International Journal of Computational Engineering Research||Vol, 03||Issue, 4||

Survey Paper Based On Hand Gesture Hex Color Matrix Vector

[1,]Sukrit Mehra, [2,]Prashant Verma , [3,]Harshit Bung , [4,]Deepak Bairagee
[1234]*Swami Vivekanad College Of Engineering, Indore*

Abstract

In order to enable a more natural communication with reality systems, automatic hand gesture recognition appears as a suitable means. Hand gesture recognition making use of digital images has been a research topic for many years. The aim of this paper is the proposal of real time vision system for its application within visual interaction environments through hand gesture recognition, using general purpose low cost software, so any user could make use of it in his office or home. The basis of our approach is a fast segmentation process to obtain the moving hand from the whole image, which is able to deal with a large number of hand shapes against different background and lighting conditions. The most important part of the recognition process is robust shape comparison carried out through Hidden Markov Model approach, which operates on edge maps. The visual memory use allows the system to handle variation within the gestures.

Keyword: *Man-Machine Interaction, Image Processing, Segmentation, Hand gesture recognition, Hidden Markov Model, Visual Memory, Range camera.*

I. INTRODUCTION

Hand gesture recognition is an important research issue in the field of Human Interaction with Computer, because of its extensive applications in virtual reality, sign language recognition, and computer games. Despite lots of previous work, building a robust hand gesture recognition system that is applicable for real-life applications.Body language is an important way of communication among humans. Gesture recognition is the mathematical interpretation of a human motion by a computing device. Gesture recognition, along with facial recognition, voice recognition, eye tracking and lip movement recognition are components of what developers refer to as a perceptual user interface (PUI)**[1]**. The goal of PUI is to enhance the efficiency and ease of use for the underlying logical design of a stored program, a design discipline known as usability. In personal computing, gestures are most often used for input commands. Recognizing gestures as input allows

computers to be more accessible for the physically-impaired and makes interaction more natural in a gaming or 3-D virtual environment.Body language is an important way of communication among humans, adding emphasis to voice messages or even being a complete message by itself. Thus, automatic posture recognition systems could be used for improving human-machine interaction. This kind of human-machine interfaces would allow a human user to control remotely through hand postures a wide variety of devices. Different applications have been suggested. Interactions between human and computer are currently**[2]** performed using keyboards, mice or different hectic devices. In addition to being different from our natural way of interacting. In gesture recognition technology, a camera reads the movements of the human body and communicates the data to a computer that uses the gestures as input to control devices or applications.

II. PROBLEM DOMAIN

The main obstacle in achieving a natural interaction between man and machine based on gestures is the lack of appropriate methods of recognition and interpretation of the gestures by the computer. Different methods can be divided into two main classes:

2.1 Vision Based

a) CAMSHIFT Algorithm Limitation

There are quite a few parameters: the number of histogram bins, the minimum saturation, minimum and maximum intensity, and the width-to-height ratio for faces **[3]**. There's also a parameter for enlarging the face region while doing Mean Shift to increase the chances of finding the maximum for skin-probability density.

b) Time Flight Camera Limitation

Background light

Although most of the background light coming from artificial lighting or the sun is suppressed, the pixel still has to provide a high dynamic range. The background light also generates electrons, which have to be stored. For example, the illumination units in today's TOF cameras **[4] [5]** can provide an illumination level of about 1 watt. The Sun has an illumination power of about 50 watts per square meter after the optical bandpass filter. Therefore, if the illuminated scene has a size of 1 square meter, the light from the sun is 50 times stronger than the modulated signal**.**

Interference

If several time-of-flight cameras are running at the same time, the cameras may disturb each others' measurements. There exist several possibilities for dealing with this problem:

- Time multiplexing: A control system starts the measurement of the individual cameras consecutively, so that only one illumination unit is active at a time.
- Different modulation frequencies: If the cameras modulate their light with different modulation frequencies, their light is collected in the other systems only as background illumination but does not disturb the distance measurement.

Multiple reflections

In contrast to laser scanning systems, where only a single point is illuminated at once, the time-of-flight cameras illuminate a whole scene. Due to multiple reflections, the light may reach the objects along several paths and therefore, the measured distance may be greater than the true distance.

c) Naïve Bayes' Classifier Limitation

Assumption of class conditional independence usually does not hold Dependencies among these cannot be modeled by Naïve Bayesian Classifier **[6].**

2.2 Data Glove Based Limitation

[1] **Cost:** Even though glove-based technology has come down in price (under $500 for the 5DT Glove), the cost of robust and complex posture and gesture recognition is going to be high if a glove-based solution is used. The cost of a tracking device and a robust glove is in the thousands of dollars. On the other hand, a vision-based solution is relatively inexpensive, especially since modern-day workstations are equipped with cameras.

[2] **User Comfort:** With a glove-based solution, the user must wear a tracking device and glove that are connected to a computer. Putting these devices on takes time, can be quite cumbersome, and can limit one's range of motion. With a vision-based solution, the user may have to wear a glove, but the glove will be extremely lightweight, easy to put on, and not connected to the computer. Applications, in which no gloves are used, give the user complete freedom of motion and provides a cleaner way to interact and perform posture and gesture recognition.

[3] **Hand Size:** Human hands vary in shape and size. This is a significant problem with glove-based solutions: some users cannot wear these input devices because their hands are too big or too small.

[4] **Hand Anatomy:** Glove-based input devices may not always fit well enough to prevent their position sensors from moving relative to the joints the sensors are trying to measure. This problem reduces recognition accuracy after extended periods of use and forces users to recalibrate the devices which can be a nuisance.

[5] **Accuracy:** In both vision- and glove-based solutions for hand posture and gesture recognition, accuracy is one of the most critical components to providing robust recognition. Both these solutions provide the potential for high levels of accuracy depending on the technology and recognition algorithms used. Accuracy also depends on the complexity and quantity of the postures and gestures to be recognized. Obviously, the quantity of possible postures and gestures and their complexity greatly affect accuracy no matter what raw data collection system is used.

[6] **Calibration:** Calibration is important in both vision- and glove-based solutions but, due to the anatomy of the hand, it is more critical with glove-based solutions. In general, a calibration procedure or step is required for every user and, in some cases, every time a user wants to run the system.

[7] **Portability:** In many applications, especially gesture to speech systems, freedom from being tied down to a workstation is important. With glove-based solutions, this freedom is generally available as long as hand tracking is not involved, since these input devices can be plugged right into a laptop computer. Vision-based solutions were originally quite difficult to use in a mobile environment due to camera placement issues and computing power requirements.

[8] **Noise:** In glove-based solutions where hand tracking is required, some type of noise is bound to be associated with the data (it can come from a variety of sources depending on the tracking technology used). Filtering algorithms are therefore necessary to reduce noise and jitter. In some cases this can get computationally expensive when predictive techniques such as Kalman filtering are used.

III. PROPOSED SOLUTION

A low cost computer vision system that can be executed in a common PC equipped with USB web cam is one of the main objectives of our approach. The system should be able to work under different degrees of scene background complexity and illumine ion conditions.

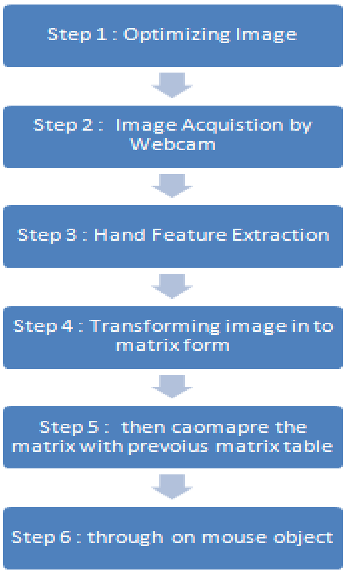

Figure 1: Proposed System Framework

The purposed technique is depending on the following approach:

- A pre image of hand is stored in the template.
- Capture the hand at initial point through frame.
- Match the captured hand with the pre hand stored in the template and marked or set the captured image as a reference image.
- Take the video frame of hand through webcam and convert the YUV color format to RGB color format.
- The RGB formatted image is saved in the form of matrix.
- Taking continues frame and match with the RGB matrix, the number of changes in the matrix is know and the movement of hand can be recognized.

3.1 Method Used: The Hidden Markov model is a stochastic process built on the top of another stochastic process, the Markov process. A time domain process exhibits first-order Markov property if the conditional probability density of the current event, given all past and present events, depends only on the most recent events. In a HMM **[7]**, the

3.2Using YUV Colorspace:YUV colorspace is a bit unusual. The Y component determines the brightness of the color (referred to as luminance or luma), while the U and V components determine the color itself (the chroma) [9]. Y ranges from 0 to 1 (or 0 to 255 in digital formats), while U and V range from -0.5 to 0.5 (or -128 to 127 in signed digital form, or 0 to 255 in unsigned form). Some standards further limit the ranges so the out-of-bounds values indicate special information like synchronization.

Figure2: YUV Image on Y-axis

These images show us rods at various points on the UV plane, extending through Y. This allows us to see how each UV point's color is changed as the Y value is increased or decreased.

3.3 YUV – RGB Conversion:

There are many slightly different formulas to convert between YUV and RGB. The only major difference is a few decimal places. The CCIR 601 Standard **[10]** specifies the correct coefficients. Since I'm lazy and haven't looked up this spec, I don't know if the following coefficients are correct or not. In any event, I've used them for many conversions with no obvious discoloration.

These formulas assume U and V are unsigned bytes.

R = Y + 1.4075 * (V - 128)
G = Y - 0.3455 * (U - 128) - (0.7169 * (V - 128))
B = Y + 1.7790 * (U - 128)

Y = R * .299000 + G * .587000 + B * .114000
U = R * -.168736 + G * -.331264 + B * .500000 + 128
V = R * .500000 + G * -.418688 + B * -.081312 + 128

JPEG/JFIF - RGB Conversion:

JPEG/JFIF files store compressed images in a YUV-like colorspace **[11][12]** that uses slightly different coefficients to convert to RGB. These formulas are:

R = Y + 1.40200 * (U - 128)
G = Y - 0.34414 * (V - 128) - 0.71414 * (U - 128)
B = Y + 1.77200 * (V - 128)

IV. CONCLUSION

The mouse works on 120dpi and the frame captured through webcam is 30-25 fps. Using HMM with color hex method we would in increase the accuracy mouse movement interface and the process of the image in the form of vector i.e. in the form of matrix would give an 60% optimum result.

REFERENCES

[1] Zhou Ren, Junsong Yuan, Zhengyou Zhang, "Robust Hand Gesture With Kinetic Sensor" , Proceeding Of The 19th ACM International Conference On Multimedia, November 28-December 01, 2011 Scottsdale, Arizona USA, Pp. 759–760. [2] Fukunaga, Keinosuke; Larry D. Hostetler (January 1975). "The Estimation Of The Gradient Of A Density Function, With Applications In Pattern Recognition". IEEE Transactions On Information Theory(IEEE) 21 (1): 32–40. Doi:10.1109/TIT.1975.1055330. Retrieved 2008-02-29.

[3] Vadakkepat, P Et.Al, "Multimodal Approach To Human-Face Detection And Tracking", Industrial Electronics, IEEE Transactions On Issue Date: March 2008, Volume: 55 Issue:3, Pp.1385 – 1393

[4] E. Kollorz, J. Hornegger And A. Barke, "Gesture Recognition With A Time-Of-Flight Camera, Dynamic 3D Imaging", International Journal Of Intelligent Systems Technologies And Applications Issue: Volume 5, Number 3-4 2008, Pp.334 – 343.

[5] B¨ohme, M., Haker, M., Martinetz, T., And Barth, E. (2007) "A Facial Feature Tracker For Humancomputer Interaction Based On 3D TOF Cameras", Dynamic 3D Imaging (Workshop In Conjunction With DAGM 2007).

[6] Pujan Ziaie Et.Al, "Using A Naïve Bayes Classifier Based On K-Nearest Neighbors With Distance Weighting For Static Hand-Gesture Recognition In A Human-Robot Dialog System" Advances In Computer Science And Engineering Communications In Computer And Information Science, 2009, Volume 6, Part 1, Part 8, Pp.308-315.

[7] A. Markov. An Example Of Statistical Investigation In The Text Of Eugene Onyegin, Illustrating Coupling Of Tests In Chains. Proceedings Of The Academy Of Sciences Of St. Petersburg, 1913

[8] L. Baum Et. Al. A Maximization Technique Occuring In The Statistical Analysis Of Probablistic Functions Of Markov Chains. Annals Of Mathematical Statistics, 41:164–171, 1970.

[9] N. Soontranon, S. Aramvith, And T.H. Chalidabhongse. Face And Hands Localization And Tracking For Sign Language Recognition. International Symposium On Commu-Nications And Information Technologies 2004 (ISCIT 2004), Sapporo, Japan, October26-29, 2004

[10] U. Ahlvers, U. Zölzer, R. Rajagopalan. Model-Free Face Detection And Head Tracking With Morphological Hole Mapping, EUSIPCO'05, Antalya, Turkey.

[11] S. Askar, Y. Kondratyuk, K. Elazouzi, P. Kauff, O. Scheer. Vision-Based Skin-Colour Segmentation Of Moving Hands For Real-Time Applications. Proc. Of. 1st Europeanconference On Visual Media Production, CVMP, London, United Kingdom, 2004.

[12] A. Hadid, M. Pietikainen, B. Martinkauppi. Color-Based Face Detection Using Skin Lo- Cus Model And Hierarchical Filtering. Proc. Of 16th International Conference On Patternrecognition 2002, Vol4, Pp. 196-200, 2002