International Journal of Engineering Research in Electronic and Communication Engineering (IJERECE) Vol 3, Issue 3, March 2016

# Implementation of FPGA Based Hand Gesture Recognition System for Blind People

[1] Salma Shapurkar [2] Bhooshan K Kalamani [3] Basavaraj I Subhedar [4] Mallikarjun Kallimani
[5] Tejashree Deshpande
[1] Assistant Professor [2][3][4][5] UG Students,
[1][2][3][4][5]Department of Electronics& Communications Engineering, Jain College of Engineering, Belagavi
[1]salma01bgm @yahoo.com [2] bhooshan0025@gmail.com [3] basusubhedar55@gmail.com
[4] manikya.kallimani@gmail.com [5] tejashreesdeshpande@yahoo.com

*Abstract-* **The proposed work describes the design and implementation of the hand gesture recognizer system for controlling the hardware appliance in real time. FPGA based implementation proposed for hand tracking system includes the image preprocessing state and feature extraction state that consists of bounding box and Center-Of-Mass based computation in terms of area segmentation. Through the features extraction state, the object's Center-Of-Mass and bounding box attributes are extracted to be applied for gesture sign classification. The main modules in this image processing system employ the developed system exploits the parallelism architecture of the FPGA to achieve real time processing. There are many applications using hand gesture as a nature control interface, such as human machine interaction and interactive entertainment. With the advancement of computer vision and machine learning, the development of vision-based hand gesture recognition systems has received more and more attention in recent years. However, much previous research was focused on software algorithms running on a PC-based platform, which makes it inapplicable to real-time applications. In this proposed work we have implemented gesture recognition system based on improved optical flow was proposed to accelerate the execution and achieve real-time operation. Experimental results showed that the proposed framework is feasible and can reach the speed of 30 frames per second.**

***Keywords*—FPGA, image processing, hand gesture recognizer system, gesture classification**

## I. INTRODUCTION

In recent years, a number of vision-based applications have been proposed for gesture recognition in [1]-[5][11]-[16]. The gesture information can provide a rich signal or message for home entertainment. Several studies developed various methods or solutions to meet the requirements of various applications. In [2], the authors built a powerful man-machine interface by using Laplacian of Gaussian (LOG) operators, motion energy, and finite state machines for feature extraction. However, a disadvantage in this system is that the motion energy requires large computation. In [3][4], a gesture recognition system was proposed for Human-Computer Interaction (HCI). The AdaBoost and Support Vector Machine (SVM) were used to detect hand location and to recognize hand type, respectively. However, a disadvantage in this system is that the pre-training process is required. In [5], a remote control system by motion detection and open finger counting was built. To enhance the accuracy of object detection, the entire candidate region is assigned a threshold for selection of detected objects. In object detection [6]-[8], the color processing is based on skin color information. In recognition, the k-cosine in [9][10] was used; and two procedures of predefined and pre-training were also required. A disadvantage in the embedded system is that of mathematical operation, such as cosine calculation. The four modules used in [11][12] include real-time hand tracking and extraction, feature extraction and gesture recognition, which were integrated for gesture application. In general, the image difference operation in temporal domain, skin color detection, edge detection and region identification were used for feature extraction.

The project aims at developing a real time system for establishing the dumb communication. For the success of the project it is equally important to study about the existing techniques which are helping the dumb to communicate. The state of art technology refers to the existing methodologies for achieving this aim. Owing to the ever increasing needs of the society, day by day technology is improving. Many of the present day researches are undertaken with the aim of building some socially relevant products which may help the people to advance technologically. Even after all these developments in the

technology, it's a bitter fact that there has not been any significant works done in helping the dumb people to communicate. Even today we depend on a translator to understand what the dumb people are trying to convey. Translators are people who can understand the sign language of the dumb. They manually decode their signs and convey the message to the world. But unfortunately in most of the cases such a translator will be absent and it makes the dumb another class of the society who have no methods to atleast convey their messages like other normal human beings. Thus we can say that the state of art technology revolves around the presence of a human translator. This inspired us to think of a technology which can be handy, easy to use as well as efficient in its function.

Two concepts of Low complexity and FPGA-based architecture were considered. Based on the analysis, the embedded gesture system is a crucial trend for future application; therefore, this study proposes an FPGA-based gesture recognition system. To avoid the design of high complexity, the developed algorithm was improved for parallel processing. Mice, keyboards, joysticks and buttons are categorized as a physical controls type of interaction for human computer interaction in virtual reality or augmented reality [1]. Generally, to control any hardware appliances, user will be furnished with these traditional controllers. With the advancement of technology conventional controllers could possibly be replaced with a more natural interface such as gesture control interfaces. The gesture control interfaces proposed include vision-based gesture recognition system, data-glove-based gesture recognition system, touch based recognition system and etc. Touch based recognition system that uses two touch pads as the interface medium was proposed to reduce the switches and the buttons which the user needs to deal with in automation [2].

## II. RELATEDWORKS

In [1] the problem of gesture recognition is narrowed down to that of hand gesture recognition and specifically deals with finger count extraction to facilitate further processing using the control so effected. A hand gesture recognition system has been developed, wherein the finger count in a certain input hand image is computed in accordance with a simple yet effective procedure. The problem of hand gesture recognition is solved by means of adopting a lucid albeit efficient algorithm which has been implemented using the Xilinx System Generator software tool. The algorithm followed is invariant to rotation and scale of the hand. The approach involves segmenting the hand based on skin color statistics and applying certain constraints to classify pixels as skin and non skin regions. A different method had been developed for Hand Gesture Recognition for Indian Sign Language which consisted the use of Camshift and HSV model and then recognizing gesture through Genetic Algorithm. In that, applying camshift and HSV model was difficult because making it compatible with different MATLAB versions was not easy and genetic algorithm takes huge amount of time for its development. [2]. In [3] in order to get hand gesture feature vectors, the system adopts a vision-based hand tracking approach by using hand gesture segmentation algorithm. The system downloads those feature vectors data from large hand gesture feature vectors data base into the on-chip cache memory of an AP (Associative processor), then performs gestures matching in an extremely short time. Although gestures recognition processing is computationally very expensive by software, latency free recognition becomes possible due to the highly parallel maximum likelihood matching architecture of the AP chip. In [4] Byung - woo min et al, presented the visual recognition of static gesture or dynamic gesture, in which recognized hand gestures obtained from the visual images on a 2D image plane, without any external devices. Gestures were spotted by a task specific state transition based on natural human articulation.In [5] a method had been developed by P Subha Rajan and Dr G Balakrishnan for recognizing gestures for Indian Sign Language where the proposed that each gesture would be recognized through 7 bit orientation and generation process through RIGHT and LEFT scan. The following process required approximately six modules and was a tedious method of recognizing signs Different software tools have been used in designing gesture base recognition system. These include implementation of the algorithm in Matlab, C++ OpenCV, and JavaScript etc. Various kinds of image processing algorithms were proposed for tracking and recognizing images effectively such as Haar Classifier algorithm introduced by Viola and Jones that had been widely used in object detection [5]. Two level approaches to achieve real time vision-based hand gesture classification was proposed [6]. Using this approach, the low level approach implements the posture recognition with Haar-like features and the AdaBoost learning algorithm while the high level approach implements the hand gesture recognition using a context-free grammar-based syntactic analysis. Besides that, using commodity PC and two low-cost web cameras, a method to track 3D position and 2D orientation of the thumb and index finger of each hand was proposed.

After going through the feature extraction process, the hand is recognized though the peak and valley detection [7]. Although the system could achieve the real time requirement, the sophisticated hardware involved allocates a major challenge for portability and flexibility of the system. When the amount of works needs to be processed increase, the ability of the system to handling the growth will be effected. Instead of implementing the algorithm through software tools in PC, there is an alternative to implement it

in FPGA using Hardware Description Language or NIOS embedded processor. The speed of image processing of two FPGAs was compared to the recent processors with SIMD instructions and multi-cores [8]. The image processing problems tested are two-dimensional filters, stereo-vision and k-means clustering.

Although the quad cores processor can achieve real-time processing when image size is small, all resources will be used up to fulfill the real-time requirement. Unlike FPGA implementation where large quantity of hardware resources would still be available and thus allows more sophisticated work to be executed. Real time applications for image processing which can be processed by processors are still very limited. For practical real time applications, FPGAs are still needed. High performance shown by FPGA is due to high parallelism architecture and huge amount of internal memory banks which can be retrieved in parallel [8]. A system which works by taking in still images of hand gesture to interpret their form and executing a command for controlling the music player was successfully created [9]. After filtering all the unnecessary pixels, the related pixels are recognized using the bounding box. To recognize the gesture input, the related pixels are counted and compared within the bounding box. By cutting the bounding box into halves along Y-axis and X-axis, the portion containing the less related pixel determines the direction pointing by the user. The interpreted information was taken by the NIOS II soft-core embedded processor to commence music playback as well as LCD screen display [9].

To achieve real time processing and increase the performance in image processing, the hand gesture recognizer system was implemented in FPGA utilizing the parallelism feature offered to provide a more natural ways of interacting with hardware devices. Bounding box algorithm is used to recognize hand, while the target object's Center-Of-Mass is computed to track its location which subsequently being used for gesture classification purpose.

## III. WORKING PRINCIPLE OF SYSTEM

Gestures are a form of nonverbal communication in which visible bodily actions are usedto communicate important messages, either in place of speech or together and in parallel withspoken words. Gestures include movement of the hands, face, or other parts of the body.Physical non-verbal communication such as purely expressive displays, or displays of joint attention differ from gestures, which communicate specific messages. Gestures are culture specific and can convey very different meanings in different social or cultural settings. Gesture is distinct from sign language. To establish a communication or interaction with deaf and mute people is of utter importance nowadays. These people interact through hand gestures or signs. Gestures are basically the physical action form performed by a person to convey some meaningful information. Gestures are a powerful means of communication among humans. In fact gesturing is so deeply rooted in our communication that people often continue gesturing when speaking on the telephone. There are various signs which express complex meanings and recognizing them isa challenging task for people who have no understanding for that language. The Sign language is very important for people who have hearing and speaking deficiency generally called deaf and mute. It is the only mode of communication for such people to convey their messages and it becomes very important for people to understand their language. Indian sign language is set of signs used in India. Different nations follow different signs as per their areas of interest and actions used. The government of India has even started an official website which gives us information about the signs used by the dumb in India and also includes lots of activities which helps the dumb to learn and live well in a society. Some gestures used in the society are as shown below.

***Figure 1.Generally used Gestures***

## IV. DESIGN AND IMPLEMENTATION OF THE SYSTEM

The design of the hand gesture recognizer system can be divided into 4 modules namely camera, image processing & control unit, display unit and the hardware appliance. In the design, CCD camera TRDB-D5M is used in the Altera Development Education 2 (DE2) board to provide fast image acquisition. The camera is switched to FREE RUN mode in order to capture images continuously. The images captured are sent to DE2 board and processed by the FPGA. Output of the image processing and control unit is then sent as control signals to the respective hardware appliance. The display unit is used to monitor the output result of the image processing module.

Colour segmentation technique [10] known as colour thresholding is used in this system to differentiate human hand from the captured image. Thresholding is an idealistic method used for stream processing as it can be carried out easily through FPGA [10].Red colour glove is used in this system to help distinguishing the human hand from the environment because of its distinct colour that is easily differentiable. By applying threshold to eachdata pixel using colour, the image captured can be divided into object image and background image. To achieve this, an initial threshold value is set randomly by testing with trial and error. Then, the intensity of red pixel is checked if it is more than the threshold value. If the condition is fulfilled, next operation is carried out to check whether the sum ofthe pixel green and blue intensity values is less than the threshold. The pixel is said belongs to the target object if these two conditions are fulfilled. Otherwise, the unwanted pixel is filtered out by converting its RGB values into gray colour. Hence, a pixel belongs to the colour class if its RGB intensity value is in the specific ranges. Under different environment, the lightning issue will dramatically affect the threshold value applied to each pixel of the image captured. Thus, the threshold value is manually controlled through the push button and toggle switch on the DE2 board. The data pixel processed through segmentation process is fed into feature extraction stage where it consists of two main image processing module which play an important role in recognizing the hand. The data pixel fed into these two main modules is processed concurrently, utilizing the parallelism architecture of FPGA. The segmentation algorithm used in this system is shown in Fig. 2.

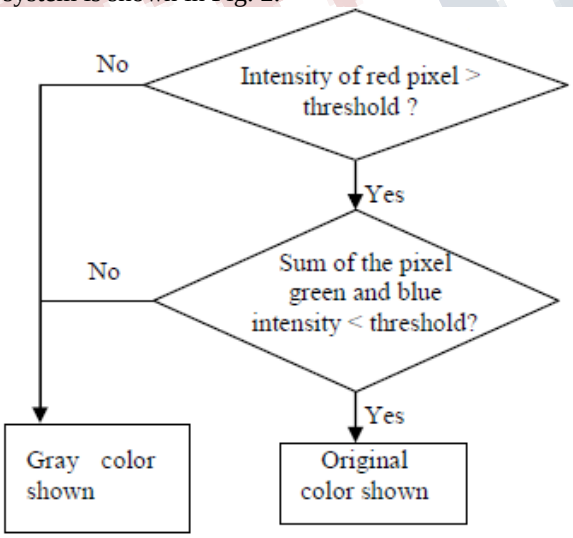


***Fig. 2: Segmentation algorithm implemented in Hand Gesture Recognizer System***

In the hand gesture recognizer system, human hand is detected and recognized by building a bounding box confining all of the related pixels inside a colour category. The bounding box is used to calculate the exact ratio, position and size of the target object by defining topmost, leftmost, and right most and the bottommost pixel belonging to the target object [11]. The first stage of the bounding box computation is the attribute identification. The pixel is checked if it is belonging to the target object. The SRAM memory is used to store and keep track on the data pixel information where bit'1' indicates the target object and bit '0' indicates the background object. The first detected pixel initializes the values for the bounding box of the target object. The pixel XY -coordinates are checked for (10, 50) coordinates values to estimate the starting point of a frame. If the starting point of a frame is detected means new frame have been started. The center point of the bounding box is computed and the related object attributes (center point and the XY-coordinates of the bounding box)are transmitted to the hand recognizing process and the entire attributes are reset back to default value. On the other hand, if no new frame detected, the pixel data belonging to the current frame is sorted into data structure (container). The attributes of the container are updated if the adjacency check ofcontainer is fulfilled. This process will be repeated until new frame starts. The computation of the bounding box's center point is shown in equation1&2 [12].

$$X_C = \frac{X_{MAX} + X_{MIN}}{2} \quad (1)$$

$$Y_C = \frac{Y_{MAX} + Y_{MIN}}{2} \quad (2)$$

Bounding box module sends out its outline max_X,max_Y position, min_X and min_Y position while center of mass provides position in term of X_position and Y_position applying in gesture sign classification module. Calculation is carried out to determine which direction the user's hand is pointing. The longest distance between the point of Center-Of-Mass and the bounding box outline is used to classify user's hand. Referring to example inFig. 3, the green dot indicates Center-Of-Mass while white line around red pixel is the bounding box outline and the white box in the middle indicates its center point. Before any computation process is carried out, no direction signal will be sending out to the hardware appliance.. Size of the bounding box will first attempt to check if any user's hand was detected by the camera. If the size of bounding box is smaller than 5 pixels square, system will consider as no input is detected and no further computation process will be

run to analysis user's input. Else, if the bounding box size is greater than 5 pixels square, computation of the longest distance between center of mass and bounding box outline will be carried out. Size of the bounding box can be calculated using Equation 5.

X_size=Xmax(bounding_box_outline)–Xmin
Y_size=Ymax(bounding_box_outline)–Ymin
(5)

The distance between center of mass and bounding box is calculated and compared to each other in four directions which are left, right, up and down. Equation 6 shows the approach to calculate the distances mentioned.

Y1 = Y_position (center of mass) - Ymin (boundingbox outline)
Y2 = Ymax (bounding box outline) - Y_position(center of mass)
X1 = X_position (center of mass) - Xmin (boundingbox outline)
X1 = Xmax (bounding box outline) - X_position(center of mass) (6)

Human hands can show different gestures depending upon the shape of the hand pattern. The different hand gestures can be distinguished on the basis of shape based features. In this section, four different shape based features viz. area, perimeter, thumb detection and radial profile of hand are described; and the proposed criterion for classification for hand gestures is defined. Shape based recognition requires very less computation effort and saves a lot of FPGA area as compared to other approaches. But the approach is not very popular as a hand can assume many shapes. So, in the proposed hand gesture recognition system four different shape based features are calculated rather than relying on a single feature as shown in the flowchart below.

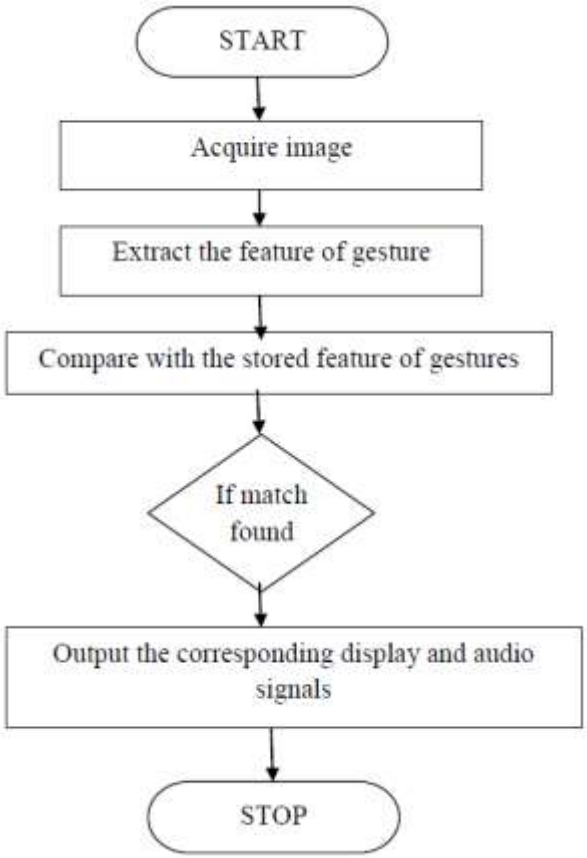


***Fig3. Flow chart of gesture recognition***

## V. CONCLUSION

The hand gesture recognizer system presented in this paper is suitable for controlling hardware appliance in real-time. The proposed gesture interface system was implemented using the Xilinx Development Education FPGA board to speed up the statistical computing or computer vision problems. The system is able to track and detect object bounded by bounding using the center of mass algorithm. Through features extraction from the result obtained, this system could be used in gesture recognition base hardware controller application to bridge the gap between the user and traditional physical hardware devices in real time.

## REFERENCES

[1] Jimeno, Puerta, "State of the Art of the VirtualReality Applied to Desing and ManufacturingProcesses", Department of ComputerTechnology and Computation. AlicanteUniversity. Pg1-15.

[2] Chung Chongyoon and Esa Rantanen, "GestureInteraction With In-Vehicle Audio and

ClimateControls", Rochester Institute of Technology,(11 May 2011).

[3] Joe Ballerini, Tiffany Ng, Alex Garcia andJeremy, "SudoGlove-Hardware Control UsingHand Gesture", (7 May 2011).

[4] Alexander Bochem, Rainer Herpers andKenneth B. Kent, " Acceleration of BlobDetection within Images in Hardware",University of New Brunswick, Dec 15, 2009,pp 1-37.

[5] Paul Viola, Micheal Jones (2001), "RapidObject Detection using a Boosted Cascade ofSimple Features", Accepted Conference onComputer Vision and Pattern Recognition.

[6] Qing Chen, Nicolas D. Georganas, Emil M.Petriu (2007), "Real-time Vision-based HandGesture Recognition Using Haar-likeFeatures", Instrumentation and MeasurementTechnology Conference.

[7] Shahzad Malik, "Real-time Hand Tracking andFinger Tracking for Interaction", December 18,2003, pp1-21.

[8] Takashi Saegusa, Tsutomo Maruyama andYoshiki Yamaguchi, "How Fast is an FPGA inimage processing", System and InformationEngineering, University of Tsukuba, pp1-6.

[9]Hand GesturecontrolledMusic Player: 2009 Innovate Asia The FPGA World competition.098.htm (26 February 2011).

[10] Color Thresholding, [Online] Available:http://www.roborealm.com/help/Color_Threshold .php (8 May 2011).

[11] C.T. Johnston, D.G. Bailey,. K.T. Gribbon"Optimisation of a colour segmentation andtracking algorithm for real-time FPGAimplementation". Institute of Information Sciences and Technology, Massey University,Palmerston North. Pg422-427.

[12] Alexander Bochem, Rainer Herpers andKenneth B. Kent, "Acceleration of BlobDetection in a Video Stream using Hardware",University of New Brunswick, May 26, 2010,pp 1-31.

[13] H. J. Kim, J. S. Lee, J. H. Park, "Dynamic hand gesture recognition using a CNN model with 3Dreceptive fields", In Proceeding of the International Conference on Neural Network and SignalProcessing, pp. 14-19, 2008.

[14] Wen-Hui Chen, Yu-Hong Lin, Shun-Jie Yang, "A generic framework for the design of visualbasedgesture control interface", The 5th IEEE Conference on Industrial Electronics andApplications (ICIEA), pp. 1522-1525, 2010.

[15] Yu Shi, Timothy Tsui, "An FPGA-Based Smart Camera for Gesture Recognition in HCIApplications", In Proceedings of the 8th Asian conference on Computer vision, pp. 718-727, 2007.

[16] Mainak Sen, Ivan Corretjer, Fiorella Haim, Sankalita Saha, Jason Schlessman, Shuvra S.Bhattacharyya, Wayne Wolf, "Computer Vision on FPGAs: Design Methodology and its Application to Gesture Recognition", In Proceedings of the IEEE Computer Society Conferenceon Computer Vision and Pattern Recognition, pp. 133-141, 2005.

[17] Y. L. Chen, C. Y. Chiang, W. Y. Liang, C. H. Chuang, "Embedded Vision-based Nighttime DriverAssistance System", Journal of Convergence Information Technology(JCIT), vol. 6, no. 2, pp.283-292, 2011.

[18] Micrium Corporate, µC/OS-II Kernel, 2011. http://micrium.com/page/products/rtos/os-ii

[19] Altera Corporation, Embedded Design Handbook, 2011. http://www.altera.com/literature/litnio2.jsp