

# A Brief Review of Vision Based Hand Gesture Recognition

GEORGIANA SIMION<sup>(1)</sup>, VASILE GUI<sup>(2)</sup>, MARIUS OTEȘTEANU

Department of Communication  
Politehnica University of Timișoara  
Bd. Vasile Pârvan No.2, Timișoara  
ROMANIA

georgiana.simion@etc.upt.ro, vasile.gui@etc.upt.ro, marius.otesteanu@etc.upt.ro

*Abstract:* - The evolution of user interfaces shapes the changes in Human-Computer Interaction (HCI). Direct use of hand as an input device is an attractive method for providing natural HCI. The applications of gesture recognition are manifold, ranging from sign language to medical rehabilitation to virtual reality. In this paper we present a brief review of vision based hand gesture recognition.

*Key-Words:* - hand gestures, recognition, model based approach, view based approach, human computer interaction, applications

## 1 Introduction

People perform various gestures in their daily lives. It is in our nature to use gestures in order to improve the communication between us. Try to imagine speaking with a person who makes no gesture. It is very difficult to understand if your message is clear for him or her, if he or she agrees with your saying, in other words it is very hard to guess what type of reaction your message produces. Between all kind of gestures that we perform, hand gestures play an important role. Hand gestures can help us say more in less time. In these days, computers have become an important part in our lives, so why not use hand gesture in order to communicate with them.

The direct use of the hand as an input device is an attractive method for providing natural Human-Computer Interaction. Two approaches are commonly used to interpret gestures for Human Computer Interaction.

*Methods Which Use Data Gloves:* Since now, the only technology that satisfies the advanced requirements of hand-based input for HCI is glove-based sensing. This method employs sensors (mechanical or optical) attached to a glove that transducers' finger flexions into electrical signals for determining the hand posture. Several drawbacks make this technology not so popular: first of all interaction with the computer-controlled environment loses naturalness and easiness the user is forced to carry a load of cables which are connected to the computer and it also requires calibration and setup procedures.

*Methods which are Vision Based:* Computer vision based techniques have the potential to provide more natural and non-contact solutions,

there are non invasive and are based on the way human beings perceive information about their surroundings. Although it is difficult to design a vision based interface for generic usage, yet it is feasible to design such an interface for a controlled environment but has no lack of challenges including accuracy, processing speed.

This paper is organized as follows: In section 2 we provide a survey on vision based hand gesture recognition. In section 3 we present various applications areas for gesture recognition and in section 4 we give the conclusions.

## 2 Problem Formulation

The approaches to Vision based hand gesture recognition can be divided into two categories: 3 D hand model based approaches and appearance based approaches [1].

### 2.1 Model based approach

Model based approaches attempt to infer the pose of the palm and the joint angles, this approach is ideal for realistic interactions in virtual environments. By large, the approach consists of searching for the kinematic parameters that brings the 2D projection of a 3D model of hand into correspondence with an edge-based image of a hand.

The model of the hand can be more or less elaborated.

A 3D model with 27 degrees of freedom (DOF) was introduced and, it has been used in many

studies and it is shown in Fig. 1 a.

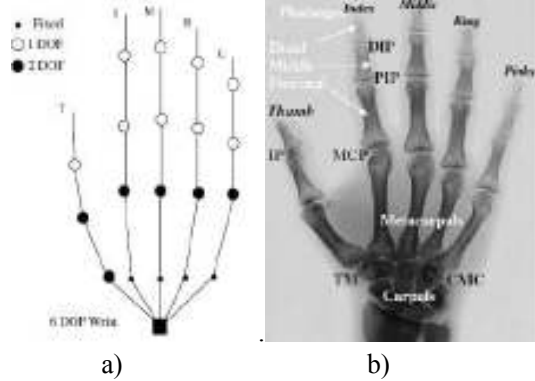


Fig.1. Skeletal hand model: (a) Hand anatomy, (b) the kinematic model according to [7]

The CMC joints are assumed to be fixed, which quite unrealistically models the palm as a rigid body. The fingers are modeled as planar serial kinematic chains attached to the palm at anchor points located at MCP joints.

Over the years the kinematic model was improved by adding extra twist motion to MCP joints [2], [3] introducing one flexion/extension DOF to CMC joints [4] or using a spherical joint for TM [5]

Rehg and Kanade [6] proposed one of the earliest model based approaches to the problem of bare hand tracking. They used a 3D model with 27 DOF for their system called DigitEyes.

Heap et al.[8] proposed a deformable 3D hand model and modeled the entire surface of the hand by a surface mesh constructed via PCA from training examples.

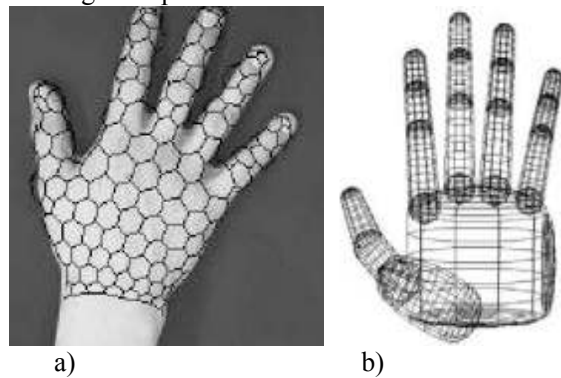


Fig.2. a) Hand tracking using 3D Point Distribution Model from [8] and b) Quadrics-based hand model from [9]

Stenger et al. [9] used quadrics as shape primitives. The use of quadrics to build the 3D model yields a practical and elegant method for generating the contours of the model, which are then compared with the image data. The pose of the hand model is estimated with an Unscented Kalman filter (UKF), which minimizes the geometric error

between the profiles and edges extracted from the images.

In [10] they have reformulated the problem within a Bayesian (probabilistic) framework. Bayesian approaches allow for the pooling of multiple sources of information (e.g. system dynamics, prior observations) to arrive at both an optimal estimate of the parameters and a probability distribution of the parameter space to guide future search for parameters. On contrary to Kalman filter approach, Bayesian approaches allow nonlinear system formulations and non-Gaussian (multi-modal) uncertainty (e.g. caused by occlusions) at the expense of a closed-form solution of the uncertainty.

In [12], a model-based visual hand posture tracking algorithm is proposed to guide a dexterous robot hand. The approach adopts a 3D model-based framework with full-DOF kinematic and an effective measurement method based on chamfer distance for both silhouette and edges. GA is integrated to traditional PF as a solution of high-dimensional and multi-modal tracking. Experimental results show a significant improvement of tracking performance compared with traditional PF.

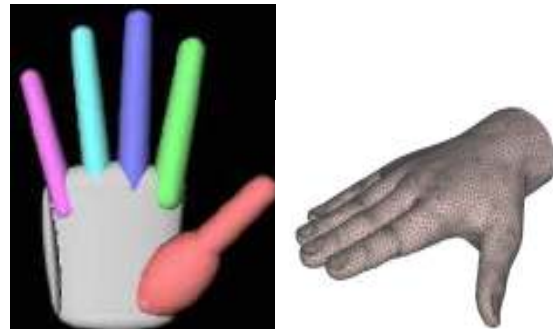


Fig.3. a) The 3D model presented in [12], b) The 3D model presented in [13]

In [13] proposed a realistic 3D model of the hand. This deformable model consists of a polygonal skin, driven by an underlying skeleton. A new pose is computed by linearly blending the motions that each skin vertex would undergo when rigidly coupled to a subset of the skeleton joints. The model is used in a particle filter framework. A novel algorithm which combines the SMD (Stochastic Meta-Descent) optimization with a particle filter to form 'smart particles' is proposed. After propagating the particles, SMD is performed and the resulting new particle set is included such that the original Bayesian distribution is not altered.

In [14,15] an approach to the recovery of geometric and photometric pose parameters of a 3D model with 28 DOF from monocular image sequences is presented.

The 3D hand pose, the hand texture and the illuminant are dynamically estimated through minimization of an objective function. Derived from an inverse problem formulation, the objective function enables explicit use of texture temporal continuity and shading information, while handling important self-occlusions and time-varying illumination. The minimization is done efficiently using a quasi-Newton method, for which was proposed a rigorous derivation of the objective function gradient.

In [16] truncated quadrics are used to build a 3D hand model where the DOF for each joint correspond to the DOF of a real hand.

Quadratic chamfer distance function is used to compute the edge likelihood and the silhouette likelihood is performed by a Bayesian classifier and online adaptation of skin color probabilities. Particle filtering is used to track the hand by predicting the next state of 3D hand model.

The 3D hand models are articulated deformable objects with many degrees of freedom; a very large image database is required to cover all the characteristic shapes under different views. Another common problem with model based approaches is the problem of feature extraction and lack of capability to deal with singularities that arise from ambiguous views.

## 2.2 Appearance based approaches

Appearance-based models are derived directly from the information contained in the images and have traditionally been used for gesture recognition. No explicit model of the hand is needed; this means no internal degrees of freedom to be specifically modeled.

When only the appearance of the hand in the video frames is known, differentiating between gestures is not as straight forward as with the model based approach. The gesture recognition will therefore typically involve some sort of statistical classifier based on a set of features that represent the hand. In many gesture applications all that are required is a mapping between input video and gesture. Therefore, many have argued that the full reconstruction of the hand is not essential for gesture recognition. Instead many approaches have utilized the extraction of low-level image measurements that are fairly robust to noise and can be extracted quickly. Low-level features that have been proposed in the literature include: the centroid of the hand region [16], principle axes defining an elliptical bounding region of the hand, and the optical flow/affine flow [17] of the hand region in a scene.

Another approach is to look for skin colored regions in the image. This is a very popular method [18], [19], [20], [21] but has some drawbacks. First, skin color detection is very sensitive to lighting conditions. While practicable and efficient methods exist for skin color detection under controlled (and known) illumination, the problem of learning a flexible skin model and adapting it over time is challenging. Lindberg [16] used scale-space color features to recognize hand gestures. Multi scale features can be found in an image at different scales. Therefore, the hand can be described as one bigger blob feature for the palm, having smaller blob features representing the finger tips which are connected by some rigid features. Furthermore, it was proposed to perform the feature extraction directly in the color space, as this allows the combination of probabilistic skin colors directly in the extraction phase. The advantage of directly working on a color image lies in the better distinction of hand and background regions, but the authors showed real time application only with no other skin colored objects present in the scene.

Another approach is to use the eigenspace. Given a set of images, eigenspace approaches construct a small set of basis images that characterize the majority of the variation in the training set and can be used to approximate any of the training images. To reconstruct an image in the training set, a linear combination of the basis vectors (images) are taken, where the coefficients of the basis vectors are the result of projecting the image to be reconstructed on to the respective basis vectors. In [17] an approach for tracking hands by an eigenspace approach is presented. The authors provide three major improvements to the original eigenspace approach formulation, namely, a large invariance to occlusions, some invariance to differences in background from the input images and the training images, and the ability to handle both small and large affine transformations (i.e. scale and rotation) of the input image with respect to the training images. The authors demonstrate their approach with the ability to track four hand gestures using 25 basis images.

In the last years is noticeable a new trend, more and more approaches use invariant local features [24], [25], [26], [27], [28], [29], [30], [31].

In [24], Adaboost learning algorithm with SIFT features is used. The Scale Invariant Feature Transform (SIFT) introduced by Lowe [32] consists of a histogram representing gradient orientation and magnitude information within a small image patch. SIFT is a rotation and scale invariant feature and is robust to some variations of illuminations,

viewpoints and noise. The accuracy of multi-class hand posture recognition is improved by the sharing feature concept. However, different features such as contrast context histogram need to be studied and applied to accomplish hand posture recognition in real time.

In [25] Bag-of-Words representation (BoW) and SIFT features is used. In a typical BoW representation, “interesting” local patches are first identified from an image, either by densely sampling, or by an interest point detector. These local patches, represented by vectors in a high dimensional space, are often referred to as the key points. The bag-of-words methods main idea is to quantize each extracted key point into one of the visual words, and then represent each image by a histogram of visual words. A clustering algorithm is generally used to generate the visual words dictionary. In [25] K-means algorithm has been used for clustering. A multi-class SVM was used to train the classifier model. In the testing stage, the keypoints were extracted from every image captured from the webcam and fed into the cluster model to map them with one (Bag-of-words) vector, which is finally fed into the multi-class SVM training classifier model to recognize the hand gesture.

In [26] the ARPD descriptor (Appearance and Relative Position Descriptor) is proposed. This descriptor includes color histogram, relative-position information, and SURF [33]. The process of constructing ARPD includes two steps: extracting SURF keypoints and color histogram from images, and computing relative-position information of every keypoint within images, the relative-position information is also included as part of ARPD. The ARPD was used in the BoW representation. The BoW was used to detect and recognize hand posture based on sliding-window framework. To meet real-time request, several approaches were proposed to speed up hand posture recognition process. In tracking process, CAMESHIFT algorithm to track hand motion and a strategy based on histogram to reinitialize tracking process were used.

In [27] compositional techniques are used for hand posture recognition. A hand posture representation is based on compositions of parts: descriptors are grouped according to the perceptual laws of grouping [34] obtain a set of possible candidate compositions. These groups are a sparse representation of the hand posture based on overlapping subregions.

The detected part descriptors are represented as probability distributions over a codebook which is obtained in the learning phase. A composition is a

mixture of the part distributions. From all candidate compositions, relevant compositions must be selected. There are two types of relevant compositions: those compositions that occur frequently in all categories and also those which are specific for a category. The category posterior of compositions is learned in the training phase, and it is a measure of relevance. The entropy of the category posterior helps to discriminate between categories. A cost function is obtained by combining the priors of the prototypes and the entropy. The process of recognition is based on bag of composition method, where a discriminative function is defined.

In [28] Maximally Stable Extremal Region (MSER) detector and color likelihood maps are used for hand tracking. Such a combination allows performing repeated figure/ground segmentation in every frame in an efficient manner.

The MSER detector is one of the best interest region detectors in computer vision [35]. MSER detection is mostly applied to single gray scale images, but the method can be easily extended for analysis of color images by defining a suitable ordering relationship on the color pixels. In general the MSER detector finds bright connected regions which have consequently darker values along their boundaries. The set of MSERs is closed under continuous geometric transformations and is invariant to affine intensity changes. Furthermore MSERs are detected at all scales. Therefore, due to these properties MSER detection is suited for segmentation purposes.

In [29], [30], [31] Haar like features are used for the task of hand detection. Haar like features focus more on the information within a certain area of the image rather than each single pixel. To improve classification accuracy and achieve realtime performance, AdaBoost learning algorithm that can adaptively select the best features in each step and combine them into a strong classifier can be used. The training algorithm based on AdaBoost learning algorithm takes a set of “positive” samples, which contain the object of interest and a set of “negative” samples, i.e., images that do not contain objects of interest.

This invariant features allowed us to model the hand as collection of characteristic parts. Key points or characteristic regions are extracted. Using such features the hand gesture is decomposed in simpler parts which are easier to recognize. This approach has major advantages: even if some parts are missing a gestures still can be recognized, so there are robust to partials occlusions, changes in view point and considerable deformations. Bag of Words

methods and compositional methods become more and more popular in hand gesture recognition. These techniques have been studied in many diverse fields such as linguistics, logic, and neuroscience, but compositionality is especially evident in the syntax and semantics of language where a limited number of letter scan form a huge variety of words and sentences. In computer vision these techniques are used in the context of a general problem: categorization. Using these techniques we address also to the semantic gap that exists between the low level features and high level representations. The hand posture is no longer modeled as a whole. These characteristic regions are assembled to form compositions; these compositions at their turn can be group in compositions of compositions and so on.

### 3 Application Areas

There is a large variety of applications which involves hand gestures. Hand gestures can be used to achieve natural human computer interaction for virtual environments, or there can be used to communicate with deaf and dumb. An important application area is that of vehicle interfaces.

In this section an overview of few application areas is given.

*Virtual Reality:* Gestures for virtual and augmented reality applications have experienced one of the greatest levels of uptake interactions [36] or 2D displays that simulate 3D interactions [37].

*Robotics and Telepresence:* When robots are moved out of factories and introduced into our daily lives they have to face many challenges such as cooperating with humans in complex and uncertain environments or maintaining long-term human-robot relationships. Telepresence and telerobotic applications are typically situated within the domain of space exploration and military-based research projects.

The gestures used to interact with and control robots are similar to fully-immersed virtual reality interactions, however the worlds are often real, presenting the operator with video feed from cameras located on the robot [38]. Here, gestures can control a robots hand and arm movements to reach for and manipulate actual objects, as well its movement through the world.

Hand gesture recognition for robotic control is presented in [24, 39]

*Desktop and Tablet PC Applications:* In desktop computing applications, gestures can provide an alternative interaction to the mouse and keyboard [40]. Many gestures for desktop computing tasks involve manipulating graphics, or

annotating and editing documents using pen-based gestures [41]. This year eyeSight introduced gesture recognition Technology for Android Tablets and Windows-based Portable Computers [50].

*Sign Language:* Sign language is an important case of communicative gestures. Since sign languages are highly structural, they are very suitable as testbeds for vision algorithms [42]. At the same time, they can also be a good way to help the disabled to interact with computers. Sign language for the deaf (e.g. American Sign Language) is an example that has received significant attention in the gesture literature [43, 44, 45 and 46].

*Vehicle interfaces:* A number of hand gesture recognition techniques for human vehicle interface have been proposed time to time [47,48]. The primary motivation of research into the use of hand gestures for in-vehicle secondary controls is broadly based on the premise that taking the eyes off the road to operate conventional secondary controls can be reduced by using hand gestures.

*Healthcare:* Wachs et al. [49] developed a hand-gesture recognition system that enables doctors to manipulate digital images during medical procedures using hand gestures instead of touch screens or computer keyboards. A sterile human-machine interface is of supreme importance because it is the means by which the surgeon controls medical information, avoiding patient contamination, the operating room and the other surgeons. The gesture based system could replace touch screens now used in many hospital operating rooms which must be sealed to prevent accumulation or spreading of contaminants and requires smooth surfaces that must be thoroughly cleaned after each procedure – but sometimes aren't. With infection rates at hospitals now at unacceptably high rates, the hand gesture recognition system offers a possible alternative.

### 4 Conclusion

In this paper a review of vision based hand gesture recognition methods has been presented. In the last years remarkable progress in the field of vision based hand gesture recognition has been done. Further research in the areas of feature extraction, classification methods and gesture representation are required to realize the ultimate goal of humans interfacing with machines on their own natural terms.

It is obviously that the near future belongs to hand gesture recognition. Probably sooner that one may think the surrounding devices will be hand gesture interfaced.

## ACKNOWLEDGMENT

<sup>(1)</sup> This paper was supported by the project "Develop and support multidisciplinary postdoctoral programs in primordial technical areas of national strategy of the research - development - innovation" 4D-POSTDOC, contract nr. POSDRU /89/1.5/S/52603, project co-funded from European Social Fund through Sectorial Operational Program Human Resources 2007-2013.

<sup>(2)</sup> This work was supported by the national grant ID 931, contr. 651/19.01.2009.

## References:

- [1] H. Zhou, T.S. Huang, Tracking articulated hand motion with Eigen dynamics analysis, *In Proc. Of International Conference on Computer Vision*, Vol 2, 2003, pp. 1102-1109
- [2] Bray M., Koller-Meier E., Gool L.V., Smart particle filtering for 3D hand tracking. *Sixth IEEE International Conference on Automatic Face and Gesture Recognition* (2004): 675
- [3] Bray M., Koller-Meier E., Muller P, Gool L.V., Schraudolph N.N., 3D Hand tracking by rapid stochastic gradient descent using a skinning model. *First European Conference on Visual Media Production* (2004): 297-302
- [4] Nirei K., Saito H., Mochimaru M., Ozawa S., Human hand tracking from binocular image sequences. *In 22th International Conference on Industrial Electronics, Control, and Instrumentation*
- [5] Kuch J.J, Huang T.S , Human computer interaction via the human hand: a hand model., *Twenty-Eighty Asilomar Conference on Signal, Systems, and Computers* (1994): 1252–56
- [6] Rehg J., Kanade T., Visual tracking of high DoF articulated structures: An application to human hand tracking. *In European Conference on Computer Vision and Image Understanding* (1994): 35–46
- [7] Ali Erol, George Bebis, Mircea Nicolescu, Richard D. Boyle, Xander Twombly., Vision-based hand pose estimation: A review. *Computer Vision and Image Understanding* 108 (2007), pp 52–73
- [8] Heap A. J., Hogg D. C., Towards 3-D hand tracking using a deformable model. *In 2nd International Face and Gesture Recognition Conference* (1996), pp 140–45
- [9] Stenger B. , Mendonc P. R. S., Cipolla R., Model-Based 3D Tracking of an Articulated Hand." *Proc. British Machine Vision Conference* 1 (2001): 63-72
- [10] Stenger B., Thayananthan A., Torr P.H.S., Cipolla R. Model-based hand tracking using a hierarchical Bayesian filter. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2006)
- [11] Jinshi Cui, Zengqi Sun, Model-based visual hand posture tracking for guiding a dexterous robotic hand, *Optics Communications* 235 (2004) 311–318
- [12] Bay M, Koller-Meier, Gool L.V., Smart particle filtering for 3D hand tracking, in: *Sixth IEEE International Conference on Automatic Face and Gesture Recognition*, Los Alamitos, CA, USA, 2004, pp 675
- [13] Martin de La Gorce, Nikos Paragios, David J. Fleet, Model-Based Hand Tracking with Texture, Shading and Self-occlusions, *IEEE Conference on Computer Vision and Pattern Recognition*, Alaska, 2008
- [14] Martin de La Gorce, David J. Fleet, and Nikos Paragios, Model-Based 3D Hand Pose Estimation from Monocular Video, *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 2011
- [15] Chutisant Kerdvibulvech, Hideo Saito, Model-Based Hand Tracking by Chamfer Distance and Adaptive Color Learning Using Particle Filter *EURASIP Journal on Image and Video Processing* 2009
- [16] New, J. R., Hasanbelliu, E. and Aguilar, M. Facilitating User Interaction with Complex Systems via Hand Gesture Recognition. *In Proc of Southeastern ACM Conf.*, Savannah 2003
- [17] Yang M. H., Ahuja N., and Tabb M., "Extraction of 2-D Motion Trajectories and its Application to Hand Gesture Recognition," in *PAMI*, 29(8) (2002) 1062–1074
- [18] Mo Z., Lewis J.P., Neumann U., Smartcanvas: a gesture-driven intelligent drawing desk system, *In 10th International Conference on Intelligent User Interfaces*, ACM Press (2005): 239-43
- [19] Martin J., Devin V. , Crowley J.L., Active hand tracking, *3rd. International Conference on Face & Gesture Recognition*, IEEE Computer Society (1998): 575
- [20] Kjeldsen R., Kender J., Toward the use of gesture in traditional user interfaces, *International Conference on Automatic Face and Gesture Recognition* (1996): 151–56
- [21] O'Hagan R.G., Zelinsky A., . Rougeaux S. Visual gesture interfaces for virtual environments, *Interacting with Computers* 14 (2002): 231–50
- [22] Lars Bretzner, Ivan Laptev and Tony Lindeberg, Hand gesture recognition using

- multiscale color features, hierarchical models and particle filtering, in *Proceedings of Int. Conf. on Automatic face and Gesture recognition*, Washington D.C., May 2002
- [23] Black M., Jepson D, Eigen tracking: Robust matching and tracking of articulated objects using a view-based representation, *In European Conference on Computer Vision*, 1996.
- [24] C C Wang, K C Wang, Hand Posture recognition using Adaboost with SIFT for human robot interaction, Springer Berlin, ISSN 0170-8643, Volume 370, 2008
- [25] Dardas, N.; Qing Chen; Georganas, N.D.; Petriu, E.M., Hand gesture recognition using Bag-of-features and multi-class Support Vector Machine, *Haptic Audio-Visual Environments and Games (HAVE)*, 2010
- [26] Yuelong Chuang, Ling Chen, Gangqiang Zhao and Gencai Chen, Hand Posture Recognition and Tracking Based on Bag-of-Words for Human Robot Interaction, *IEEE International Conference on Robotics and Automation Shanghai International Conference Center* May 9-13, 2011, Shanghai, China
- [27] Simion G., Gui V., Oteşteanu M. "A Compositional Tehnique for Hand Posture Recognition : New Results." *Wseas Transactions on Communications* 8 (8) (2009): 805-21
- [28] Michael Donoser and Horst Bischof, Real Time Appearance Based Hand Tracking Pattern, *ICPR 2008*
- [29] R. Lienhart and J. Maydt, An extended set of Haar-like features for rapid object detection, in *Proc. IEEE Int. Conf. Image Process.*, 2002 vol. 1, pp. 900–903
- [30] Andre L. C. Barczak, Farhad Dadgostar, Real-time hand tracking using a set of co-operative classifiers based on Haar-like features, *Res. Lett Inf. Math Sci.*, 2005, Vol. 7, pp 29-42
- [31] Qing Chen , N.D. Georganas, E.M Petriu, "Real-time Vision based Hand Gesture Recognition Using Haar-like features *IEEE Transactions on Instrumentation and Measurement*, 2007
- [32] Lowe, David G., Object recognition from local scale-invariant features, *Proceedings of the International Conference on Computer Vision*, 1999, Vol 2, pp. 1150–1157
- [33] H. Bay, A. Ess, T. Tuytelaars and L. Van Gool, SURF: Speeded up robust features, in *CVIU*, 2008, 110(3), pp. 346-359
- [34] Bruce V., Green P. R., Georgeson M. A., *Visual Perception: Physiology, Psychology, and Ecology*, Psychology Press East Sussex, UK, 3rd edition 1996
- [35] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool, A comparison of affine region detectors, *International Journal of Computer Vision*, 65(1-2):43–72, 2005
- [36] Sharma, R., Huang, T. S., Pavovic, V. I., Zhao, Y., Lo, Z., Chu, S., Schulten, K., Dalke, A., Phillips, J., Zeller, M. & Humphrey, W., Speech/Gesture Interface to a Visual Computing Environment for Molecular Biologists, In: *Proc. of ICPR'96*, Vol 2, pp 964-968
- [37] Gandy, M., Starner, T., Auxier, J. & Ashbrook, D. "The Gesture Pendant: A Self Illuminating, Wearable, Infrared Computer Vision System for Home Automation Control and Medical Monitoring". *Proc. of IEEE Int. Symposium on Wearable Computers*. (2000), 87-94
- [38] Goza, S. M., Ambrose, R. O., Diftler, M. A. & Spain, I. M, Telepresence Control of the NASA/DARPA Robonaut on a Mobility Platform, In: *Conference on Human Factors in Computing Systems*. ACM Press, (2004) 623–629
- [39] Malima A, Ozgur E, Cetin M, A fast algorithm for Vision based hand gesture recognition for robot control, *14th IEEE conference on Signal Processing and Communications Applications*, April 2006
- [40] Stotts D., Smith J, Gyllstrom M, K. Facespace: Endoand Exo-Spatial Hypermedia in the Transparent Video Facetop, In: *Proc. of the Fifteenth ACM Conf. on Hypertext & Hypermedia*. ACM Press, (2004) 48–57
- [41] Smith, G. M. & Schraefel. M. C., *The Radial Scroll Tool: Scrolling Support for Stylus-or Touch-Based Document Navigation*, In *Proc. 17th ACM Symposium on User Interface Software and Technology*. ACM Press, (2004) 53–56
- [42] Valli C., Lucas C., *Linguistics of American Sign Language: An Introduction*, Washington, D. C.: Gallaudet University Press, (2000)
- [43] Martinez A., Wilbur, B., Shay R., Kak, A. "Purdue RVL-SLLL ASL Database for Automatic Recognition of ASL, In *IEEE Int. Conf. on Multimodal Interfaces*, (2002) 167–172
- [44] Starner, T., Weaver, J., Pentland, A. , Real-Time American Sign Language Recognition using Desk and Wearable Computer Based Video, *PAMI*, 20(12) (1998) 1371–1375



- [45] Vogler, C. & Metaxas, D. "A Framework for Recognizing the Simultaneous Aspects of American Sign Language" *Comp. Vision and Image Understanding*, 81(3) (2001) 358–384
- [46] Waldron, M. "Isolated ASL Sign Recognition System for Deaf Persons". *IEEE Transactions on Rehabilitation Engineering*, 3(3) (1995) 261–271
- [47] Dong Guo Yonghua, Vision-Based Hand Gesture Recognition for Human-Vehicle Interaction, *International Conference on Control, Automation and Computer Vision*, 1998
- [48] Pickering, Carl A. Burnham, Keith J. Richardson, Michael J. Jaguar, A research Study of Hand Gesture Recognition Technologies and Applications for Human Vehicle Interaction, *3rd Conference on Automotive Electronics*, 2007
- [49] Juan P. Wachs, Helman I. Stern, Yael Edan, Michael Gillam, Jon Handler, Craig Feied, Mark Smith, *A Gesture-based Tool for Sterile Browsing of Radiology Images*, *Journal of the American Medical Informatics Association*, 2008
- [50] <http://www.eyesight-tech.com/news-and-events/#news>