ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization *Volume 4, Special Issue 11, September 2015*

National Conference on Emerging Trends in Computing, Communication & Control Engineering [NCET 2K15]

On 4th September, 2015

Organized by

Department of CSE, ECE & EEE, Magna College of Engineering, Chennai-600055, India.

Hand Gesture Recognition System

[1]V.Michael Maria Lavanya, [2]J.Ancy, [3]M.Anitha, [4]E.Mala

[1]UG Student, Department of CSE, Rajas Engineering College, Vadakkangulam, India

[2]UG Student, Department of CSE, Rajas Engineering College, Vadakkangulam, India

[3]UG Student, Department of CSE, Rajas Engineering College, Vadakkangulam, India

[4]UG Student, Department of CSE, Rajas Engineering College, Vadakkangulam, India

ABSTRACT: The aim of this paper is to provide with a real time application based on hand gesture recognition. Virtual touch is a concept whereby any normal surface with no internal circuitry and hardware for touch sensitivity, is converted into a touch sensitive surface by using image processing techniques.Using such a technique the paper presents the transformation of a plain white wall into a large touch wall. The virtual touch wall aims at building a gesture sensitive environment for users, to provide them new ways of interaction in their work atmosphere and in their recreational space. The hardware requirements of the system are kept on the minimum , limiting only to a simple USB webcam and a PC .Other alternatives to the PC have also been presented , such as the Beagle Board , which is a mini version of a PC and very inexpensive .The approach is based on simple and fast motion detection trying to eliminate unnecessary regions of interest and a recognition algorithm based on the hand contours , their convexities and a further comparison of hand shapes based on HU moment matching which works on image contours .The paper also presents a way of improving the distorted hand contours by distinguishing between external contours and holes and explicitly filling up holes .The use of a predefined set of contours for each gesture helps in better recognition .A simple state machine is implemented to further improve reliability .A total of 5 gestures have been implemented and tested on the PC and the Beagle Board and the performance compared. The Linux XLIB library and the X display functions have been utilized to control mouse cursor actions and other display properties of the wall.

KEYWORDS*:* Gesture Recognition, HU Moments, Hand Tracking, Contours, Xlib, Beagle Board, virtual gesture wall.

I. INTRODUCTION

One of the greatest gifts of almost all living things in this world is the ability to communicate through vocals and actions. Human and machine interaction by far has only been through simple means of communication like a mouse ,keyboard or switches .Voice recognition and gesture recognition are two powerful and more natural means of communicating with machines as is with human beings.There are innumerable instances where conventional applications can be replaced with hand gestures. In many cases special gloves or markers have been used for efficient detection and tracking which constraints the user and other vases demand a very simple and plain background.A very real time method of hand gesture recognition based on convexity defects is presented but an analysis based only on convexity defects is dependent on smoothness of background.Implements gesture recognition based on Hausdroff distance but is not real time especially when implemented on dedicated systems such as Beagle Board This paper attempts to solve these problems by proposing a realtime gesture recognition system with no userconstraints and with any kind of background environment.Further the paper presents a practical application of gesture recognition by implementing an interactive gesture based virtual touch wall. Such a system consists of a projector, a projecting surface and a processing unit. Interaction between the user and the projecting surface is based on gestures than on conventional methods such as mice and keyboard. Such a system provides a new dimension for users in their work, gaming and recreational aspects of their lives.

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 4, Special Issue 11, September 2015

National Conference on Emerging Trends in Computing, Communication & Control Engineering [NCET 2K15]

On 4th September, 2015

Organized by

Department of CSE, ECE & EEE, Magna College of Engineering, Chennai-600055, India.

II. SYSTEM ARCHITECTURE AND HARDWARE BLOCKS

Based on hardware level abstraction, the entire system is sub divided into three major hardware modules. They are the Beagle Board module, the Camera Module and the Projection Module.

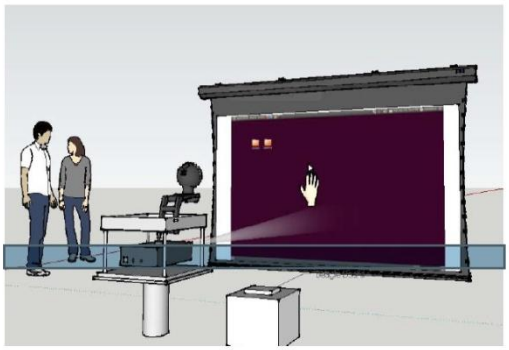


Fig. 1. Physical system setup

A. The Beagle Board Module

The role of this module is to accept frames from the camera module and perform several image processing algorithms on the frame and display its output on the projection module. Beagle board is a low-cost, single board computer developed by Texas Instruments. It is equipped with OMAP-3530 System on chip processor, which includes an ARM Cortex-A8 core. Ubuntu Natty 11.04 is installed on the Board using an external SD card, prior to which the card is set up with an MSDOS type boot partition and the Linux File System partition[12]. The OS which is in the form of a tar file is untarred using the command:

tar –xjvf armel-rootfs.tgz

This function copies the root file system to the file system partition. The boot files namely MLO, U-boot and Uimage are copied to the Boot partition manually in specific order. The SD card is then placed on the SD/MMC card slot of the board and the user button is used to boot the OS. Fig. 2a shows the Beagle Board and its internal architecture.

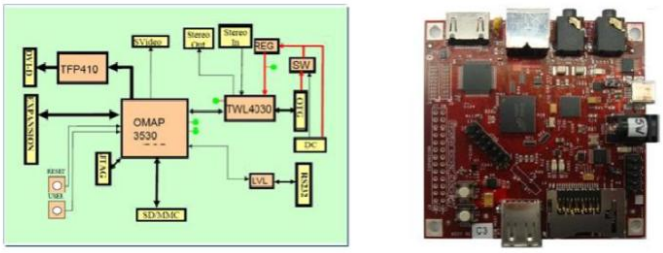


Fig. 2(a). The Beagle Board and its System architecture

Opencv which is a library of programming functions is used to perform computer vision algorithms and operations on the Beagle Board or any Linux system

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 4, Special Issue 11, September 2015

National Conference on Emerging Trends in Computing, Communication & Control Engineering [NCET 2K15]

On 4th September, 2015

Organized by

Department of CSE, ECE & EEE, Magna College of Engineering, Chennai-600055, India.

B. The Camera Module

The camera provides frames to the Beagle Board in real time. The webcams tested and used are the Logitech C100 and the Logitech C200.

C. The Projection module

The projection module consists of a projector to project the operating system of the Beagle Board and a projecting surface. Any smooth white surface works well with the system.

III. WORK FLOW

There are five stages in the working of the system as shown in the Flow diagram in Fig. 3.a. The five stage are initialization, Acquisition, Pre-Processing before analysis, Posture Recognition and Execution. These five stages occur in succession ass long the system is active.

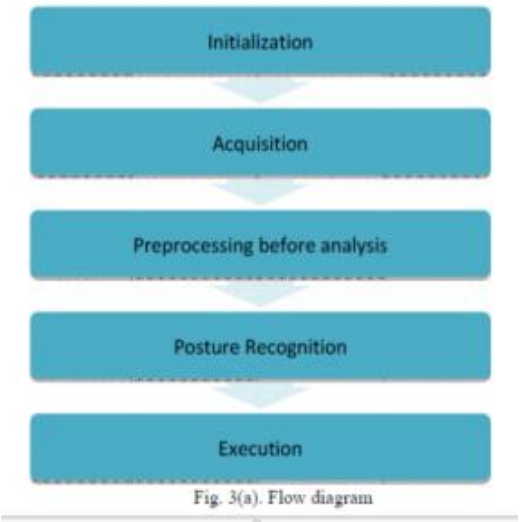


Fig. 3(a). Flow diagram

A. Initialization

A set of recognizable postures are stored in the system memory at the startup of the system.These postures correspond to the contours of different hand gestures .They are very essential in the later stages of Hand Recognition.

B. Acquisition

Onnce the system is begun, every frame from the webcam is acquuired and processed .Each frame is a 3 channel GB fram with each channel having 8 bit depth The size of ach fram is 640X480

C. Pre Processsing Before Annalysis

A series of imagge processingg operations arre undertaken on each frame captuured from the webcam These operations motiondetection,segmentation,smoothing.

IV. GESTURE RECOGNITION MODULE

The steps involved in recognizing hand gestures are as shown in Fig.

A. Motion Detection

Every frame from webcam is converted into a gray scale image and a pixel wise subtraction is performed between the current frame

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 4, Special Issue 11, September 2015

National Conference on Emerging Trends in Computing, Communication & Control Engineering [NCET 2K15]

On 4th September, 2015

Organized by

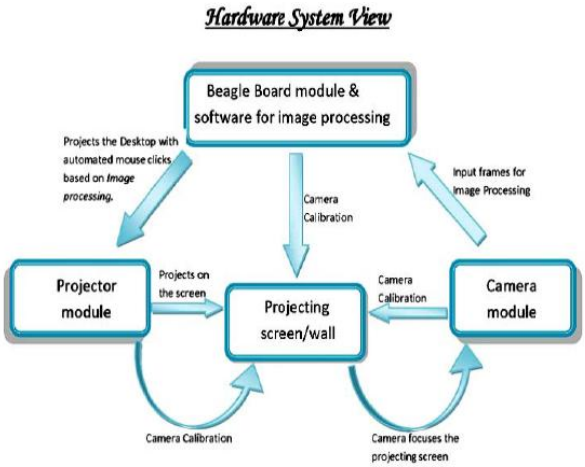
Department of CSE, ECE & EEE, Magna College of Engineering, Chennai-600055, India.

and previously acquired frame. The resulting output is then smoothed using a Gaussian filter to move noisy pixels .A threshold is now applied on smoothed image thereby changing it into a binary image with maximum value 2555 and minimum value 0 .2555 corresponds to pure white and 0 corresponds to pure black. The requirement of a fixed background is eliminated by using motion detection .The background subtraction image would immediately adapt to new background with an error only in the first frame after change. Change in background should not be continuous in time domain, but can be discreet at random time intervals. A simple threshold of 10 is experimented to work well

If image(i,j) >10 ;
then image(i,j)=255;
Else image(i,j)==0;

B. Hole Filling and Skin Segmentation

Output of motion detection does not yield the full area of the object under motion .As seen in Fig. 4.1b there are many black regions or holes with in white boundaries in the motion detection image .For this reason, a hole filling operation is perform ed .Dilation reduces number of holes, but hand contour will still not be devoid of holes. Fig. 4.2a show Dilation output .At this stage multiple dilations are not performed as this increases size of unnecessary regions in the image.



A logical AND is performed between he dilated image and the original frame. This new image is now converted into YCrCb color space and an upper and lower bound are set to color segment it This process is not accurate, but will suffice because most off the image iis eliminated in the motion detection image leaving apart only very mall areas which move. The lower mit is (0,113,,67) and the upper limit is (255,173,127) .Combination off background subtraction and skin segmentation is shown inn Fig. 4.2b which is the final output.

C. Finding appropriate Contours

A contour analysis gives groups of pixels that have some label due to their connectivity inn binary image. After such an analysis the image is treated as a set of contours and not as a collection of discreet pixels. After contour analysis individ ual contours with a minimum size off 4000 pixels are isolated.

Fig. shows the hardware block diagram of the entire system

D. Contour Cropping and Contour Approximations

The most useful information of the lies in its fingers .Hence the hand contour is cropped to the upper portion, leaving only the orientation of the fingers to analyze .Fig. 4.4a shows the upper half contour.This output contour is still uneven in shape

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 4, Special Issue 11, September 2015

National Conference on Emerging Trends in Computing, Communication & Control Engineering [NCET 2K15]

On 4th September, 2015

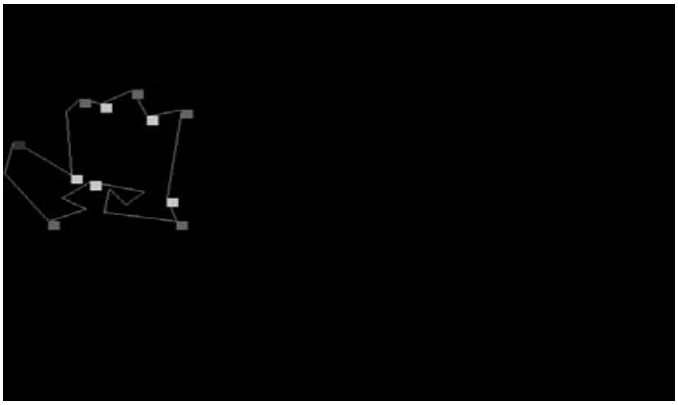
Organized by

Department of CSE, ECE & EEE, Magna College of Engineering, Chennai-600055, India.

with very small distortions.Hence a polygon approximation of the contour is done to approximate the upper part of the hand as a polygon .This yields a contour derived of straight line segments making it easier in finding convexities .The result of polygon approximation is shown in Fig. 4.4b.

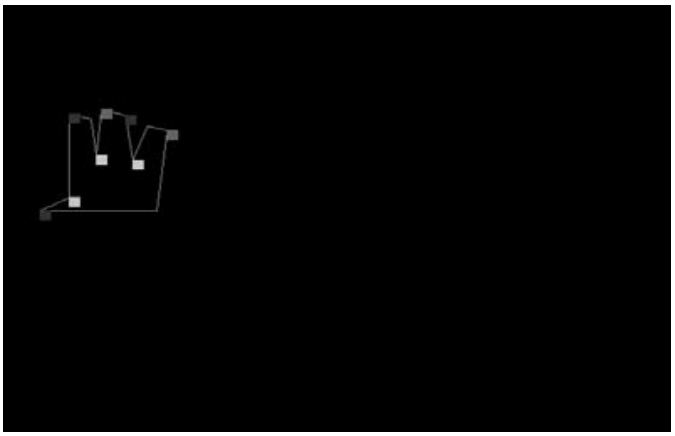
E. Finding Convexity Defects

Convexity defects are valley points. In particular, convexity defects are sequences of contour points between two consecutive contour vertices on the contour Hull. Fig. 4.5a , 4.5b , 4.5c , 4.5d , 4.5e shows the convexity defects in a palm , open thumb , V shaped fingers , a fist , and in three fingers .These convexity defects are very useful in providing information about the location and state of the fingers . In the figures below, the pure white markers are the convexity defects or the valley points and the dark markers are the start and end points respectively.



F. Eliminating Convexity Defects

As shown in Fig. not all the convexity defects provide useful information .Some defects are formed especially due to incomplete and distorted contours. Most of the defects which are on the underside of the contour are not related to the fingers. Every defect has a start and an end point .The not removed .They are marked as special defects, as they are required at later stages.



A minimum bounding rectangle is a rectangle of minimum area which bounds the contour. The rectangle and coordinates

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 4, Special Issue 11, September 2015

National Conference on Emerging Trends in Computing, Communication & Control Engineering [NCET 2K15]

On 4th September, 2015

Organized by

Department of CSE, ECE & EEE, Magna College of Engineering, Chennai-600055, India.

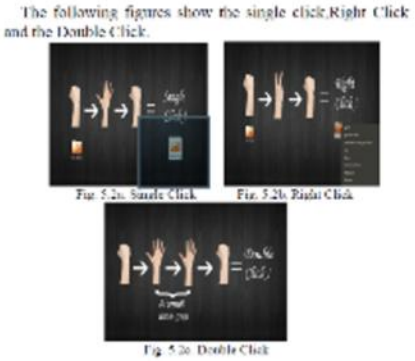
of its vertices are calculated. In case the first defect starts with a special defect, then the horizontal distance from the special defect to its next defect (in the updated list of defects which has some unnecessary defects removed) is measured. In case the first defect is not a special defect, then the horizontal distance between the nearest vertical edge of the bounding rectangle and the defect is measured .After that, the distance between every two consecutive defects is measured. That is (defect2.xcoordinate – defect1.xcoordinate) (defect3.xcoordinate – defect2.xcoordinate) and so on are measured until the last defect is reached .If the last defect is a not a special defect , then distance between last defect and nearest vertical side of bounding rectangle is measured .If it is a special defect , then it is ignored . Based on measured distances, a specific range is provided to approximate the number of fingers .In this case, with a distance greater than 75, number of fingers is approximated as 3 .With distance greater than 50 number of fingers approximated is 2 and a distance greater than 15 one finger is approximated .

G. Contour Matching

That is if two fingers were found out, it is first compared with the V shaped set of contours. If 5 fingers are found, then it is first compared with the Palm set of contours. Hu moments are said to be invariant to scale, translation and Rotation.

H. Experiments

Experiments on a set of five people with varying skin tone have resulted in the following outcomes .The following table illustrates the accuracies of the output of the gesture recognition module.



V. LINKING MODULE

This module is the follow up of gesture recognition. Once a gesture is recognized,

A. Pointer Movement via Hand Tracking

In the gesture recognition module, a contour analysis on each of the input frames resulted in identification of ligible hand contours in each of the frames. These hand contours were further verified using HU Moments. From the hand contour, the coordinates of the center of the hand are found out . They are similar to any other shell scripting code snippets. These

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 4, Special Issue 11, September 2015

National Conference on Emerging Trends in Computing, Communication & Control Engineering [NCET 2K15]

On 4th September, 2015

Organized by

Department of CSE, ECE & EEE, Magna College of Engineering, Chennai-600055, India.

snippets use the X display as their interface.Xlib can draw simple solids , receive events ,write text in various fonts and also do primitive inter process communication. The Xwarp pointer function which is an inbuilt Linux function is used to move the pointer to the desired location in the X Display. pointer by setting offsets relative to the current position of the pointer.
The following figures illustrate Pointer Movement via and Tracking. Consider two discreet time intervals t1 and t2 (t1<t2). At each time instant the movement of the hand is
Tracked.

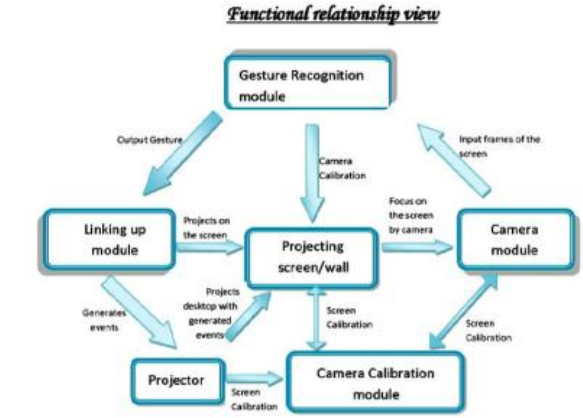


Fig. 3(b). Functional relationship between the three software modules

B. Mouse Clicks Via Gesture Recognition and Permissible States
Depending on the output gesture recognized , several states and functionalities have been defined for the entire system.Three gestures have been selected and assigned particular tasks. The fist, the palm and the V shaped Fingers.To achieve these functionalities , the Xlib is used again.There are three basic types of pointer events in the X-Library.They are Button press ,Button release , Motion Notify. The motion Notify event is generated by the system whenever the pointer is moved and when the pointer motion begins and ends in a window. The Button press event is used to signify the left and right clicks and the Button release event is used to signify the left and right click releases. The different functions used to perform the above functionalities are XCloseDisplay, XFlush ,XWarppointer.The table below shows the permissible states for an event such as mouse clicks to occur.

VI. CAMERA CALIBRATION MODULE

The camera calibration module is pivotal for proper functioning of the entire system.Camera calibration is a necessary step in 3D computer vision in order to extract metric information from 2D images.Improper calibration in the proposed system could mean that the pointer would not move in accordance to the movement of the hand as the distance covered by the hand and the distance covered by the pointer are not the same .Hence the accuracy of the entire system would falter with improper calibration techniques . The procedure is as follows:

A checkerboard pattern is displayed on the screen and camera is focused at the grid to capture a series of images. Positions of pattern corners are known with respect to a co-ordinate system. (The Resolution of Desktop screen can be used as a measure).Images Corners are obtained from the captured Images.Obtain Equations that that describe imaging and contain internal parameters of camera.

The captured images from the camera are processed to obtain the location of the corners. This can be done using an

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization *Volume 4, Special Issue 11, September 2015*

National Conference on Emerging Trends in Computing, Communication & Control Engineering [NCET 2K15]

On 4th September, 2015

Organized by

Department of CSE, ECE & EEE, Magna College of Engineering, Chennai-600055, India.

openCV code which uses color segmentation to know the position of square. This process is repeated for all the four corners through which the borders of the screen are obtained. Then, two horizontally aligned squares are recognized to obtain the distance between them. This can also be done using the above technique of color segmentation. The results now obtained from the images are according to the calibration settings of the camera. Having known the desktop resolution, the distances on the desktop are mapped to the distances obtained from the images. A linear function is obtained which describes the relation between the distances.According to these results and the linear function, the calibration settings of the camera are changed to minimize the calibration error.

REFERENCES

1. A. S. Pentland, *Computer Vision For Human-MachineInteraction;*eds.Cambridge University Press,ch1, pp.3-21, 1998.
2. T. F. William and D. W. Craig, "Television Control by Hand Gestures," in *IEEE Intl. Workshop on Automatic Face and GestureRecognition ,Zurich* ,June 1995.
3. S. Ju, M. Black, S. Minneman, D. Kimber, "Analysis of Gesture and Action in Technical Talks for Video Indexing," in *IEEE Conf. on Computer Vision and Pattern Recognition,CVPR 97*
4. VanDam, "Post-WIMP user interfaces," in Communications of the ACM, vol 40, pp. 63-67, February 1997.
5. [7] C. Maggioni. Gesturecomputer. "New Ways of Operating a Computer," in *Proc. Int´l Workshop Automatic Face and Gesture Recognition*, 1995.
6. [8] C. Manresa , J. Varona , R. Mas, and F. J. Perales "Real –Time Hand.
7. Tracking and Gesture Recognition For Human –Computer Interaction," in *Electronic Letters on Computer Vision and ImageAnalysis* 0(0), pp.1-7, 2000.
8. E. Sanchez-Nielsen, L. Anton-Canalis, and M. Hernandez-Tejera "Hand Gesture Recognition For Human-Machine Interaction," *Journalof WSCG* ,vol. 12, no. 1-3 , ISSN 1213-6972.