ISSN: 2277-5528
Impact Factor: 2.745 (SIJF)

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & MANAGEMENT

# REPLACEMENT OF OBJECTS IN A VIDEO FRAME USING HAND GESTURE RECOGNITION

**Shaikh Shabnam *, Dr Shah Aqueel**
* Department of Computer Engineering, Shri.Jagdish Prasad Jhabarmal Tibrewala University, Rajasthan, India#Principal, Maulana Mukhtar Ahmad Nadvi Technical Campus, Malegaon

## ABSTRACT

The main idea behind this paper is to make human computer interaction a touch less technology. This touch less technology is applied to video inpainting technique, extracting and replacing the object in frame of video in video processing. For easy communication between human and computer some in-put output devices are designed over the years. Extensions are done in these input output devices to make the computer more intelligent and allow human to do more difficult communication with the computer. Gesture recognition is one of the extension for input output devices.Using the concept of gesture recognition, we can move finger at computer screen so that the cursor will move accordingly.

The aim of this paper is to design an application that is able to do foreground background segmentation of the object in video using real time hand gesture. Hand gestures are detected and tracked in real time to use their actions as mouse events.The user performs a gesture in front of a camera, which is linked to the computer. The picture of the gesture is then processed to identify the gesture indicated by the user. Once the gesture is identified corresponding control action assigned to the gesture is actuated.
The goal and objective is to extract and replace the objects by the other objects at the foreground background of the frame in video processing. All this is done through the dynamic hand gestures.

**Keywords**: gesture recognition, foreground .background, dynamic hand gestures, video processing.

## INTRODUCTION

Several approaches have been studied for solving a problem of recognition of hand postures/gestures. Pattern recognition has been utilized for this purpose. One particular way is decision theoretic approach. There are several methods such as distance classifiers, template matching, conditional random field model (CRF), dynamic time warping model (DTW), Bayesian network, Fishers linear discriminant model, time-delayed neural networks (TDNN), fuzzy neural networks, discriminant analysis. Some studies have followed hybrid models.
Some examples are: k-nearest neighbor combined with Bayesian classifier; least squares estimator with ANFIS network; incorporation of Markov chains and independent component analysis; hybrid statistical classifiers; use of self-organizing feature maps, simple recurrent network with hidden Markov model.
Asanterabi Malima *et. al.* [2] approach to the hand gesture recognition problem a robot control context involved the use of markers on the finger tips. An associated algorithm is used to detect the presence and color of the markers, through which one can identify which fingers are active in the gesture. The inconvenience of placing markers on the user's hand makes this an infeasible approach in practice. Vafadar and Alireza Behrad [3] used Template Based**:** In this approach the data obtained is compared against some reference data and using the thresholds, the data is categorized into one of the gestures available in the reference data. This is a simple approach with little calibration but suffers from noise and doesn't work with overlapping gestures. Byung-Woo Min *et. al.* [4] developed Hidden Markov Model (HMM) which is commonly used and has been widely exploited for temporal gesture recognition. An HMM consists of states and state transitions with observation probabilities. For watch gesture a separate HMM is trained and the recognition of the gesture is based on the generation of maximum probability by a particular HMM. This method also suffers from training time involved and complex working nature as the results are unpredicted because of the hidden nature. For the gesture recognition,
Wing Kwong Chung *et.al.*[5] has presented a hand gesture recognition modal based on "A Real-time Hand Gesture Recognition based on Haar Wavelet Representation.In addition to voice and controller pads, hand gestures can also be an effective way of communication between humans and robots or even between auditory handicapped people and robots.
Mu-Chun Su [6] suggested a method using Neural Network which is based on modeling of the human nervous system element called neuron and its interaction with the other neurons to transfer the information. Each node

consists of and the input function which computes the weighted sum and the activation function to generate the response based on the weighted sum.

Byung-Woo Min *et al*.[7] developed a method for gesture recognition using Hidden Markov Model. This method has been widely exploited for temporal gesture recognition. Bhuyan, *et.al.* [8] have proposed the advantage of VOP based method for segmentation of hand image. The proposed acceleration feature works efficiently only when the spatial end position of preceding gesture is different from start position of next gesture in the connected gesture sequence.

Shewta and Pankaj [9] have proposed that ANN provides a good and powerful solution for gesture recognition. Artificial Neural Networks are applicable to multivariate non-linear problems .It has fast computational ability. Gesture recognition is an important for developing alternative human-computer interaction modalities.

Zhou Ren *et.al.* [10] have worked in the direction of hand gesture recognition by making use of kinetic sensor which is very much different from the normal web camera. Hand gesture based Human-Computer-Interaction (HCI) is one of the most natural and intuitive ways to communicate between people and machines, since it closely mimics how human interact with each other. Hamid A Jalab [11] have proposed the succeeds to extract features from hand gesture image based on hand segmentation using both wavelet network an ANN.

Qing and Nicolas [12] proposed a method that used the formal grammar to represent the hand gestures and postures however limited. This method involves simple gestures requiring the fingers to be extended in various configurations which are mapped to the formal grammar specified by specific tokens and rules. The system involves tracker and glove. This system has poor accuracy and very limited gesture set.

Lee and Yangsheng Xu [13] developed a glove-based gesture recognition system that was able to recognize 14 of the letters from the hand alphabet, learn new gestures and able to update the model of each gesture in the system in online mode, with a rate of 10Hz. Over the years advanced glove devices have been designed such as the Sayre Glove, Dexterous Hand Master and Power Glove

Spatio-temporal vector Analysis method was proposed by Vafadar and Behrad [14] which used to track the movement of the hand in the images of the scene and track the motion in the sequence of image. The information about the motion is obtained by the derivatives and it is assumed that under static background, hand motion is the fastest changing object of the scene. Then using the refinement and variance constraint flow field is refined. This flow field captures the characteristics of the given gesture

## SYSTEM DESCRIPTION

Once the hand gesture events are detected we can select input video for doing operations on it. Suppose we have taken input as video then frames are extracted from input video.

ISSN: 2277-5528
Impact Factor: 2.745 (SIJF)

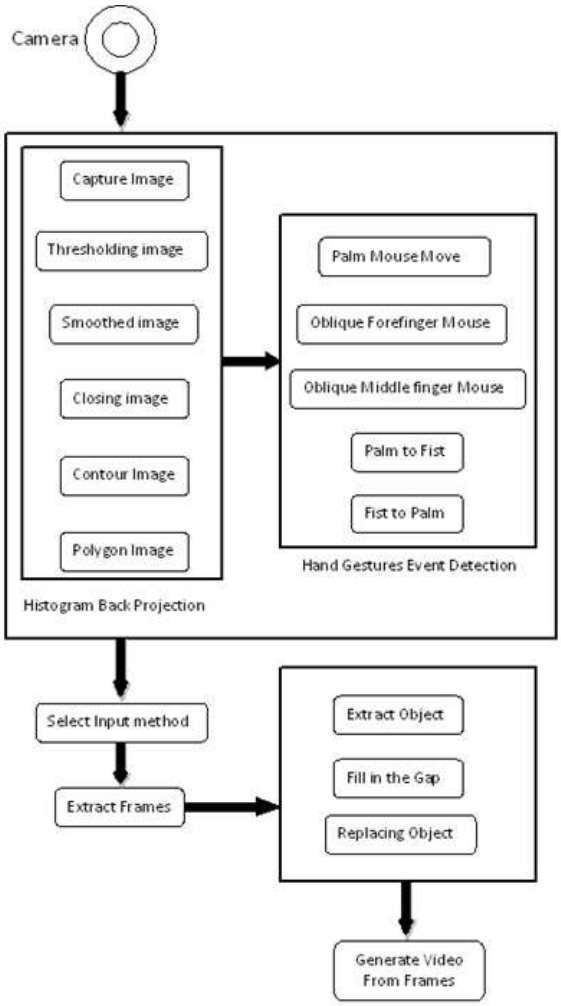


***Figure1: Architectural Components of the system***

PixMix algorithm is applied on those extracted frames. This algorithm takes the selected frame and does the operations. To have a continuity in the video we take the before and after frame corresponding to the current frame. Then the first step in the algorithm is to remove the object. Then the hole or gap is filled with the other object, putting back in video, is done with this algorithm. Depending on these operations algorithm generate video from re-covered frames.

## DATA FLOW DIAGRAM

The graphical representation which shows the information flow, transformations applied as the data moves from input to the output stage is known as data flow diagram (DFD). A system or the software at different level of abstraction can be represented by the data flow diagram. The data flow diagram can be divided into different levels so as to represent the functional details and the flow of the information. The mechanism provided by the DFD is functional modeling and modeling the information flow.

The zero level DFD also known as context model or the fundamental system. It can be used to represent the complete software as one bubble. The input and the output is shown by the arrows. Extra processes (bubbles) and information flow paths can be shown as the level 0 DFD is divided further to expose more detail.

ISSN: 2277-5528
Impact Factor: 2.745 (SIJF)

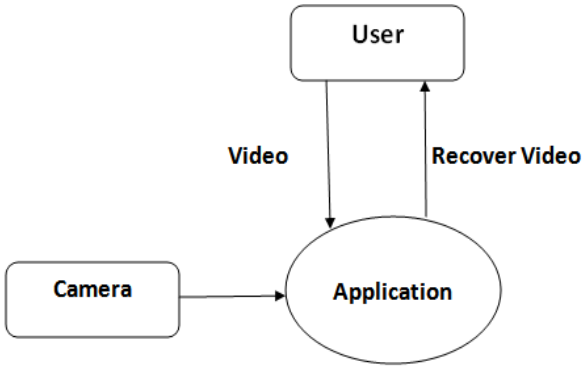


*Figure 2: Data Flow Diagram for Level 0*

In DFD level 0, user gives a input video which is stored on the disk to application with the help of webcam. Application performs required operations on the video by using hand gestures. Application then gives back the recovered video to user.

In DFD level 1, hand gestures are captured through webcam. Processing is done on hand gestures to detect events of hand, for this through histogram back projection method and hand gesture event detection methods are used. With the help of these events video is selected which is degraded. Frames from that video is extracted and then Pix-Mix algorithm is applied on that frames for video Inpainting methods such as extracting object, replacing object etc. Then final recovered video is generated.

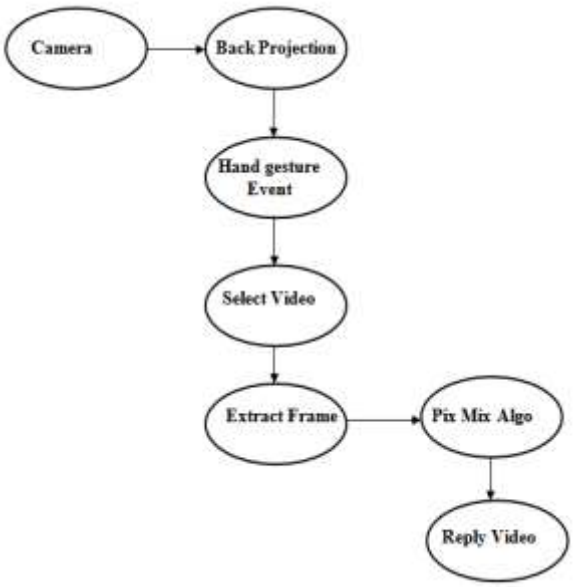


*Figure 3: Data Flow Diagram for Level 1*

## HISTOGRAM BACK PROJECTION ALGORITHM

The back projection is the re-application of the modified histogram to the original image, functioning as a look-up table for pixel brightness values to record how well the pixels t the distribution of pixels in a histogram model. We described above how to find skin-color pixel in a color image. Thus, we have a histogram of fresh color then we can use back projection to find fresh color areas in an image.

ISSN: 2277-5528
Impact Factor: 2.745 (SIJF)

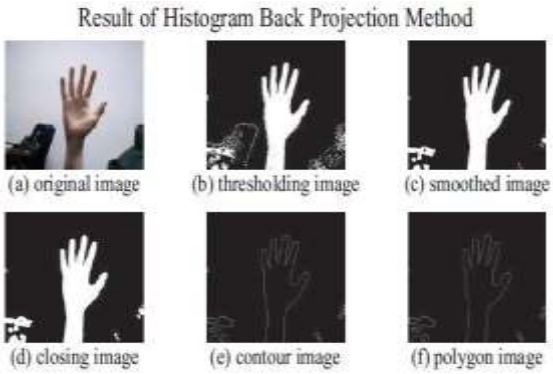


***Figure 6: Step by step contour extraction of hand region***

## CONTOUR FINDING

A contour is a list of points that represent, in one way or another, a curve in an image. An image contour is necessary in recognition. The centroid of interior region of hand and characteristic points of a contour of the region represent the structural features of hand. If we are drawing a contour, it is common to approximate a contour representing a polygon with another contour having fewer vertices. The polygonal approximation of the shape consists on finding significant vertices along the contour such that these vertices constitute a good approximation of the original contour. A classic approach to this problem is to take the high curvature points (i.e., points with high absolute value of curvature) as significant vertices. Contour representation of binary image in Figure 6.6(d) is depicted in Figure 6.6(e). We used Suzuki's algorithm for contour finding. The centroid of the region and the vertices of a polygon that approximates a region contour constitute characteristic points that are used for defining a structural representation of an image. For a contour depicted in Figure 6.6(e), its corresponding polygon is shown in Figure 6.6(f).

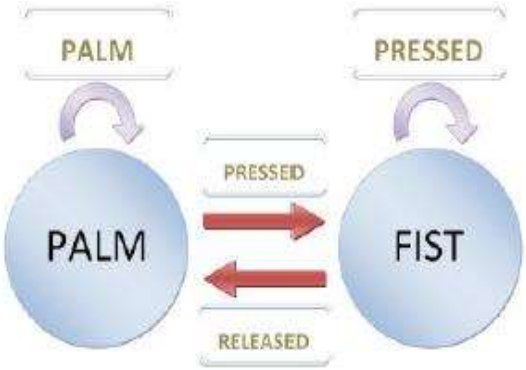


***Figure 7: The FSM of mouse released and pressed gestures with two states***

## GESTURE RECOGNITION

By the recognition of 4 postures we have handled 3 mouse actions already which are they mouse move, left click and right click. To recognize other gestures, we needed a mixture of these postures in a sequence. For this, we defined a gesture as a sequence of states and modeled by a finite state machine (FSM). FSM has been well received in various areas that are in common since it basically uses string matching between a data sequence and the state sequence of an FSM. Gesture recognition is also relevant with FSM in this aspect.

### i) Mouse Pressed and Released

These mouse actions are sequences of palm and fist postures. As shown in the diagram, mouse is pressed if posture turns from palm to fist or remain as fist. On the other hand, if posture returns to palm from fist it is described as

mouse released action as shown in Figure 6.7.

**ii) Double Left Click**

When oblique fore finger posture (mouse left click action) is detected in a frame, a timer is triggered and when another left click action is detected in another frame we look for if previous frame was palm and the timer is less than 500 milliseconds, if such is the case then we label it as double left click action.

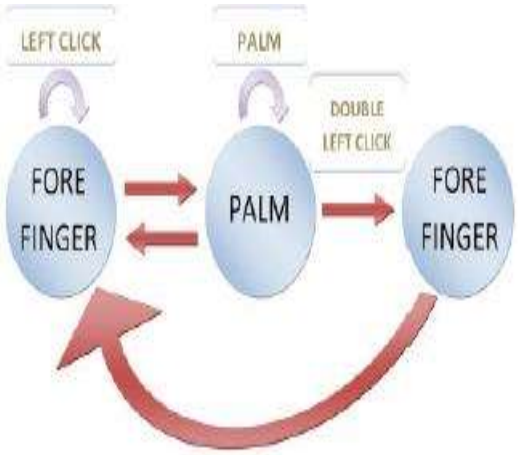


**Figure 8: The FSM of double left click gesture with three states**

## PIXMIX ALGORITHM FOR VIDEO OPERATIONS

The presented video inpainting system will be able to offer a coherent good quality effect for a web camera of low resolution with active translations and rotations. The object has to be static in nature and the background has to be planar for the homographic determination of the contour of the objects. For a static camera or if the movements of the camera are limited a consistent video stream can be presented for the nonplanar background. Our method supports real time applications with the camera that moves about a static object that has to be removed. It does not support the situations where the object moves about or the situations where both the camera and object are dynamic. The approach does not take into consideration the situations where the unknown background becomes visible at a later stage.

These conditions would occur if an object with a major volumetric expansion (perpendicular to the surface) in combination with dynamic camera movements has to be detached. An earlier created image area is then substituted by the real image content. Additional fading mechanisms is required for obtaining a coherent view. On the other hand, as a real-time video inpainting system lacks information about future video frames (in difference to eliminating objects from a video which is prerecorded). The system developed will be able to produce a good quality result and a sound video stream. The finger-print based object assortment method is an important improvement. If the individual characteristics of the background and the object are more and the characteristic difference of the background and the object is less than it becomes difficult to make a consistent selection.

## CONCLUSION

In this paper a system for real time hand gesture recognition is developed. The primary goal of the research is to create a system that can identify human generated gestures and use this information for device control and various operations on image restorations. This system is aimed to make access the system environment simpler so that the person who is unaware about the system can handle it with few simple gesture commands. This system very useful for the physically disabled users who lack the strength and precision used to operate the traditional input devices. In future, Using gestures like eye blinks, head motions, hand gestures would help them to communicate effectively.

## REFERENCES

[1] Attila Licsar and Tamas Sziranyi, IEEE Int, “Hand Gesture based film Restoration ACM Cybernetics” pp. 215-22, 2004.

[2] Asanterabi Malima, Erol Özgür, and Müjdat Çetin “A Fast Algorithm For Vision-Based Hand Gesture Recognition for Robot Control”IEEE Int, Conf. Systems, Man, Cybernetics. pp. 289-296, 2008.

[3] Vafadar and Alireza Behrad "Human Hand Gesture Recognition Using spatio- Temporal Volume for Human Computer Interaction" IEE Int. Symposium on Telecommunications, pp.713-718,2008.

[4] Byung-Woo Min, Ho-Sub Yoon, Jung Soh, Yun-Mo Yang and Toshiaki Ejima, "Hand Gesture Recognition Using Hidden Markov model" IEEE Int Conf. systems,Man, and Cybernetics, 1997 , pp 4232-4235.1997.

[5] Wing Kwong Chung, Xinyu Wu and Yangsheng Xu, "A Realtime Hand Gesture Recognition based on Haar Wavelet Representation" Proceedings of the 2008 IEEE Int. Conf. on Robotics and Biomimetics , pp.336-341,2008.

[6] Mu-Chun Su "A Fuzzy Rule-Based Approach to Spatio-Temporal Gesture Recognition" IEEETrans on systems, Man, And Cybernetics –Part C: Applications and Reviewws, Vol.30No 2, pp276-281,2000.

[7] Byung-Woo Min, Ho-Sub Yoon, Jung Soh, Yun-Mo Yang and Toshiaki Ejima, "Hand Gesture Recognition Using Hidden Markov model" IEEE Int. Conf. systems,Man, And Cybernetics, 1997 , pp 4232-4235,1997.

[8] M K Bhuyan, D Ghosh, P K Bora,"Feature Extraction from 2D Gesture Trajectory in Dynamic hand Gesture Recognition" 1–4244–0023–6/06/ IEEE pp. 432-441,2006.

[9] Shewta Yewale ,Pankaj Bharne,"Artificial neural network approach for hand gesture recognition" International journal of engineering Science and technology, Vol 3 No.4,pp. 2603-2608, 2011.

[10] Zhou Ren Jing jing Meng Junsong Yuan,Zhengou Zhang, "Robust hand gesture recognition with kinect Sensor" ACM multimedia 2011 conference ,pp. 978-979.2011.

[11] Hamid A Jalab,"Static hand Gesture recognition for human computer interaction" Asian Network for Scientific Information information technology journal, pp. 1-7,2012.

[12] Qing Chen, Nicolas D. Georganas, "Hand Gesture Recognition Using Haar-Like Features and a Stochastic Context-Free Grammar", IEEE transactions on instrumentation and measurement, vol. 57, no. 8, pp.1562-1571,2008.

[13] Christopher Lee and Yangsheng Xu, "Online, interactive learning of gestures for human robot interfaces" Carnegie Mellon University, The Robotics Institute, Pittsburgh, Pennsylvania, USA,1996.

[14] Vafadar,Alireza Behrad,"Human Hand Gesture Recognition Using spatio-Temporal Volume for Human Computer Interaction" IEE Int. Symposium on Telecommunications, pp.713-718,2008.