



GLOBAL JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY  
Volume 11 Issue 22 Version 1.0 December 2011  
Type: Double Blind Peer Reviewed International Research Journal  
Publisher: Global Journals Inc. (USA)  
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

# A Computer Vision Based Collaborative Augmented Reality Method for Human-Computer Interaction

By Akhil Khare, Vinaya Kulkarni, Dr. Akhilesh Upadhayay

*Abstract* - Computer vision is becoming very popular now a days since it can hold a lot of information at a very low cost. With this increasing popularity of computer vision there is a rapid development in the field of virtual reality as it provides an easy and efficient virtual interface between human and computer. At the same time much research is going on to provide more natural interface for human-computer interaction with the power of computer vision .the most powerful and natural interface for human-computer interaction is the hand gesture. Hand replaces the currently used cumbersome and inefficient devices like mouse and keyboard and with the bare hands one can easily communicate with the computer. This paper explores a system where hand gesture can be effectively used as a password in the login process for authentication of a person using just a simple web camera. Also this technique does not need any special device like head-mounted display, gloves or any special camera that operates beyond visible spectrum. So with this idea, with a simple video camera and bare hands, a person can interact with computer.

*Keywords* : Computer vision, Human-Computer interaction, Gesture recognition, Haar-like feature.

*GJCST Classification* : 1.2.10



A COMPUTER VISION BASED COLLABORATIVE AUGMENTED REALITY METHOD FOR HUMAN-COMPUTER INTERACTION

*Strictly as per the compliance and regulations of:*



# A Computer Vision Based Collaborative Augmented Reality Method for Human-Computer Interaction

Akhil Khare <sup>α</sup>, Vinaya Kulkarni <sup>Ω</sup>, Dr. Akhilesh Upadhayay <sup>β</sup>

**Abstract** - Computer vision is becoming very popular now a days since it can hold a lot of information at a very low cost. With this increasing popularity of computer vision there is a rapid development in the field of virtual reality as it provides an easy and efficient virtual interface between human and computer. At the same time much research is going on to provide more natural interface for human-computer interaction with the power of computer vision .the most powerful and natural interface for human-computer interaction is the hand gesture. Hand replaces the currently used cumbersome and inefficient devices like mouse and keyboard and with the bare hands one can easily communicate with the computer. This paper explores a system where hand gesture can be effectively used as a password in the login process for authentication of a person using just a simple web camera. Also this technique does not need any special device like head-mounted display, gloves or any special camera that operates beyond visible spectrum. So with this idea, with a simple video camera and bare hands, a person can interact with computer.

**Keywords** : *Computer vision, Human-Computer interaction, Gesture recognition, Haar-like feature*

## 1. INTRODUCTION

With the rapid increase in human and computer interaction an easy and natural interface is getting much more value than it was previously. Now a day's keyboard and mouse are used as the main interface for transferring information and commands to the computer. In our day to day life we human uses our vision and hearing as a main source of information about our environment. Therefore a much research is going on for providing more natural interface for human-computer interaction based on computer vision. Hand gesture is most popular and effective medium for communication in virtual environment because it conveys information very effectively and naturally. The purpose of this project is to develop new perceptual interfaces for human-computer-interaction based on visual input captured by computer vision systems. In the initial days different cumbersome devices were imposed on users such as head mounted display, digital gloves etc. These devices had limited the users movement and feels uncomfortable to the user. On the other hand vision based gesture recognition system that uses bare hand is becoming very popular because it does not need any device to impose on user's body Instead, it provides a natural hand gesture recognition interface system for human-computer interaction. The whole

process of hand gesture recognition is broadly divided in three steps first is the segmentation that is the hand is separated from the background using different methods such as colour segmentation method. Then the features of the hand are extracted that is the feature detection and with the help of extracted features multiple hand gestures are categorized in to three groups communicative, manipulative and controlling gesture. Communicative gesture is used to express an idea or concept. Manipulative gesture is used to interact with virtual objects in virtual environment. To control a system controlling gestures are used.

Previous methods suffers from the limitation of lightening changes and less accuracy. Also in some methods different devices were used such as head mounted display or hand gloves etc. In some methods two cameras were used as a well as sometimes a 3D sensor was also used. So a new method has been invented for gesture recognition that uses haar like structure along with topological features and color segmentation technique to identify and classify different hand postures and gives satisfactory performance with higher accuracy when applied to human-computer interaction for personal authentication. This method makes use of a single camera to capture the image as well as no special device or sensor is needed.

In this paper we focus our attention to vision based recognition of hand gesture for personal authentication where hand gesture is used as a password. Different hand gestures are used as password for different personals. This method could also be used for blind people who can use their hand gesture as a password for the login process. Hand gesture has been used mostly to convey some commands to the computer. This system is introduces a new application of hand gesture that is the personal authentication.

The remainder of this paper is structured as follows: section II takes a short review on different methods described in various papers. The hand gesture classification and phases are discussed in section III. Section IV discusses proposed system and also discusses how it is different from the existing systems and finally the conclusion.

## II. LITERATURE ANALYSIS

One of the method proposed by Rokade et al [1] uses the technique of thinning of segmented image, but it needs more computation time to detect different hand postures. One method is based on elastic graph matching, but it is sensitive to light changes [2]. In a system proposed by Stergiopoulou and Papamarkos YCbCr color segmentation model was used but the background should be plane and uniform [3].

In one method CSS features were used by Chin-Chen Chang for hand posture recognition [4]. In the method presented by this paper a boost cascade of classifiers trained by Adaboost and haar like features are used to accelerate the evaluation speed used to recognize two hand postures for human-robot recognition. It uses haar like features along with color segmentation method to improve the accuracy in detecting the hand region and then the topological method is used to classify different hand postures.

In the method proposed by Shuying Zhao [5] for hand segmentation Gaussian distribution model (for building complexion model) is used. With Fourier descriptor and BP neural network an improved algorithm is developed that has good describing ability and good self learning ability. This method is flexible and realistic. In the system proposed by Wei Du and Hua Li statistic based and contour based features are used for stable hand detection [6].

In a system developed by Utsumi [7] a simple hand model is constructed from reliable image features. This system uses four cameras for gesture recognition. In a system known as fingermouse developed by Quek the hand gesture replaces mouse for certain actions [8]. In this system only one gesture that is pointing gesture is defined and for mouse press button shift key on the keyboard is used. Segan has developed a system [9] that uses two cameras to recognize three gestures and hand tracking in 3D. By extracting the feature points on hand contour the thumb finger and pointing finger are detected by the system.

In the system presented by Triesch multiple dcues such as motion cue stereo cue, color cue are used for robust gesture recognition algorithm[10]. This system is used in human robot interaction that helps robot to grasp objects kept on the table. In the system real time multihand posture recognition system for human-robot interaction haar like feature and topological features were used along with color segmentation technique [11]. This method gives accurate results and a rich set of features could be extracted.

Compared with the traditional interaction approaches, such as keyboard, mouse, pen, etc, vision based hand interaction is more natural and efficient. Yikai Fang, Kongqiao Wang, Jian Cheng and Hanqing Lu proposed a robust real time hand gesture recognition

method [12] in their paper "A real time hand gesture recognition method". In this method, firstly, a specific gesture is required to trigger the hand detection followed by tracking; then hand is segmented using motion and color cues; finally, in order to break the limitation of aspect ratio encountered in most of learning based hand gesture method, the scale-space feature detection is integrated into gesture recognition. Applying the proposed method to navigation of image browsing, experimental results show that this method achieves satisfactory performance.

Wei Du and Hua Li presented a real-time system in "Vision based gesture recognition system with single camera" for human-computer interaction through gesture recognition and hand tracking[10]. Stable detection can be achieved by extracting two kinds of features: statistic-based feature and contour-based feature. Unlike most of previous works, our system recognizes hand gesture with just one camera, thus avoids the problem of matching image features between different views. This system can serve as a natural and convenient user input device, replacing mouse and trackball.

## III. GESTURE CLASSIFICATION AND MODELLING

Hand gesture is a movement of hand(s) and arm(s) that are used as a means to express an idea or to convey a command to control an action. Hand gesture can be classified in a several ways. For HCI applications the most commonly used and suitable classification divides hand gesture in to three groups communicative, Manipulative and controlling gestures. To express an idea or a specific concept communicative gestures are used. It may be used as a substitute for verbal communication. Communicative gesture is similar to sign language and as in sign language it also requires a high structured set of gestures. To interact with objects in an environment manipulative gestures are used. This is generally used to manipulate virtual objects in virtual environment. Controlling gestures as the name indicates used to control a system or to locate an object. One of the application of controlling gestures is controlling mouse movements on the desktop. The major steps in hand gesture analysis are analysis of hand motion, modeling of hand(s) and arm(s), mapping the motion features to the model and interpreting the gesture in the time interval. But generally speaking analysis of hand gesture is totally application dependent.

From the psychological point of view hand gesture consists of three phases, these are preparation,, nucleus and retraction phase. Preparatory phase includes bringing the hand from resting position to the starting posture of the gesture. Sometimes this phase is very short and many times it is combined with the

retraction phase of the previous gesture. The next is nucleus phase that includes the main concept of gesture and has a definite form and that is used as a command to the computer. The last phase is retraction that shows the resting position of hand after completing the gesture. If the gesture is succeeded by another gesture then the retraction phase may be very short or not present. The preparatory and retraction phases are usually and hand movements are faster compared to the nucleus phase. However identifying starting and ending position of the nucleus phase is quite difficult as there are variations in the preparatory and retraction phase.

#### IV. PROPOSED SYSTEM

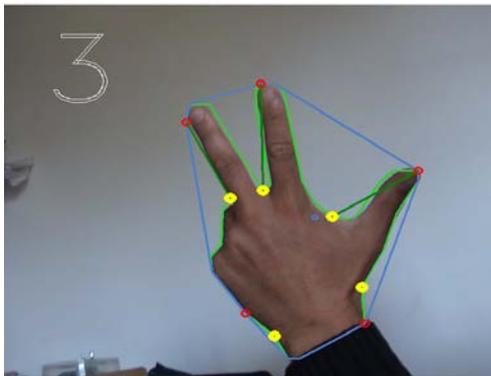


Fig 1(a)



Fig 1(b)

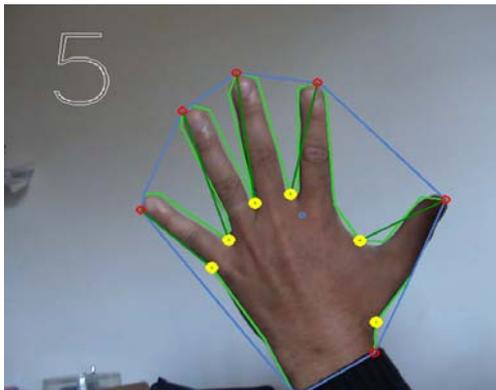


Fig 2(a)



Fig 2(b)

Fig 1 : (a) and 2(a) original picture and fig 1(b) and 2(b) segmentation result

##### a) Video Capturing

Video capture is the process of converting an analog video signal—such as that produced by a video camera—to digital form. The resulting digital data are referred to as a digital video stream, or more often, simply video stream. This is in contrast with screen casting, in which previously digitized video is captured while displayed on a digital monitor.

The video capture process involves several processing steps. First the analog video signal is digitized by an analog-to-digital converter to produce a raw, digital data stream. In the case of composite video, the luminance and chrominance are then separated; this is not necessary for S-Video sources. Next, the chrominance is demodulated to produce color difference video data. At this point, the data may be modified so as to adjust brightness, contrast, saturation and hue. Finally, the data is transformed by a color space converter to generate data in conformance with any of several color space standards, such as RGB and YCbCr. Together, these steps constituted video decoding, because they "decode" an analog video format such as NTSC or PAL. Special electronic circuitry is required to capture video from analog video sources. At the system level this function is typically performed by a dedicated video capture card. Such cards often utilize video decoder integrated circuits to implement the video decoding process.

##### b) Image extraction from video

Here we have to select captured video as input. We are now ready to start extracting frames from the videos. After getting frame from video start to extract images from those frames. Store those extracted files in particular folder.

##### c) Image enhancement and Remove noise

Noise reduction is the process of removing noise from a signal. Noise reduction techniques are conceptually very similar regardless of the signal being processed, however a priori knowledge of the characteristics of an expected signal can mean the



implementations of these techniques vary greatly depending on the type of signal. The median filter is a nonlinear digital filtering technique, often used to remove noise. Such noise reduction is a typical pre-processing step to improve the results of later processing (for example, edge detection on an image). Median filtering is very widely used in digital image processing because under certain conditions, it preserves edges while removing noise. The main idea of the median filter is to run through the signal entry by entry, replacing each entry with the median of neighboring entries. The pattern of neighbors is called the "window", which slides, entry by entry, over the entire signal. For 1D signal, the most obvious window is just the first few preceding and following entries, whereas for 2D (or higher-dimensional) signals such as images, more complex window patterns are possible (such as "box" or "cross" patterns). Note that if the window has an odd number of entries, then the median is simple to define: it is just the middle value after all the entries in the window are sorted numerically.

#### d) Background suppress

An algorithm that detects and removes background shadows from images in which the pattern set occupies the upper-most intensity range of the image and the image is background dominant outside the pattern set is presented. The algorithm will remove background shadows and preserve any remaining texture left behind by the shadow function. A mathematical model of the histogram modification function of the shadow-removal algorithm is developed. An analysis of the sequential nature of the algorithm is included along with simulated results to verify the mathematical model developed and to show the effectiveness of the algorithm in removing background pattern shadows.

#### e) Hand region segmentation

The initial step of hand gesture recognition is the detection of hand region from the background. This step is also known as hand detection. It involves detecting and extracting hand region from background and segmentation of hand image. Previous methods made use of following two approaches that is the color based model and statistical based model. This system uses the additional third approach i.e. haar like feature with adaboost technology. Different features such as skin colour, shape, motion and anatomical models of hand are used in different methods. The output of this step is a binary image in which skin pixels have value 1 and non-skin pixels have value 0.

**Haar-like detector:** First step is conversion of input image to an integral image since haar-like features can be calculated from an integral image with a greater speed. A rich set of haar like features can be computed from the integral image. The integral image at the point  $p(x,y)$  contains the sum of the pixel values left and above this point. It is defined as,

$$P(x,y) = \sum I(X,Y)$$

$$x' \leq x, y' \leq y$$

each haar like feature is composed of two connected black and white rectangles. The value of a haar likefeature is obtained by subtracting the sum pixel values of the white rectangle from the black rectangle. Single haar like feature is not able to recognize hand region with high accuracy. The adaboost learning algorithm can considerably improve the overall accuracy stage by stage by using a linear combination of these individually weak classifiers. This combination makes the processing faster and robust.

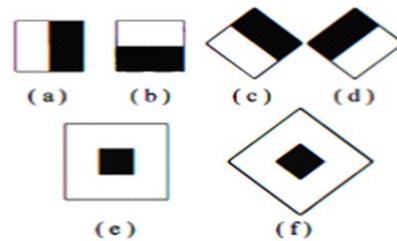


Fig.3 : Haar-like features

Different colour models can be used for hand detection such as **YCbCr**, **RGB**, **YUV** etc. but the proposed system uses **YCbCr** color model since it is useful in compression applications. Also **YCbCr** separates **RGB** in to luminance & chrominance information. Following equation is used to transform **RGB** values in to **YCbCr** color space'. The characteristics of hand shape such as topological features could be used for hand detection. Learning detectors from pixel values: Hands can be found from their appearance and structure such as Adaboost algorithm. 3D model based detection: Using multiple 3D hand models multiple hand postures can be estimated.

#### f) Feature extraction and gesture recognition

The next important step is hand tracking and feature extraction. Tracking means finding frame to frame correspondence of the segmented hand image to understand the hand movement. Following are some of the techniques for hand tracking.

- Template based tracking: If images are acquired frequently enough hand can be tracked. It uses correlation based template matching. by comparing and correlating hand in different pictures it could be tracked.
- Optimal estimation technique: Hands are tracked from multiple cameras to obtain a 3D hand image.
- Tracking based on mean shift algorithm: To characterize the object of interest it uses color distribution and spatial gradient. Mean shift algorithm is used to track skin color area of human hand.

Two types of features are there first one is global statistical features such as centre of gravity and second one is contour based feature that is local feature that includes fingertips and finger-roots. Both of these features are used to increase the robustness of the system. Hand posture can be distinguished using the number of fingers of the hand and if the number of fingers are same then the angle between two fingers can be measured to recognize the specific gesture.

The goal of hand gesture recognition is interpretation of the meaning gesture of the hands location and posture conveys. From the extracted features multiple hand gestures are recognized. Different methods for hand gesture recognition can be used such as template matching, method based on principle component analysis, Boosting contour and silhouette matching, model based recognition methods, Hidden Markov Model (HMM). Hand gesture is movement of hands and arms used to express an idea or to convey some message or to instruct for an action. From psychological point of view hand gesture has three phases.

#### *g) Register user*

The Register User action registers the user information with the installer to identify the user of a product. it provides a unique user id for every user. a large set of postures and gestures is stored on the computer one for each individual.

#### *h) Login*

When a user wants to login he/she has to perform the desired hand gesture. This hand gesture can be performed using single hand. That gesture will be compared with the already recorded gesture that works as a password for that particular person, if that gesture matches with the performed gesture then only that person will be authenticated and will be allowed to access his/her account or product. Basic idea is that the number of fingers are counted and the password is created Ex, 123,432,531,23,4532,123451 etc. the password can be any combination of the numbers from 0,1,2,3,4,5. This password performed by the user is authenticated by the system and he/she will be allowed to access the application or is rejected the access.

This proposed system could be used by any application to authenticate the authorized user. The major benefit of this system is that it could be used by blind users also, but the accuracy is the major concern, the system may not give accurate results in intricate background.

## V. CONCLUSION

Vision based hand gesture recognition has major applications in human-computer interaction as well as in intelligent service robot. This paper describes a collaborative vision based hand gesture recognition system where a hand gesture could be effectively used by a person as her password for the personal

authentication. This system provides an easy interface for human-computer interaction. This system will provide a more efficient system with greater accuracy that makes use of both the hands as well as the drawback of previous techniques have been tried to remove such as complexion problem could be effectively removed by using background model alongwith the complexion model. In multihand gesture recognition method a rich set of features could be extracted using haar like feature and topological features with greater accuracy.

## REFERENCES REFERENCES REFERENCIAS

1. R. Rokade, D. Doye, and M. Kokare, "Hand Gesture Recognition by Thinning Method," Digital Image Processing, 2009 IEEE International Conference on, 2009, pp. 284-287.
2. Stergiopoulou, N. Papamarkos, "Hand gesture recognition using a neural network shape fitting technique," Engineering Applications of Artificial Intelligence, Vol. 22, Issue 8, December 2009, pp. 1141-1158.
3. C.C. Chang, "Adaptive multiple sets of css features for hand posture recognition," Neurocomputing, vol. 69,2006, pp.2017-2025.
4. Shuying Zhao, Wenjun Tan, Chengdong Wu, Chunjiang Liu, Shiguang Wen, "A Novel Interactive Method of Virtual Reality System Based on Hand Gesture recognition" 2009 Chinese Control and Decision Conference (CCDC 2009).
5. Wei Du Hua Li, "Vision based gesture recognition system with single camera", proceedings of ICSP2000, 2000 IEEE.
6. Akira Utsumi, Tsutomu Miyasato, Fumio Kishino, Ryohei Nakatsu, Hand Gesture Recognition System using Multiple Cameras, Proceedings of ICPR'96, vol 1~6 6 7 - 6 7A1u, wst 1996.
7. Francis K.H. Quek, Non-Verbal Vision-Based Interfaces, Vision Interfaces and Systems Laboratory Technical Report 1994, Electrical Engineering and Computer Science Department, The University of Illinois at Chicago.
8. Jakub Segen and Senthil Kumar, Human-Computer Interaction using gesture Recognition and 3D Hand Tracking, Proceedings of ICIP'98, Vol. 3, pp188-192, Chicago, October 1998.
9. Jochen Triesch, Christoph von der Malsburg, Robotic Gesture Recognition, Proceedings of ECCV'98, pp 2 3 3 -2 2 4 1998.
10. Cao Chuqing, Li Ruifeng, Ge Lianzheng "Real time multihand posture recognition" , International Conference On Computer Design And Applications (ICDDA 2010)
11. Yikai Fang, Kongqiao Wang, Jian Cheng, Hanqing Lu, "A Real-Time Hand Gesture Recognition Method," Multimedia and Expo, 2007 IEEE International Conference on, July 2007, pp.995-998.