11.

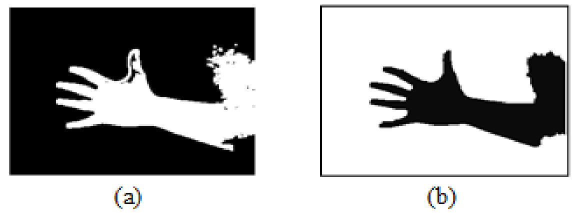


Fig. 11. Gesture pixels detection and region segmentation using a hybrid histogram backprojection and a skin model approach: (a) skin pixels; (b) region segmentation

Although the hybrid approach provides certain independence on the illumination, the system is still quite sensitive to the lighting conditions. That is in general one of the biggest problems of all image recognition systems with an image capturing sensor which works in the visible part of the spectrum. The transformation of the input image to YCrCb colour space provides some invariance and stability for gesture recognition of images captured under natural daylight conditions, but artificial light sources severely worsens the accuracy of gesture pixels detection based on human skin model. This limitation can be overcome using the glove mode of the system. Examples for detecting skin pixels in a room illuminated by a yellow light bulb are shown on fig. 12.

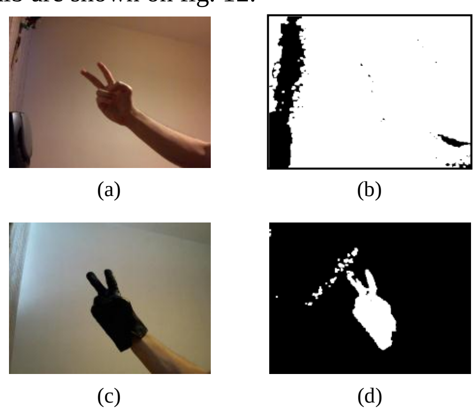
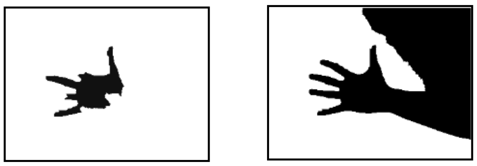


Fig. 12. Gesture pixels detection using a yellow light bulb illumination: (a) input image in a hand mode; (b) skin pixels for the hand gesture; (c) input image in a glove mode; (d) skin pixels for the glove gesture;

The accuracy of the gesture pixels detection depends on the resolution of the histogram for the Cr and Cb channels and the parameters of the histogram backpropagation.

At high resolution the gesture pixel criterion covers pixels with diverse values that decreases the average probability of pixels to be part of the gesture region and can lead to fragmentation in the segmentation of the gesture region. On the other hand too low histogram resolution can cause large number of background pixels to fall close to the detection criterion. The results for histogram resolution influence are shown on fig. 13.



(a) (b)

Fig. 13. Histogram resolution influence on the gesture pixels detection: (a) skin pixels for histogram resolution 200×200 pixels; (b) skin pixels for histogram resolution 40×40 pixels

V. CONCLUSION

The presented system allows user interface to be remotely controlled based on a hand gesture recognition. The algorithm for static gesture recognition uses a hybrid approach for skin pixels detection that combines skin model and histogram backpropagation approaches. The gesture region segmentation is implemented through watershed algorithm and morphological operations for seed markers selection. The palm centre is localised as a combination of the palm minimal covering circle centre and the positions of the palm convexity defect points. The gesture classification utilizes a hierarchical decision tree with rules based on the palm geometric skeleton and the palm definition circle.

The system provides two working modes using either hand skin or hand glove and allows initial calibration of the hand gesture pixel values. The glove mode is recommended in artificial light illumination conditions.

The utilized algorithm involves simple and fast processing thus making the system performance relevant for real time applications and allows further usage in embedded systems.

REFERENCES

[1] Mauser S., Burgert O., 2014, Touch-Free, Gesture-Based Control of Medical Devices and Software Based on the Leap Motion Controller, Studies in health technology and informatics Vol.196, pp.265-270.

[2] Murthy G., Jadon R., 2009, A Review of Vision-Based Hand Gestures, International Journal of Information Technology and Knowledge Management, Vol.2, No.2, pp.405-410.

[3] M. Bhuiyan, R. Picking, 2009, Gesture-Controlled User Interfaces, What Have We Done and What's Next?, Proc. of the Fifth Collaborative Research Symposium on Security, E-Learning, Internet and Networking (SEIN‘2009), Darmstadt, Germany, pp.59-60.

[4] Garg P., Aggarwal N., Sofat S., 2009, Vision Based Hand Gesture Recognition, World Academy of Science, Engineering and Technology, Vol.49, №1, pp.972-977.

[5] Yan X., Aimaiti N., 2011, Gesture-Based Interaction and Implication for the Future, Master Thesis, Department of Computing Science, Umea University, Sweden.

[6] Vezhnevets V., Sazonov V., Andreeva A., 2003, A Survey on Pixel-Based Skin Color Detection Techniques, Proc. of International Conf. GRAPHICON’2003, pp. 85-92.

[7] Basilio J., Torres G., Perez G., Medina L., Meana H., 2011, Explicit Image Detection using YCbCr Space Color Model as Skin Detection, Proc. of the 2011 American Conference on Applied Mathematics and the 5th WSEAS International Conference on Computer Engineering and Applications, pp.123-128.

[8] Zheng J., Hao C., Fan Y., Zang X., 2004, Adaptive Skin Detection Under Unconstrained Lighting Conditions Using a Bigaussian Model and Illumination Estimation.

[9] Chen W., 2011, Real-Time Palm Tracking and Hand Gesture Estimation Based on Forearm Contour, M. S. Thesis, Dept. Inform. Eng., National Taiwan University of Science and Technology.

[10] Ghotkar A., Kharate G., 2012, Hand Segmentation Techniques to Hand Gesture Recognition for Natural Human Computer Interaction, International Journal of Human Computer Interaction (IJHCI), Vol.3, №1, pp.15.

[11] Hobbs N., Chi L., 2012, US Patent 8942881.

[12] Swain M., Ballard D., 1992, Indexing via Color Histograms, Active Perception and Robot Vision, Springer Berlin Heidelberg, pp.261-273.

[13] Beucher S., 1991, The Watershed Transformation Applied to Image Segmentation. Scanning Microscopy-Supplement, pp.299.

[14] Rivest J., Delhomme J., 1991, Marker-Controlled Picture Segmentation Applied to Electrical Logging Images, Proc. SPIE, Nonlinear Image Processing II, San Jose, California.

[15] Youssef M., 2011, Hull Convexity Defect Features for Human Action Recognition, Doctoral Dissertation, University of Dayton, Dayton, OH, USA.

[16] Wen X., Niu Y., 2010, A Method for Hand Gesture Recognition Based on Morphology and Fingertip-Angle, Proc. of 2nd International Conference on Computer and Automation Engineering (ICCAE’2010), Vol. 1, pp.688-691.

[17] Microsoft Xbox, www.xbox.com.

[18] Leap Motion, www.leapmotion.com.

[19] Albiol A., Torres L., Delp E., 2001, Optimum Color Spaces for Skin Detection, 2001. Proc. of IEEE International Conference on Image Processing, Vol. 1, pp. 122-124.

[20] OpenCV Library, www.opencv.org.

[21] Dadgostar F., Sarrafzadeh A., 2006, An Adaptive Real-time Skin Detector Based on Hue Thresholding: A Comparison on Two Motion Tracking Methods, Pattern Recognition Letters, pp. 1342-1352.

[22] Rosenfeld A., Pfaltz J., (1966), Sequential Operations in Digital Processing, Journal of the ACM, Vol.13, №4, pp. 471-494.

[23] Suzuki S., Abe K., 1985, Topological Structural Analysis of Digitized Binary Images by Border Following, Journal Computer Vision, Graphics and Image Processing, Vol. 30, №1, pp.32-46.

[24] Graham R.L., 1972, An Efficient Algorithm for Determining the Convex Hull of a Finite Planar Set, Information Processing Letters, Vol. 1, pp.132-133.

[25] Skyum S., 1990, A Simple Algorithm for Computing the Smallest Enclosing Circle, DAIMI Report Series, №314.