BULETINUL INSTITUTULUI POLITEHNIC DIN IAŞI
Publicat de
Universitatea Tehnică „Gheorghe Asachi” din Iaşi
Tomul LVII (LXI), Fasc. 6, 2011
Secţia
ELECTROTEHNICĂ. ENERGETICĂ. ELECTRONICĂ

SKIN VIDEO SEGMENTATION AND TRAJECTORY RECOGNITION IN HUMAN–COMPUTER INTERFACE APPLICATIONS

BY

CIPRIAN DAVID[1,*], VASILE GUI[1], PEKKA NISULA[2] and VEIJO KORHONEN[2]

[1]"Politehnica" University of Timişoara
[2]Oulu University of Applied Sciences
Oulu, Finland

Received, May 31, 2011
Accepted for publication: July 26, 2011

Abstract. In this paper we address two problems involved in video segmentation for human–computer interface applications. First we propose an adaptive method for hand skin segmentation to enhance the hand feature detector. A new analysis colour space is used in order to obtain a reliable adaptive skin detector. The colour space is chosen so that the image has a sparser representation than those using the classical colour spaces. Mean shift mode detection is used to segment the cluster corresponding to the skin. Secondly, a dynamic hand gesture segmentation approach is presented. We are concerned with locally linear hand gestures. Trajectory points are obtained from the hand feature detector. The trajectory is filtered in order to obtain a smooth continuous version. Combining Radon space analysis and a few static postures we generate an encoding sequence of the gesture that can be used further to attach an action to be performed by the computer.

Key words: video segmentation; skin detection; gesture recognition; human–computer interface.

1. Introduction

By using body language humans are trying to emphasize different parts of their speech, hence its great importance. This inspired a lot of people to

*Corresponding author: *e-mail*: ciprian.david@etc.upt.ro

consider body language recognition as a communication protocol with the computer. Human–computer interactions (HCI) find many applications, as suggested by different authors (Penland, 1998; Liscar & Sziranyi, 2004). There are different types of information that an HCI can rely on. Such examples are: hand trajectory and posture, facial expression, head or eye movement, etc. Detailed surveys of different types of information and HCI can be found in the works performed by Jaimes & Sebe (2007) and Karray *et al*. (2008).

Our work relies on hand posture and trajectory recognition as a mean for designing an HCI. In this framework we address two problems. The first one concerns the enhancement of the hand segmentation by using a skin detector. The second problem is represented by the actual design of a communication protocol by hand posture recognition and trajectory segmentation.

Gestural interfaces have the advantage that they involve natural and simple movements for the user. This is the main reason for their popularity in many application areas like remote control, mixed reality or virtual reality.

Traditionally, a gestural HCI chain (Fig. 1), starts with background/ foreground segmentation. The foreground contains the interest object (*i.e*. the hand), to be detected. The next step uses some feature detections to identify the interest object, and finally the posture or the trajectory of the interest object is analysed and recognized so that the corresponding action can be performed.

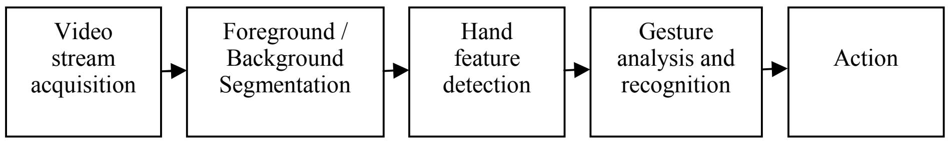


Fig. 1 – Gestural HCI chain.

Our work aims to improving the hand feature detection by including an adaptive skin detector. Also, we propose a method for gesture recognition that combines both dynamic trajectories and static postures recognition. The rest of the paper is organized as follows. Section **2** gives an overview of the state of the art concerning skin detection approaches and gesture recognition algorithms. In section **3** we present our approach. We start by shortly presenting our choice of background segmentation and hand feature detection approaches. Next, based on the analysis of previous work, an adaptive video skin detector is introduced. The section is ended by our gesture recognition algorithm presentation. Section **4** is dedicated to some results, conclusions and future development and research directions.

2. Related Work

2.1. Skin Detection

Numerous applications involve a skin detection stage. Such examples are found in video surveillance, face recognition, hand tracking, image indexing, etc. A particular type of approaches tries to construct a general model of the skin tone. Some of the research works in that sense include the use of Gaussian mixture models (GMM) (Jedynak *et al*., 2002), Bayesian estimation if beta mixture models (BMM) (Ma & Leijon, 2010), non parametric statistical models (Brand & Mason, 2000), some other parametric models (Jones & Rehg, 2002; Lu & Yoo, 2002), the recent proposed random forest classifier (Khan *et al*., 2010), or predefined colour space thresholds (Cheddad *et al*., 2009).

In an HCI application or a video application, a general model of the skin is not desirable. Usually, this type of applications is expected to work in various dynamic conditions. This poses a major problem for such general classifiers. The case of variable illumination is one example where the general skin tone model can give a wrong segmentation. Skin-like coloured objects can be present in the scene and a careful adjustment of the model parameters is required. Depending on the situation, optimum parameter selection is not trivial and in many cases almost impossible due to the model limitations.

Another difficult choice is represented by the colour space to be used for skin detection. Albiol *et al*. (2001) argues that every colour space has an optimum skin detection scheme. Nevertheless, there is an extensive work related to different colour spaces, each one reporting better results than the previous ones. The most popular colour spaces for skin segmentation are: Red-Green-Blue (RGB) (Ma & Leijon, 2010; Brand & Mason, 2000; Jones & Rehg, 2002; Aznaveh *et al*., 2008); Hue-Saturation-Value (HSV) (Hasem, 2009; Berens & Finlayson, 2000), or Luminance-Red chrominance-Blue chrominance (YCrCb) space (Lee & Yoo, 2002; Cheddad *et al*., 2009).

In our opinion a suitable choice of colour space is the space allowing the sparsest possible representation of the image. This way the corresponding skin vectors tend to cluster more obviously and the cluster identification or optimum parameter selection becomes a relatively simple procedure. This idea will be argued in the next section.

2.2. Gesture Recognition

Gesture recognition generally involves hand detection and recognition. In hand gesture recognition we can distinguish between two cases: hand posture recognition and trajectory segmentation.

Hand pose recognition is a challenging problem to solve, and usually involves a classification of the posture based on some detected hand features. Extensive research work are performed concerning this subject, and some of the research directions use Gaussian mixture models (Roomi *et al*., 2010), hidden

Markov models (Rigoll *et al*., 1998), model matching (Nielsen *et al*., 2004), fuzzy logic (Anderson *et al*., 2004; Bedregal *et al*., 2006), or some other features (Gui *et al*., 2009). However, the problem of hand pose recognition is a difficult one and we believe that for HCI applications using a complex model is not suitable due not only to computational reasons, but also because such methods are very complex and not always with satisfactory results.

A better choice for designing a human–computer communication protocol seems to be trajectory recognition. It is also a natural and simple action for the user. Not necessarily for HCI applications, but there is some work concerning trajectory segmentation. Some interesting ideas are presented by Shaogang *et al*. (1999) involving recognition by clustering hidden Markov states, or by Black & Jepson (1998), which the condensation algorithm for segmentation.

Our focus concerns the locally linear gestures. There is no maximum limit on the number of segments that compose the trajectory. Separators between different dynamic gestures are identified by a few static postures. These postures are intrinsically defined by the choice of hand feature detector. More details are given in the next section.

3. Proposed Approach

3.1. HCI Framework

As already mentioned, Fig. 1 shows the HCI general block scheme. Video acquisition is done by using a common webcam.

Our choice of foreground/background segmentation method is the one proposed by Kim *et al*. (2005). The background model is obtained by employing a codebook representation. For each pixel a separate codebook model is computed. The foreground is obtained from the difference between the current frame and the background model. It is demonstrated that the method is adaptable to illumination changes and can be updated very fast, thus being suitable for dynamic backgrounds. The model efficiency was tested successfully in both cases of indoor and outdoor backgrounds. An important advantage is the use of codebook representation permitting real-time processing, crucial for some HCI applications.

The hand feature detection is based on a robust clustering in the sparse feature space of extracted finger strips. The output of the detector is represented by fingertips positions. Also, the finger orientation can be extracted. An extension of the approach offers also the solution to detecting multiple fingers. The hand tracker robustness is proved in a series of tests for occlusion and accuracy. Multiple finger detection allows us to define a few hand postures, to be used as separators between dynamic gestures. However, the tracker performs more robustly in the case of one finger detection and this case is considered as default. Further details can be obtained from a work published by Gui *et al*. (2005).

3.2. Skin Video Segmentation

First problem that arises in the case of skin segmentation is choosing the colour space. A sparse representation of the image is thought to be more suitable in this case. In such a representation we expect that skin pixels will cluster in a more obvious manner, leaving a simpler segmentation task.

Classical colour spaces used for skin detection include RGB, HSV or YCrCb spaces. A very common approach is to define a threshold in order to extract the skin pixel cluster. General models assume a fixed threshold. Our aim is to conceive an adaptive algorithm, robust to illumination changes and camouflages.

In the case of using the RBG colour space a general set of thresholds for skin segmentation is given by the following equations (Kovac *et al*., 2003):

$$\begin{cases} R > 95, G > 40, B > 20, \\ \max\{R,G,B\} - \min\{R,G,B\} > 15, \\ |R - G| > 15, R > G, R > B. \end{cases} \tag{1}$$

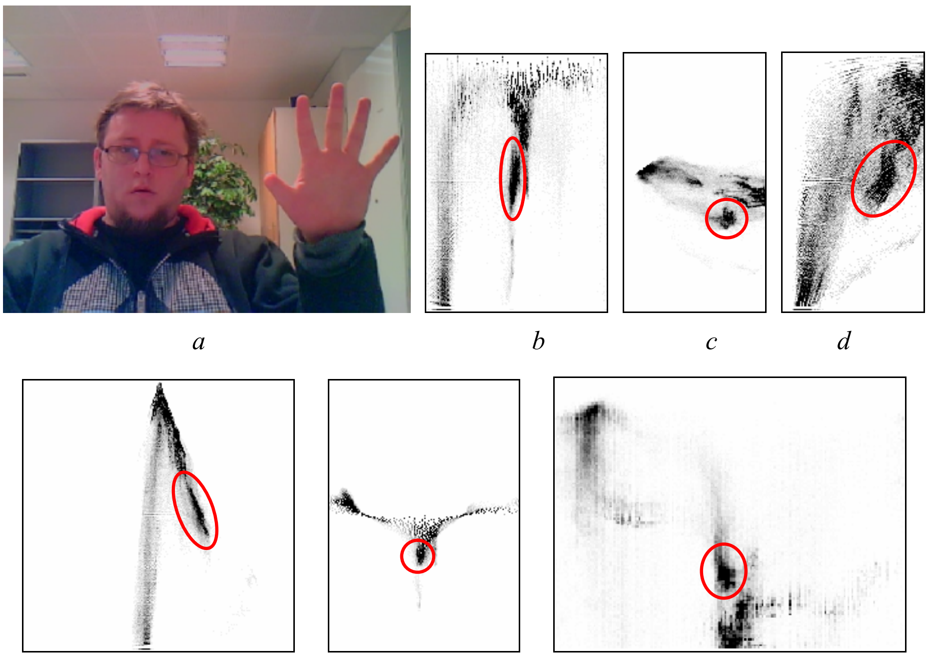
Even though this model is not at all robust, it offers a first helping clue in deciding the colour space, that there is an important red channel component for skin pixels. Some authors prefer normalizing the components but there is still a lack of robustness. In a recent work published by Porle *et al*. (2007), because it has the least skin representation, the normalized B-component is omitted, offering the second indication for the suitable colour space in skin detection.

a *b*

Fig. 2 – Skin detection results: *a* – threshold on Y-component; *b* – thresholds on H- and S-components.

Mainly using the Y-component of the YCrCb space, some interesting results are presented in a paper published by Khan *et al*. (2010). Zarit *et al*. (1999) have obtained good results using HSV space. Fig. 2 shows a result obtained with the approach proposed by Cheddad *et al*. (2009) (Fig. 2 *a*) and a skin segmentation obtained by imposing some thresholds on the *H* and *S* components (Fig. 2 *b*). There are still some false positives and false negatives in these detection results, but they demonstrate the potential of the components involved (*i.e. Y*, *H*, *S*).

Summing the above conclusions we investigate the use of an extended colour space containing the luminance, red chrominance, hue, and saturation components: YCrHS. In Fig. 3 the six corresponding histograms of the YCrHS space are shown, for an example image. These histograms represent all the possible projections on a 2-D plane defined by the four axes. The skin pixel clusters can be clearly distinguished (ellipses).



a *b* *c* *d*

e *f* *g*

Fig. 3 – YCrHS colour space projections: *a* – example image, histograms; *b* – HS; *c* – YCr; *d* – YS; *e* – CrS; *f* – HCr; *g* – HY.

Our proposition for skin detection is using only the HY and CrS histograms. Using all of them is redundant and experiments showed good results in using only these two. The algorithm starts by computing the HY and CrS histograms only for finger points obtained from the hand feature detector. These initial histograms give an indication about the skin cluster location in the next frame. The cluster is detected using a mean shift mode detector (Comaniciu & Meer, 2002). Having the skin corresponding histogram we perform a back projection in the image, thus obtaining two skin images. Logical

"AND" between the two skin images and the foreground, followed by the morphological operation used by Kim *et al.* (2005) to fill eventual detection gaps, will provide the skin segmented image. Further, this is used to complete the loop by applying the hand feature detector.

In Fig. 4 a general block scheme of our approach is presented. In the first frame we use the points detected by the hand feature detector to compute the two histograms. From these histograms an approximate location of the skin cluster is obtained. This hint about the location is used by the mean shift mode detector to segment the entire skin cluster in the histogram. A back projection operation gives the skin points of the frame. The last stage is represented by the gap filling morphological operation, and we obtain the segmented skin frame, which from now on will be used instead of the initialization step, to compute the two histograms of the following and find the approximate cluster location in the next frame. From the second frame, the skin segmentation result represents the input of the hand feature detector.

The most important contribution of our approach is the use of a new colour space for the representation of the skin. The YCrHS space assures a sparser representation of the image and thus the skin cluster is better separated from other image pixels. Also, the use of mean shift mode detection to segment the skin cluster has the advantage of robustness, and more, we do not use a threshold for the segmentation, hence the adaptability of our approach.

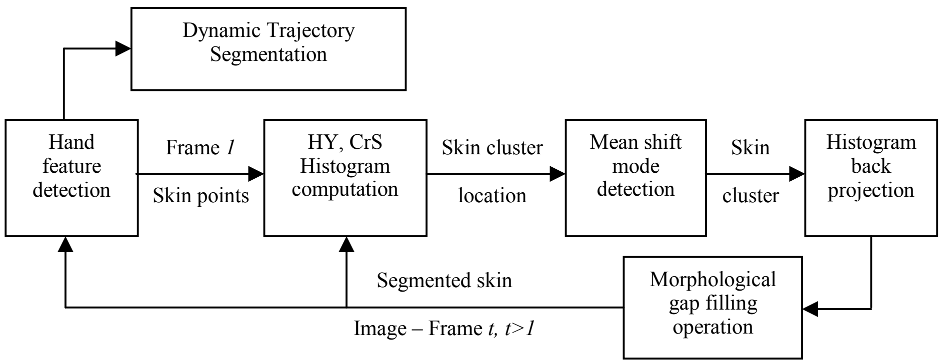


Fig. 4 – Video skin segmentation block scheme.

3.3. Trajectory Segmentation

The hand feature detector provides a sparse collection of trajectory points. The first step is to construct a smooth continuous trajectory by employing a tensor voting filtering technique. Each point linear character is encoded in a tensor containing local orientation information by means of a principal component analysis (PCA) on its neighborhood. Two voting (accumulation) stages will propagate these tensors. PCA on the accumulated tensors offers a local linear saliency map which also represents the reconstructed trajectory. Details about tensor voting technique can be found in a paper published by Tang *et al*. (2000).

Filtering the sparse trajectory eliminates outliers due to possible false detections of the hand tracker. A Radon space transform is performed on the trajectory. Each segment of the trajectory is encoded by the corresponding orientation. This orientation is obtained by mean shift mode detection (Comaniciu & Meer, 2002). The resulting sequence of directions identifies a specific gesture. We find that a quantization of the orientation space in 8 directions (*i.e*. N, S, E, W, SW, SE, NW, and NE) is sufficient enough to define a large number of dynamic gestures for an HCI application. The hand tracker allows static gestures to be defined, according to the detected hand posture: one finger (1F), two fingers (2F) and multiple fingers (*m*F). These static gestures are used to treat pauses between different gestures. For example, *m*F state will start the application, 2F state represents a *click*. Between two *clicks* trajectory points are collected but only if the 1F state is detected. 1F posture is preferred because the tracker is proven to be more robust in this case. Also, 1F state is considered as the default pose.

4. Results

Our experiments were conducted using a simple webcam to acquire the video sequence. The frame resolution is 640×480. Initially, the size of the mean shift mode detection window in the YCrHS space is obtained from the histograms of the skin points detected by the hand feature tracker in the first frame. For the following frames the window size is adaptively computed from the corresponding newly segmented skin points.

For the dynamic trajectory segmentation we used a rectangular window of size 41 in the case of PCA employed by tensor voting filtering. The curvature parameter value was chosen to be 0.2, and mean shift mode detection in the Radon space uses a 10×10 processing window. The implementation of both segmentation algorithms was made using the OpenCV library, and both algorithms work in real time.

Some skin segmentation results are shown in Figs. 5 and 6. In the first case a simple non perturbing background is present. The segmentation has good results. Pixels used to calculate the two histograms are marked by dark gray. Light gray (Fig 4 *b*) area represents the area where the hand tracker found one of the 1F, 2F or *m*F states. A partial camouflage case is presented in Fig. 6. Good skin segmentation is obtained. Still some false negatives are detected, partially due to the camouflage and partially due to the adaptation to histograms obtained from finger points. Also, the face skin is removed in this case because of the adaptation combined with the morphological gap filling operation. However, for the purpose intended (to be used with the specific case of hand tracker) the result is more than satisfying. The finger strips feature space is not affected.

Compared to classical image skin segmentation proposals our approach has the advantage of using a sparser representation in the new YCrHS space. This representation allows an adaptable strategy by employing the mean shift mode detection, and, of course, by exploiting the fact that we deal with video sequences instead of images. The adaptive skin cluster detection of our approach makes the entire algorithm robust to illumination changes and also increases the robustness in camouflage situations. For classical skin detectors illumination changes represent a serious problem, and moreover, in camouflage situations a careful tuning of the parameters is needed, which is not a trivial operation.

a *b*

Fig. 5 – Skin detection in a non perturbed case: *a* – initial hand detection, *b* – skin segmentation result in the next frame.

Fig. 6 – Skin detection in a partial camouflaged case.

Some final results of the trajectory segmentation are shown in Fig. 7. Experiments on synthetic generated trajectories proved the reliability of the dynamic hand segmentation algorithm. A highly perturbed, hypothetic, trajectory gave a 96% recognition rate for an experiment with 100 realizations; and 100% recognition rate for moderate perturbations, which are more characteristic for a human gestural input. For these simulations, in both cases, we used a set of 100 synthetic perturbed trajectories.

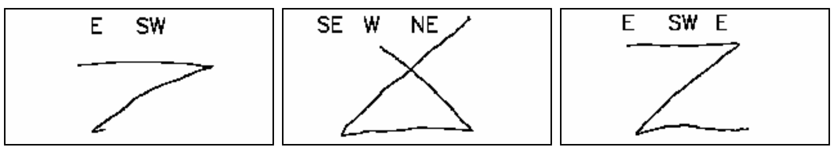


Fig. 7 – Trajectory segmentation results.

Our trajectory segmentation approach is also more robust compared to somewhat similar proposals concerning trajectory segmentation. Most of the previous proposals use a set of predefined gestures to train a model. Our approach proposes a methodology to define the set of gestures. Our proposal allows a large set of possible dynamic gestures to be defined. Also, using a trained gesture model which is not fully robust, some false detection can appear. From another point of view, previous proposals attack the problem of inter-gesture pauses. We solved this problem by defining a small number of static hand poses that are used to mark the pauses between consecutive dynamic gestures.

In conclusion, the paper proposed two segmentation methods to be used in HCI applications. The first concerns adaptive video skin segmentation used to enhance hand detection. It is shown that even in perturbed situations, like camouflage, good results are obtained.

The second segmentation approach concerns dynamic gesture recognition. The algorithm is validated by a perfect recognition score, obtained for moderate perturbed synthetic trajectories that simulate real human gestural inputs. Future work will include enhancing more the skin detection robustness for difficult situations of background and for trajectory segmentation extending the approach for circular trajectories.

Acknowledgments. This paper was supported by the project "Development and Support of Multidisciplinary Postdoctoral Programmes in Major Technical Areas of National Strategy of Research – Development – Innovation" 4D-POSTDOC, contract no. POSDRU/89/1.5/S/52603, project co-funded by the European Social Fund through Sectoral Operational Programme Human Resources Development 2007-2013.

This work was supported by the national grant ID 931, contr. 651/19.01.2009 and by the UbiGo project, financed by The European Regional Development Fund, Council of Oulu Region, Oulu University of Applied Sciences and City of Oulu, Finland.

REFERENCES

Albiol A., Torres L., Delp E., *Optimum Color Spaces for Skin Detection*. Proc. of IEEE Internat. Conf. on Image Process., Thessaloniki, Greece, October 7-10, 2001, 122-124.

Anderson L., Purdy J., Viant W., *Variations on a Fuzzy Logic Gesture Recognition Algorithm*. Proc. of the 2004 ACM SIGCHI Internat. Conf. on Adv. in Comp. Entert. Technol., Singapore, June 3-5, 2004, 280-283.

Aznaveh M. M., Mirzaei H., Roshan E., Saraee M., *A New and Improved Skin Detection Method Using RGB Vector Space*. Proc. of IEEE Internat. Multi-Conf. on Syst., Sign. A. Dev., Amman, Jordan, July 20-22, 2008, 1-5.

Bedregal B.C., Costa A.C.R., Dimuro G.P., *Fuzzy Rule-Based Hand Gesture Recognition*. Artif. Intell. in Theory a. Practice, IFIP, **217**, Springer Boston, 2006, 285-294.

Berens J., Finlayson G.D., *Log-Opponent Chromaticity Coding of Colour Space*. Proc. of IEEE Internat. Conf. on Pattern Recogn., Barcelona, Spain, September 3-8, 2000, **1**, 206–211.

Black M.J., Jepson A.D., *Recognizing Temporal Trajectories Using the Condensation Algorithm*. Third IEEE Internat. Conf. on Autom. Face a. Gesture Recogn., Nara, Japan, April 1998.

Brand J., Mason J. S., *A Comparative Assessment of Three Approaches to Pixel-Level Human Skin-Detection*. Proc. of IEEE Internat. Conf. on Pattern Recogn., Barcelona, Spain, Sept. 3-7, 2000, **1**, 1056-1059.

Cheddad A., Condell J., Curran K., McKevitt P., *A Skin Tone Detection Algorithm for an Adaptive Approach to Steganography*. Signal Process., **89**, 2465-2478 (2009).

Comaniciu D., Meer P., *Mean Shift: A Robust Approach Toward Feature Space Analysis*. IEEE Trans. Pattern Anal. Machine Intell., **24**, *5*, 603-619 (2002).

Gui L., Thiran J.-P., Paragios N., *Cooperative Object Segmentation and Behavior Inference in Image Sequences*. Internat. J. of Comp. Vision, **84**, *2* (2009).

Gui V., Popa G., Nisula P., Korhonen V., *Finger Detection in Video Sequences Using a New Sparse Representation*. Acta Techn. Napoc. (accepted for publishing in 2011).

Hashem H.F., *Adaptive Technique for Human Face Detection Using HSV Color Space and Neural Networks*. Radio Sci. Conf. (NRSC 2009), National, New Cairo, March 17-19, 2009.

Jaimes A., Sebe N., *Multimodal Human Computer Interaction: A Survey*. Comp. Vision a. Image Underst., **108**, *1-2*, 116-134 (2007).

Jedynak B., Zheng H., Daoudi M., Barret D., *Maximum Entropy Models for Skin Detection*. Techn. Rep. Publ. IRMA, **57**, *13* (2002).

Jones M. J., Rehg J. M., *Statistical Color Models with Application to Skin Detection*. Internat. J. of Comp. Vision, **46**, *1*, 81-96 (2002).

Karray F., Alemzadeh M., Saleh J. A., Arab M. N., *Human-Computer Interaction: Overview on State of the Art*. Internat. J. on Smart Sensing a. Intell. Syst., **1**, *1*, 137-159 (2008).

Khan R., Hanbury A., Stoettinger J., *Skin Detection: a Random Forest Approach*. Proc. of 2010 IEEE 17th Internat. Conf. on Image Process. , Hong Kong, September 26-29, 2010.

Kim K., Chalidabhongse T.H., Harwood D., Davis L., *Real-Time Foreground–Background Segmentation Using Codebook Model*. Real-Time Imaging, **11**, *3*, 167-256 (2005).

Kovac J., Peer P., Solina F., *Human Skin Colour Clustering for Face Detection*. Proc. of Internat. Conf. on Comp. as a Tool, Slovenia, September 22-24, 2003, **2**, 144-148.

Lee J. Y., Yoo S. I., *An Elliptical Boundary Model for Skin Color Detection*. Proc. of the Internat. Conf. on Imag. Sci., Syst., a. Technol., Las Vegas, Nevada, USA, June 24-27, 2002.

Licsar A., Sziranyi T., *Hand Gesture Recognition in Camera-Projector System*. Internat. Workshop on Human-Comp. Interaction, Lecture Notes in Computer Sci., **3058**, Springer, Prague, May 16, 2004, 81-91.

Ma Z., Leijon A., *Human Skin Color Detection in RGB Space with Bayesian Estimation of Beta Mixture Models*. 18th Europ. Signal Proc. Conf. (EUSIPCO-2010), Aalborg, Denmark, August 23-27, 2010.

Nielsen E.S., Canalis L.A., Tejera M.H., *Hand Gesture Recognition for Human–Machine Interaction*. J. of WSCG, **12**, *1-3* (2004).

Penland A., *Smart Rooms: Machine Understanding of Human Behavior*. Comp. Vision for Human–Machine Interaction, Cambridge University Press, 1998, 3-21.

Porle R.R., Chekima A., Wong F., Sainarayanan G., *Wavelet-Based Skin Segmentation for Detecting Occluded Arms in Human Body Pose Modeling System*. Proc. of Internat. Conf. on Intell. A. Adv. Syst., Malaysia, November 25-28, 2007, 764-769.

Rigoll G., Kosmala A., Eickeler S., *High Performance Real-Time Gesture Recognition Using HMM*. In *Gesture and Sign Language in Human–Computer Interaction*, (I. Wachsmuth *et al*., Eds.), LNAI, **1371**, Springer, 1998, 69-80.

Roomi S. M. M., Priya R. J., Jayalakshmi H., *Hand Gesture Recognition for Human-Computer Interaction*. J. of Comp. Sci., **6**, *9*, 1002-1007 (2010).

Shaogang G., Walter M., Psarrou A., *Recognition of Temporal Structures: Learning Prior and Propagating Observation Augmented Densities via Hidden Markov States*. Proc. Seventh Internat. Conf. on Comp. Vision, Kerkyra, Greece, September 20-27, 1999, **1**, 157-162.

Tang C.-K., Lee M.-S., Medioni G., *Tensor Voting*. Ed. Elsevier, 2000.

Zarit B.D., Super B., Francis J., Quek K.H., *Comparison of Five Color Models in Skin Pixel Classification*. Proc. of the Internat. Workshop on Recogn., Anal., a. Tracking of Faces a. Gestures in Real-Time Syst., Corfu , Greece, September 26-27, 1999, 58–63.

SEGMENTAREA VIDEO A PIELII ŞI RECUNOAŞTEREA TRAIECTORIILOR ÎN APLICAŢII CU INTERFEŢE OM–MAŞINĂ

(Rezumat)

Se prezintă două metode implicate în segmentarea video pentru aplicații cu interfețe om–mașină. Prima propunere se referă la o metodă adaptivă de segmentare video a pielii pentru îmbunătățirea detectorului de mână. Un nou spațiu color de analiză este utilizat pentru o mai mare adaptivitate şi robustețe. Acest spațiu este ales astfel încât imaginea să aibă o reprezentare cât mai « sparse » față de spațiile color clasice. Detecția de mod prin algoritmul „mean shift" este utilizată pentru a segmenta pielea. În a doua parte este introdusă o metodă de segmentare a gesturilor dinamice. Sunt considerate traiectorii local liniare. Punctele care formează traiectoria sunt obținute de la un detector de caracteristici ale mâinii. Această traiectorie este filtrată pentru a se obține o versiune de traiectorie continuă şi netedă. Combinând o analiză în spațiul transformatei Radon şi utilizarea câtorva gesturi statice, se generează o secvență care codează gestul dinamic şi care poate fi mai departe folosită pentru a defini o acțiune a fi realizată de calculator.