

February 15, 2024

# Hand Gesture Recognition: Solving a Computer Vision Problem

Tasnimul Hasnat  
190041113  
CSE-1

# Contents

1 Introduction .....	3
2 Solution Approaches .....	3
2.1 Glove Based Sensor Approach .....	3
2.2 Camera-Vision Based Approach .....	4
3 Recognition Methods .....	5
3.1 Color-Based Recognition .....	5
3.1.1 Color-Based Recognition Using Glove Marker .....	5
3.1.2 Color-Based Recognition of Skin Color .....	5
3.2 Appearance-Based Recognition .....	6
3.3 Motion-Based Recognition .....	6
3.4 3D Model-Based Recognition .....	6
3.5 Deep-Learning Based Recognition .....	7
4 Conclusion .....	7
References .....	7

# 1 Introduction

Hand gestures are an aspect of body language that can be conveyed through the center of the palm, the finger position and the shape constructed by the hand. Hand gestures can be classified into static and dynamic. As its name implies, the static gesture refers to the stable shape of the hand, whereas the dynamic gesture comprises a series of hand movements such as waving. There are a variety of hand movements within a gesture; for example, a handshake varies from one person to another and changes according to time and place. The main difference between posture and gesture is that posture focuses more on the shape of the hand whereas gesture focuses on the hand movement. Hand gestures offer an inspiring field of research because they can facilitate communication and provide a natural means of interaction that can be used across a variety of applications. The main approaches to hand gesture research can be classified into the wearable glove-based sensor approach and the camera vision-based sensor approach.[1, 2]

## 2 Solution Approaches

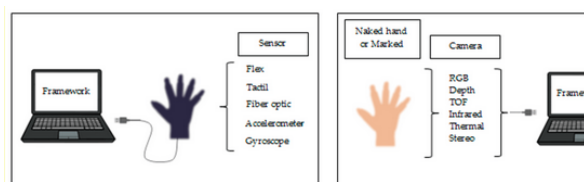


Figure 1: Different techniques for hand gestures. (a) Glove-based attached sensor either connected to the computer or portable; (b) computer vision-based camera using a marked glove or just a naked hand.

As illustrated in Figure 1, hand gestures for human–computer interaction (HCI) started with the invention of the data glove sensor. It offered simple commands for a computer interface. The gloves used different sensor types to capture hand motion and position by detecting the correct coordinates of the location of the palm and fingers. Various sensors using the same technique based on the angle of bending were the curvature sensor, angular displacement sensor, optical fiber transducer, flex sensors and accelerometer sensor. These sensors exploit different physical principles according to their type.

### 2.1 Glove Based Sensor Approach

In the glove base approach, the sensors detected a physical response according to hand movements or finger bending. The data collected were then processed using a computer connected to the glove with wire. This system of glove-based sensor could be made portable by using a sensor attached to a microcontroller.

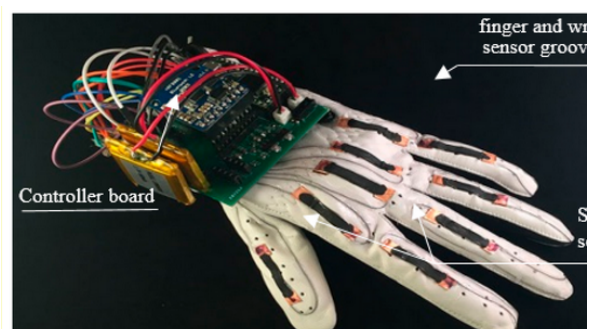


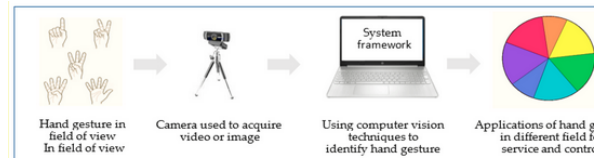
Figure 2: Sensor-based data glove

Although the techniques mentioned above have provided good outcomes, they have various limitations that make them unsuitable for the elderly, who may experience discomfort and confusion

due to wire connection problems. In addition, elderly people suffering from chronic disease conditions that result in loss of muscle function may be unable to wear and take off gloves, causing them discomfort and constraining them if used for long periods. These sensors may also cause skin damage, infection or adverse reactions in people with sensitive skin or those suffering burns. Moreover, some sensors are quite expensive. These drawbacks led to the development of promising and cost-effective techniques that did not require cumbersome gloves to be worn. These techniques are called camera vision-based sensor technologies.

## 2.2 Camera-Vision Based Approach

With the evolution of open-source software libraries, it is easier than ever to detect hand gestures that can be used under a wide range of applications like clinical operations [3], sign language [4], robot control [5], virtual environments, home automation [6], personal computer and tablet, Using computer vision techniques to identify gestures. Where the user perform specific gesture by single or both hand in front of camera which connect with system framework that involve different possible techniques to extract feature and classify hand gesture to be able control some possible application. gaming [7].



*Figure 3: Using computer vision techniques to identify gestures. Where the user perform specific gesture by single or both hand in front of camera which connect with system framework that involve different possible techniques to extract feature and classify hand gesture to be able control some possible application.*

These techniques essentially involve replacement of the instrumented glove with a camera. Different types of camera are used for this purpose, such as RGB camera, time of flight (TOF) camera, thermal cameras or night vision cameras. Algorithms have been developed based on computer vision methods to detect hands using these different types of cameras. The algorithms attempt to segment and detect hand features such as skin color, appearance, motion, skeleton, depth, 3D model, deep learn detection and more.

In this paper we will discuss the various computer vision based approaches, their benefits and the challenges they might face.

## 3 Recognition Methods

### 3.1 Color-Based Recognition

#### 3.1.1 Color-Based Recognition Using Glove Marker

This method uses a camera to track the movement of the hand using a glove with different color marks, as shown in Figure 3. This method has been used for interaction with 3D models, permitting some processing, such as zooming, moving, drawing and writing using a virtual keyboard with good flexibility [4]. The colors on the glove enable the camera sensor to track and detect the location of the palm and fingers, which allows for the extraction of geometric model of the shape of the hand. The advantages of this method are its simplicity of use and low price compared with the sensor data glove. However, it still requires the wearing of colored gloves and limits the degree of natural and spontaneous interaction with the HCI. The color-based glove marker is shown in Figure 4.



Figure 4: Color-based recognition using glove marker

#### 3.1.2 Color-Based Recognition of Skin Color

Skin color detection is one of the most popular methods for hand segmentation and is used in a wide range of applications, such as object classification, degraded photograph recovery, person movement tracking, video observation, HCI applications, facial recognition, hand segmentation and gesture identification. Skin color detection has been achieved using two methods. The first method is pixel based skin detection, in which each pixel in an image is classified into skin or not, individually from its neighbor. The second method is region skin detection, in which the skin pixels are spatially processed based on information such as intensity and texture. Color space can be used as a mathematical model to represent image color information. Several color spaces can be used according to the application type such as digital graphics, image process applications, TV transmission and application of computer vision techniques.

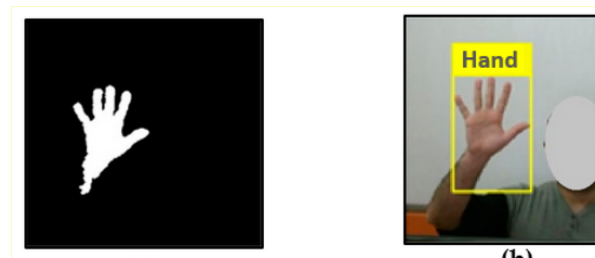


Figure 5: Example of skin color detection. (a) Apply threshold to the channels of YUV color space in order to extract only skin color then assign 1 value for the skin and 0 to non-skin color; (b) detected and tracked hand using resulted binary image.

### 3.2 Appearance-Based Recognition

This method depends on extracting the image features in order to model visual appearance such as hand and comparing these parameters with feature extracted from the input image frames. Where the features are directly calculated by the pixel intensities without a previous segmentation process. The method is executed in real time due to the easy 2D image features extracted and is considered easier to implement than the 3D model method. In addition, this method can detect various skin tones. Utilizing the AdaBoost learning algorithm, which maintains fixed feature such as key points for a portion of a hand, which can solve the occlusion issue, it can separate into two models: a motion model and a 2D static model.

A study by Chen et al. [6] proposed two approaches for hand recognition. The first approach focused on posture recognition using Haar-like features, which can describe the hand posture pattern effectively used the AdaBoost learning algorithm to speed up the performance and thus rate of classification. The second approach focused on gesture recognition using context-free grammar to analyze the syntactic structure based on the detected postures.

### 3.3 Motion-Based Recognition

Motion-based recognition can be utilized for detection purposes; it can be extracts the object through a series of image frames. The AdaBoost algorithm utilized for object detection,

characterization, movement modeling, and pattern recognition is needed to recognize the gesture. The main issue encounter motion recognition is this is an occasion if one more gesture is active at the recognition process and also dynamic background has a negative effect. In addition, the loss of gesture may be caused by occlusion among tracked hand gesture or error in region extraction from tracked gesture and effect long-distance on the region appearance. Two stages for efficient hand detection were proposed in. First, the hand detected for each frame and center point is used for tracking the hand. Then, the second stage matching model applying to each type of gesture using a set of features is extracted from the motion tracking in order to provide better classification where the main drawback of the skin color is affected by lighting variations which lead to detect non-skin color. A standard face detection algorithm and optical flow computation was used by to give a user-centric coordinate frame in which motion features were used to recognize gestures for classification purposes using the multiclass boosting algorithm.

### 3.4 3D Model-Based Recognition

he 3D model essentially depends on 3D Kinematic hand model which has a large degree of freedom, where hand parameter estimation obtained by comparing the input image with the two-dimensional appearance projected by three-dimensional hand model. In addition, the 3D model introduces

human hand feature as pose estimation by forming volumetric or skeletal or 3D model that identical to the user's hand. Where the 3D model parameter updated through the matching process. Where the depth parameter is added to the model to increase accuracy.

### 3.5 Deep-Learning Based Recognition

The artificial intelligence offers a good and reliable technique used in a wide range of modern applications because of using a learning role principle. The deep learning used multilayers for learning data and gives a good prediction out result. The most challenges facing this technique is required dataset to learn algorithm which may affect time processing.

## 4 Conclusion

Hand gesture recognition addresses a fault in interaction systems. Controlling things by hand is more natural, easier, more flexible and cheaper, and there is no need to fix problems caused by hardware devices, since none is required. From previous sections, it was clear to need to put much effort into developing reliable and robust algorithms with the help of using a camera sensor has a certain characteristic to encounter common issues and achieve a reliable result. Each technique mentioned above, however, has its advantages and disadvantages and may perform well in some challenges while being inferior in others.

## References

1. Fang, Z.: Computer gesture input and its application in human computer interaction. MINIMICRO SYSTEMS-SHENYANG-. 20, 418–421 (1999)
2. Mitra, S., Acharya, T.: Gesture recognition: A survey. IEEE Transactions on Systems, Man, and Cybernetics. 37, 311–324 (2007)
3. Kramer, R. K., Majidi, C., Sahai, R., Wood, R. J.: Soft curvature sensors for joint angle proprioception. In: 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 1919–1926 (2011)
4. Pansare, J. R., Gawande, S. H., Ingle, M.: Real-time static hand gesture recognition for American Sign Language (ASL) in complex background. (2012)
5. Bergh, M. Van den, Carton, D., De Nijs, R., Mitsou, N., Landsiedel, C., Kuehnlenz, K., Wollherr, D., Van Gool, L., Buss, M.: Real-time 3D hand gesture interaction with a robot for understanding directions from humans. In: 2011 Ro-Man. pp. 357–362 (2011)
6. Desai, S., Desai, A.: Human Computer Interaction through hand gestures for home automation using Microsoft Kinect. In: ComNet 2016. pp. 19–29 (2017)
7. Murthy, G., Jadon, R.: A review of vision based hand gestures recognition. International Journal of Information Technology and Knowledge Management. 2, 405–410 (2009)