

Real Time Hand Gesture Recognition for Human Machine Communication Using ARM Cortex A-8

¹Vignesh. S PG Scholar, ²Mr.P.Saravanan.

M.Tech Asst.Professor

^{1,2}*Electrial and Electronics Engineering-M.E Embedded Sys Tech, Ganadipathy Tulsi's Jain Engineering college/Anna University, India.*

Abstract: *A novel method proposes for human machine communication using ARM Cortex A-8 processor. Gesture is a form of non-verbal communication in which visible bodily actions communicate particular messages. A novel method proposes for human-machine communication via gesture. The real time system employs a USB web camera, Beagle board XM and HDMI Monitor. Web cam for capturing sequence of image to handle image recognition. Haar classifier used for object recognition was in real time gesture detector and working with only image intensities made the task. Beagle board XM for act as mini CPU which is interface to monitor. The board consists of ARM cortex-A8 Processor which is take real time video for capture gesture image to control mouse moment and implement its functions. Haar classifier there not continuous and not differentiable one to another signal.*

Keyword: *Real time hand gesture, Image recognition, Skin test, k^{th} nearest neighbor algorithm, Beagle board XM, USB web camera, Haar classifiers, Real time operating system.*

I. Introduction

Human computer interaction is one of the important area of research were people try to improve the computer technology. Now days we find smaller and smaller devices being used to improve technology. Hand gesture and object recognition is another area of research. A simple interface like embedded keyboard, folder-keyboard and mini keyboard already exists in today's market. However there interface need some amount of space to use and cannot be used while moving. Touch screen are also globally used which are good control interface and are being used in many applications. However, touch screen cannot be applied to desktop system because of cost and other hardware limitations. By applying vision technology[1], colored substance, image comparison technology and controlling the mouse by natural hand gesture[2] using Beagle board XM act as mini PC. Haar classifiers which are used for filter, noise reduce in an image and gesture image detector [3].

Different type of gestures stored in a database, computer vision [4] can estimate each gesture [5]. The gesture image capture from real time USB web camera to calculate pixel, scaling of image [6] and labeling each gesture [7]. Implements gesture recognition[8] based on hausdroff distance but is not real time, especially when implemented on dedicated system such as Beagle board XM. In many cases special gloves or markers have been used for efficient detection [1] and tracking which constraints the users and other vases demand a very simple and plain background [9] [10]. In a very real time method of hand gesture recognition[12] [15]based on convexity defect is presented but an analysis based only on convexity defects is dependent on smoothness of background [6].

Image recognition is the process of identifying and detecting an object or a feature in a digital image or video[10]. This concept is used in many applications like systems for factory automation, toll booth monitoring, and security surveillance. Typical post[11] image recognition algorithms include: Optical character recognition, Pattern and gradient matching, Face or Hand recognition, License plate matching, Scene change detection. This can implement using mat lab is familiar for image recognition.

A different gesture[13] templates, finding the best match has been, hand finding convexities and Hu moment matching are implement. 'Mouse less' is an invisible computer mouse that provides the familiarity of interaction of a physical mouse without actually needing a real hardware mouse. Despite the advances in computing hardware technologies, the two-button computer mouse has remained the predominant means to interact with a computer. The Mouse less invention removes the requirement of having a physical mouse altogether but still provides the intuitive interaction of a physical mouse that users are familiar with

II. System Block Diagram

A predefined gesture data are stored in a board which is Database. A real time USB camera capturing real time gesture image and pass into Beagle board XM fig.1. The board can covert image into frame and

compare the image data's execute recognition. Database and real time gesture image data these two data's are comparing and image recognition when two data are same the corresponding mouse functions can be execute. If gesture data are mismatch the function doesn't execute in correct manner with respective time.

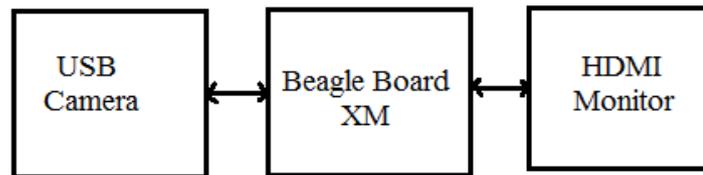


Fig. 1 System Block Diagram

In this we are design a portable computer with arm cortex A-8 processor. The gesture is recognized by using Haar classifiers. A pre defined hand gesture is stored to control the mouse movement by hand. Real time input image capture from USB web camera device send it to beagle board XM to operate image recognition and act for corresponding mouse operation execution. The gesture positive and negative sample taken for mouse cursor movement, right click and left click. Through hand movement we can control the mouse action without using the mouse. This concept is called hand recognition.

III. Operating System

Ubuntu is an operating system based on the Linux kernel and the Linux distribution Debian, with Unity as its default desktop environment. Ubuntu, it can be RTOS and it has been desktop environment. It is distributed as free and open source software. Ubuntu is composed of many software packages, the majority of which are distributed under a free software license. The main license used is the GNU General Public License (GNU GPL) which, along with the GNU Lesser General Public License (GNU LGPL), explicitly declares that users are free to run, copy, distribute, study, change, develop and improve the software. On the other hand, there is also proprietary software available that can run on Ubuntu. Ubuntu operating system is straightforward to use- Ubuntu OS is documented well and searches are performed fast and simply. With ubuntu, one may be able to produce a dual boot while not encountering abundant problems. With ubuntu, creating a partition and then sharing with different operating system is easy. Ubuntu operating system has the integrated software upgrade tool which operates smoothly within the background and updates the system plus all the installed applications. Ubuntu additionally has the "straightforward ubuntu" that installs each application from flash browser plug-in to video drivers among other applications.

IV. Hardware Description

BeagleBoard-xM act as mini CPU, delivers extra ARM Cortex™ -A8 MHz and extra memory with 512MB of low-power DDR RAM, enabling engineers, innovators, and hobbyists to take their imagination to another level. The BeagleBoard-xM features an open hardware design that improves upon the laptop-like performance and expandability of the original BeagleBoard while keeping at hand-held power levels. Direct connectivity is supported by the on-board four-port hub with 10/100 Ethernet while maintaining a tiny 3.25" x 3.25" footprint. The BeagleBoard-xM is intended as a community-supported platform that can be used as the basis for building more complete development systems and as a target for community software baselines.

USB webcam is a device used to capture images, audio and videos. Webcam are either inbuilt or can be externally attached with the laptop or the computer. The most common application of a webcam is video chat, video recording, image capturing. A webcam generally consist of a lens, an image and sound sensors, electronic circuitry to process the data and send it to PC. A variety of webcams with different features are available in the markets. Depending on the features and quality, the technologies used and the prices vary. However the basic function remains the same.

Camera Specifications: Optical lens with CMOS sensor, 25 Megapixel (Interpolated), Frame rate up to 30fps, Video resolution: 320x240, 640x480, Manual switch for LED, 4 built-in LEDs, USB 2.0 interface, Support Windows XP SP2/VISTA/7, Support LINUX Kernel 2.6.27.7 version. A camera is an optical instrument that records images that can be stored directly, transmitted to another location, or both. These images may be still photographs or moving images such as videos. A device for recording visual images in the form of photographs, or video signals.

Cameras utilize the same basic design: light enters an enclosed box through a converging lens and an image is recorded on a light-sensitive medium. A shutter mechanism controls the length of time that light can enter the camera. Most photographic cameras have functions which allow a person to view the scene to be recorded, allow for a desired part of the scene to be in focus, and to control the exposure so that it is not too

bright or too dim. A monitor has four components as shown below: initialization, private data, monitor procedures, and monitor entry queue. The initialization component contains the code that is used exactly once when the monitor is created, The private data section contains all private data, including private procedures, that can only be used within the monitor. Thus, these private items are not visible from outside of the monitor. The monitor procedures are procedures that can be called from outside of the monitor. The monitor entry queue contains all threads that called monitor procedures but have not been granted permissions

V. Simulation Results

A database is an organized collection of data. The data are typically organized to model relevant aspects of reality in a way that supports processes requiring this information. Formally, collection of data and stored in a folder for refer model data. the term "database" refers to the data itself and supporting data structures. Databases are created to operate large quantities of information by inputting, storing, retrieving, and managing that information. Image recognition is the process of identifying and detecting an object or a feature in a digital image or video. This concept is used in many applications like systems for factory automation, toll booth monitoring, and security surveillance. Image recognition have steps to implement correct match database and input image. They are Input image, Skin threshold image, Final image hand region, Extracted hand region and K^{th} Nearest Neighbor image. This process are execute correct matching of image recognition.

The input image fig.2 is the image on which your perform the search using the models in the database. Input image from camera reads a grayscale or color image from the file specified by the string filename. Skin detection is the process of finding skin-colored pixels and regions in an image or a video. This process is typically used as a preprocessing step to find regions that potentially have human faces and limbs in images. Several computer vision approaches have been developed for skin detection.



Fig. 2 Input Image

A skin detector typically transforms a given pixel into an appropriate color space and then use a skin classifier to label the pixel whether it is a skin or a non-skin pixel. A skin classifier defines a decision boundary of the skin color class in the color space based on a training database of skin-colored pixels. Final image hand region Image are finally produce and selecting contour region of hand which is used for execute the operation. Extracted hand region fig.3 portion of a matrix can be extracted and stored in a smaller matrix by specifying the names of both matrices and the rows and columns to extract. k^{th} nearest neighbor in pattern recognition, the k -nearest neighbor algorithm is a non-parametric method for classifying objects based on closest training examples in the feature space.

Fig.4 k -NN is a type of instance-based learning, or lazy learning where the function is only approximated locally and all computation is deferred until classification. The k -nearest neighbor algorithm is amongst the simplest of all machine learning algorithms an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its k nearest neighbors.

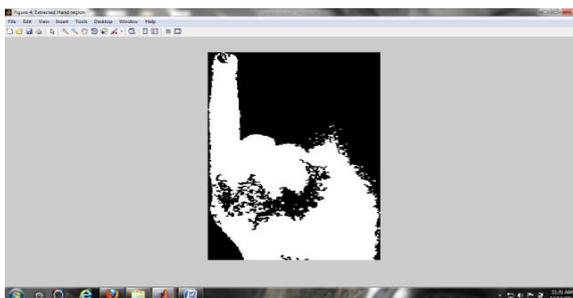


Fig.3 Extracted hand region

When the input data to an algorithm is too large to be processed and it is suspected to be notoriously redundant (e.g. the same measurement in both feet and meters) then the input data will be transformed into a reduced representation set of features (also named features vector). Transforming the input data into the set of features is called Feature extraction. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input. Feature extraction is performed on raw data prior to applying K-NN algorithm on the transformed data in Feature space.

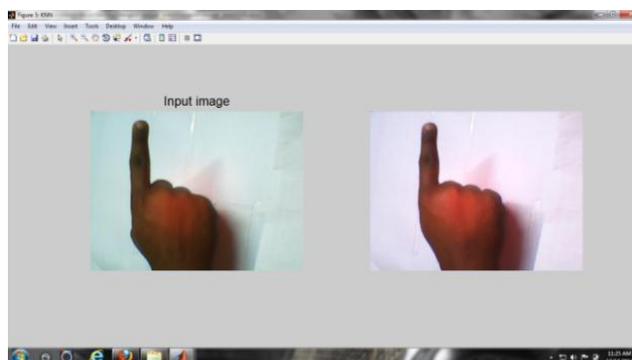


Fig.4 Kth Nearest Neighbor

Mouse cursor movement produce changes the cursor position fig.5 Sequences allow to move the cursor around the full screen using through gesture. Fig.7 Left click is usually a primary action of the mouse. By default in many operating systems, selects an object executes or opens the object. Its reduced time needed to complete the action. One-click phrase has also been used to apply to the commercial field as a competitive advantage. Right click usually a secondary action of the mouse. By default selecting operation are right click.

After recognition gesture right click operation are execute. Reduce time needed to complete the action. Skin test fig.6 RGB color space is the most commonly used color space in digital images. It encodes colors as an additive combination of three primary colors: red(R), green (G) and blue (B). RGB Color space is often visualized as a 3D cube where R, G and B are the three perpendicular axes. One main advantage of the RGB space is its simplicity. However, it is not perceptually uniform, which means distances in the RGB space do not linearly correspond to human perception. In addition, RGB color space does not separate luminance and chrominance, and the R,G, and B components are highly correlated.

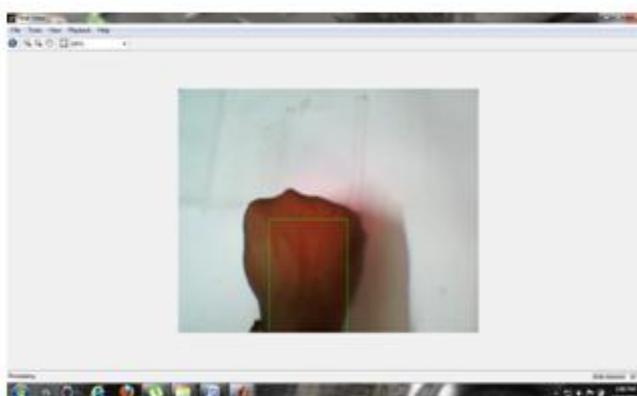


Fig.5 Cursor Movement

The luminance of a given RGB pixel is a linear combination of the R, G, and B values. Therefore, changing the luminance of a given skin patch affects all the R, G, and B components. In other words, the location of a given skin patch in the RGB color cube will change based on the intensity of the illumination under which such patch was imaged! This results in a much stretched skin color cluster in the RGB color cube. skin patches from images of Asian people taken at arbitrary random illumination are plotted in the RGB space. Skin detection is the process of finding skin-colored pixels and regions in an image or a video. This process is typically used as a preprocessing step to find regions that potentially have human faces and limbs in images. Several computer vision approaches have been developed for skin detection. A skin detector typically transforms a given pixel into an appropriate color space and

ways of interaction for human and machine. In future we can implement accuracy and efficiency for add more gestures for different operation

References

- [1]. Erol A, Bebis G, Nicolescu M, et al. Vision-based hand pose estimation: A review. *Computer Vision and Image Understanding*, 2007, 108(1-2): 52-73. www.elsevier.com/locate/cviu.
- [2]. T. F. William and D. W. Craig, "Television Control by Hand Gestures," in *IEEE Intl. Workshop on Automatic Face and Gesture Recognition*, Zurich, June 1995.
- [3]. Wang F, Ngo C W, Pong T C. Simulating a smart board by real-time gesture detection in lecture videos. *IEEE Transactions on Multimedia*, 2008, 10(5): 926-935.
- [4]. Juang C F, Chang C M, Wu J R, et al. Computer vision-based human body segmentation and posture estimation. *Systems, Man and Cybernetics, Part A: Systems and Humans*, *IEEE Transactions on*, 2009, 39(1): 119-133.
- [5]. Okatani T, Deguchi K. Auto calibration of a projector camera system. *Pattern Analysis and Machine Intelligence*, *IEEE Transactions on*, 2005, 27(12): 1845-1855.
- [6]. Lowe D G. Distinctive image features from scale-invariant key points. *International Journal of Computer Vision*, 2004, 60(2): 91-110.
- [7]. Chang F, Chen C J, Lu C J. A linear-time component-labeling algorithm using contour tracing technique. *Computer Vision and Image Understanding*, 2004, 93(2): 206-220. www.elsevier.com/locate/cviu.
- [8]. A. Bobick and A. Wilson, "A state-based technique for the
- [9]. summarization and recognition of gesture," in *Proc. IEEE Fifth Int. Conf. on Computer Vision*, Cambridge, pp. 382-388, 1995.
- [10]. Wang F, Ngo C W, Pong T C. Lecture video enhancement and editing by integrating posture, gesture, and text. *IEEE Transactions on Multimedia*, 2007, 9(2): 397-409.
- [11]. Lee H K, Kim J H. An HMM-based threshold model approach for gesture recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1999, 21(10): 961-973.
- [12]. C. Manresa , J. Varona , R. Mas, and F. J. Perales "Real -Time Hand Tracking and Gesture Recognition For Human -Computer Interaction," in *Electronic Letters on Computer Vision and Image Analysis* 0(0), pp.1-7, 2000.
- [13]. 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*
- [14]. A. VanDam, "Post-WIMP user interfaces," in *Communications of the ACM*, vol 40, pp. 63-67, February 1997
- [15]. E. Hunter, J. Schlenzig, and R. Jain. "Posture Estimation in
- [16]. Reduced-Model Gesture Input Systems," in *Proc. Int'l Workshop Automatic Face and Gesture Recognition*, pp. 296-301, 1995.
- [17]. C. Maggioni. *Gesturecomputer*. "New Ways of Operating a Computer," in *Proc. Int'l Workshop Automatic Face and Gesture Recognition*, 1995.