# An American Sign Language Detection System using HSV Color Model and Edge Detection

A. Sharmila Konwar, B. Sagarika Borah, C. Dr.T.Tuithung

***Abstract*---The work presented in this paper is aimed to design an automatic vision based American Sign Language detection system and translation to text. To detect the human skin color from the image, HSV color model is used. Then edge detection is applied to detect the hand shape from the image. A set of morphological operation is applied to get a refined output for the sign language recognition.**

***Index Terms*---Edge detection; HSV; Morphological operation.**

## I. Introduction

SIGN language is the oldest and the natural form of language for communication. Communication is the process of exchange of thoughts and messages in various ways such as speech, signals, behavior and visuals [1]. Deaf and dumb (D&M) people make use of their hands to express different gestures to express their ideas with the other people. Gestures are the nonverbally exchanged messages and these gestures are understood with vision. This nonverbal communication of D&M people is called sign language. It has been adopted as an integral part of our day to day communication process. A sign language is similar to any other language in that is has vocabulary and grammar, but it uses visual modality for information exchange. For interaction between normal people and D&M people a language barrier is created as sign language structure is different from normal text. So they depend on vision based communication for interaction.

F.A.Sharmila Konwar,Student,M.Tech in Information Technology,North Eastern Regional Institute of Science and Technology,Nirjuli,Arunachal Pradesh,India,Email:chansharmi@gmail.

S.B.Sagarika Borah,Student, M.Tech in Computer science and Engineering, ,North Eastern Regional Institute of Science and Technology,Nirjuli,ArunachalPradesh,India,Email:sgrk1017@rediffmail.com

T.C.Dr.T Tuithung,Associate Professor,Computer Science and Engineering,North Eastern Regional Institute of Science and Technology,Nirjuli,ArunachalPradesh,IndiaEmail:t.tuithung@yahoo.com

If there is a common interface for them that converts the sign language to text the gestures can be easily understood by the other people. So research has been made for a vision based interface system where common man and the D&M people can enjoy communication without really knowing each other's language.

The aim is to develop a user friendly human computer interfaces (HCI) where the computer understands the human language. There are various sign languages all over the world, namely American Sign Language (ASL), French Sign Language, British Sign Language (BSL), Japanese Sign Language and work has been done on other languages all around the world. A sign language recognition system can be described as a two phase system.

The first phase is the training phase. A database is created by capturing images either by web camera or digital camera. The steps in this phase are preprocessing, feature extraction and training the classifier.

In the second phase –testing phase, the steps are image acquisition and then preprocessing, feature extraction and classification based on the testing phase.

There are different approaches for image acquisition. The gloved based approach uses electromechanical devices to provide exact hand configuration, and position. But it is expensive and not user friendly. The next approach is the vision based approach. One type of vision based approach uses color gloves for gesture recognition. The next approach uses bare hand as input which is user friendly, but has challenges like complex background, environment variations and presence of other skin color objects with the hand object.

Various sign language recognition systems have been designed till date. Kulkarni [1] used HSV color model for segmentation purpose and features are extracted using histogram technique and Hough algorithm. Feed forward Neural Networks with three layers are used for gesture classification. Wysoski et al. presented rotation invariant postures using boundary histogram. Hyeon-Kyu Lee and Jin H. Kim [2] used HMM (Hidden Markov Model) for classification. Stergiopoulou used YCbCr color space and detected an approximation of hand shape morphology using Self-Growing and Self-Organized Neural Gas (SGONG) network. P.V.V.Kishore, P.Rajesh Kumar, E.Kiran Kumar & S.R.C.Kishore [2] used Sobel edge operatorand DCT (Discrete

978-1-4799-3358-7/14/$31.00 ©2014 IEEE

Cosine Transform) for feature extraction.Ravikiran J, Kavi Mahesh, Suhas Mahishi, Dheeraj R, Sudheender S, Nitin V Pujari [3]proposed a method using canny edge detection and then boundary tracing to recognize the number of fingers open in a gesture.

The other different techniques and tools that have been applied are PCA, histogram orientation, OTSU's segmentation algorithm, Segmentation and Extraction with edge detection for feature extraction and classifier like Finite State Machine (FSM),fuzzy clustering, Genetic Algorithms (Gas).

This project is designed to create a user friendly interface to convert sign language to text and vice versa. The main aim of considering ASL is that it is a static gesture language and it has a standard database. The system is a vision based approach. All the signs are represented with bare hands and so it eliminates the problem of using any artificial devices for interaction. As an experimental approach five ASL hand gestures for the alphabets "A,B,C,D,E" are used for detection and recognition. The images are captured in a static background. The existing approaches usually take pre-captured images of sign language as their inputs. The proposed method is a vision based method using a web cam to capture images as this approach is rarely used in this purpose.

This paper describes a simple methodology for hand gesture detection using skin color detection and edge detection for sign language recognition.

## II. METHODOLOGY

The system designed to detect the input sign language consist of the 6 steps. The algorithm of the proposed methodology is described using the following block diagram in figure 1.

### *A. Image capture*

We use a web camera of 10 megapixels to capture the input images for creation of the database. The RGB images are then resized to 2000x2000 to get a detailed result of the hand shape for further computation.

### *B. Preprocessing*

Pre-processing is a common name for operations with images at the lowest level of abstraction -both input and output are intensity images.Pre-processing improves the image data by suppressing unwanted distortions or enhances some image features important for further processing. As the proposed method is a vision based sign language detection system it is important to check whether the input is a hand or not. The challenges for a vision based hand detection system are varying lighting condition, geometric variation, complex background and skin color detection, so variation in human skin color complexion requires the efficient hand tracking and segmentation for any gesture recognition. From various researches it is found that color is very powerful descriptor for object detection. So for the segmentation purpose we are using color information which is invariant to rotation and geometric variation of the hand [5].

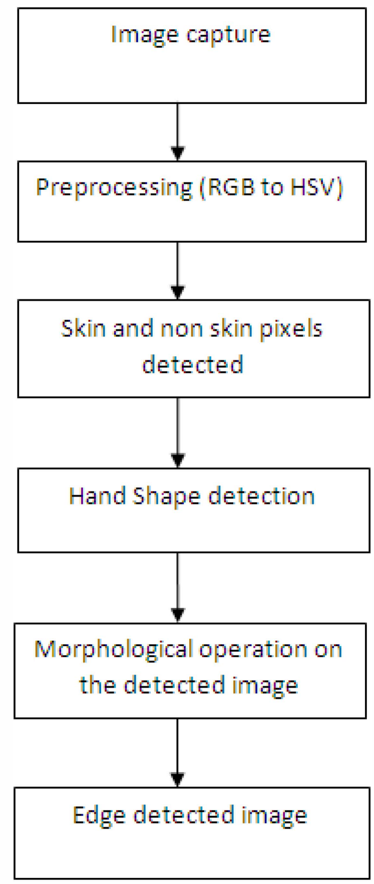


Figure 1.The flow diagram of the sign language detection algorithm

*1) Skin detection:* Skin detection is typically used as preprocessing step to indicate the existence of humans in an image or a video, face detection system, sign language detection system and many other computer vision based methods. A skin detector detects the pixels of the skin portions from the input image and classifies it as skin or a non skin pixel. Then the skin pixels are used for hand shape detection and application for further processing. Human skin color does not fall randomly in a given color space, but it is clustered at a small area in the color space. Forsyth and Fleck [6] described that human skin color has a restricted range of hues and is not deeply saturated, since the appearance of skin is formed by a combination of blood (red) and melanin (brown, yellow). Variety of color spaces has been used in skin detection literature with the aim of finding a color space where the skin color is invariant to illumination conditions and variation of skin tone in different parts of the world. The choice of the color spaces affects the shape of the skin class, which affects the detection process. Color models are useful for to specify a particular color in standard way. It is space-coordinated system within which any specified color represented by single point.

The different color models that are used for skin color detection are:

i) HSI, HSV, HSL - Hue Saturation Intensity (Value, Lightness)
ii) YCrCb
iii) RGB
In the proposed method HSV color space is being used for preprocessing of the sign languages.

*2) HSV color model:* The HSV color model is a perceptual color model. It is called a perceptual model as it represents color in terms that non technical people understand. HSV color model is a deformation of the RGB color space and it can be mapped to RGB color space by a nonlinear transformation of RGB model. HSV color model defines color according to the three basic features of the color:

Hue (H) is the basic feature of color such as: red, yellow, green .It ranges from 0 to 360 in the color space.

Saturation(S) describes the color purity. It indicates the amount of grey in the color space. It ranges between 0 and 100% or 0 to 1. When the value is ’0,’ the color is grey and when the value is ’1,’ the color is a primary color.

Luminance (Value) is the brightness of the color and varies with color saturation. It ranges from 0 to 100%. When the value is ’0′ the color space will be totally black. The color space brightness up and shows various colors with the increase in the value. Alvy Ray Smith promoted this model in 1978[7]. [Skarbek and Koschan 1994] noted that the transformation of RGB to HSV is invariant to highlights at white light sources, and also, to ambient light and surface orientation relative to the light source. This model allows user to specify the boundary of the skin pixels only in terms of hue and saturation. Since I,L or V give brightness information it is normally ignored to reduce illumination dependency of skin color.

*C. Skin pixel detection Algorithm*

We used the HSV color model described in [6] for skin detection. The HSV color model is represented as a hexacone in 3D space. The central vertical axis represents the Intensity [9]. Hue is defined as an angle in the range [0,360] relative to the Red axis with red at angle 0, green at 120, blue at 240 and red again at 360.Saturation is the depth or purity of the color and is measured as a radial distance from the central axis with value between 0 at the center to 1 at the outer surface [8][9]. The saturation is represented as the distance from the center of the circle. Highly saturated colors are on the outer edge of the cone, whereas gray tones (which have no saturation) are at the very center. The brightness is determined by the colors vertical position in the cone.

For determining whether the pixels are skin or non- skin the resized image is first converted from RGB to gray using the matlab function rgb2gray.From the HSV color model a range is calculated to fix a boundary for skin pixels.

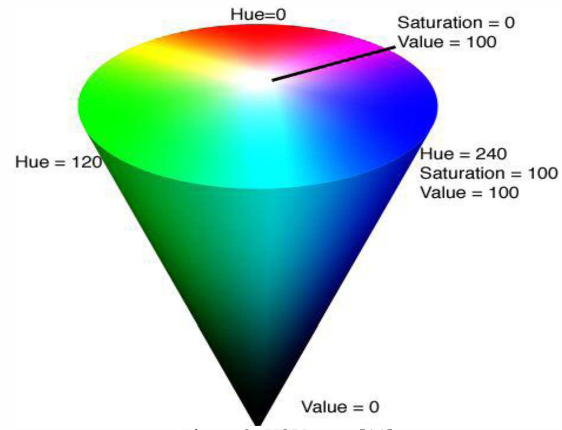

Figure 2: HSV cone [11]

The range of the pixels was calculated from the HSV cone by first estimating the color of skin and then determining the angle. Then this angle was converted to revolutions using the matlab function convang. From the cone it is seen that color of the cone changes from red to yellow within the range approximately 0 to 120.The value greater than 120 is green. So the hue range is assigned to be less than 0.25(120 degree).The saturation range is assigned from.98 to.198 in a similar manner.

*D. Edge detection*

The next step after hand shape detection is to find the exact shape of the hand from the detected skin pixels. For this purpose edge detection is used. Edge is the discontinuity in image intensity from one image pixel to another pixel. An edge detected image can be used as an input for data compression, matching of image, feature extraction. Detection of edges filters out useless data and stores only the necessary information needed for future work under the condition that the important structural properties of the image are not loosed.

For the sign language detection system optimal edge detection technique is canny edge detection technique. This algorithm satisfies various criteria of a good edge detector like including most edges by minimizing the error rate, marking edges closely as possible to the actual edges to maximize localization, and marking edges only once when a single edge exists for minimal response [9].

The canny algorithm follows the following steps:

Step 1:

Gaussian filter is used to filter out any noise in the original image. A suitable mask is calculated, for Gaussian smoothing .With bigger values for σ(Gaussian value),it increases the size of the mask which decreases the accuracy of the detected edges.

Step 2:

A 2-D spatial gradient measurement is performed on the image to find the edge strength by taking the gradient of the filtered image. The 3x3 convolution mask is applied by Sobel operator to find the edge strength at each point of the image. The 3x3 convolution mask estimating the gradient in both x direction and y direction is show below:

| | | |
|---|---|---|
| -1 | 0 | +1 |
| -2 | 0 | +2 |
| -1 | 0 | +1 |

Gx

| | | |
|---|---|---|
| +1 | +2 | +1 |
| 0 | 0 | 0 |
| -1 | -2 | -1 |

Gy

Figure 3: Masks for Sobel operator

The magnitude, or edge strength, of the gradient is then approximated using the formula:

$$|G| = |Gx| + |Gy| \quad (1)$$

Step 3:

The gradient in the x and y direction gives the edge directions. The edge direction is computed by the formula:

$$\text{Theta} = \text{invtan}\left(\frac{Gx}{Gy}\right) \quad (2)$$

Step 4:

Once the edge direction is known, the next step is to relate the edge direction to a direction that can be traced in an image. So if the pixels of a 5x5 image are aligned as follows:

| | | | | |
|---|---|---|---|---|
| x | x | x | x | x |
| x | x | x | x | x |
| x | x | a | x | x |
| x | x | x | x | x |
| x | x | x | x | x |

Let us consider the pixel "a". There are only four possible directions for the surrounding pixels of a. When it is 0 degrees it is in the horizontal direction, 45 degrees means along the positive diagonal), 90 degrees is in the vertical direction, or 135 degrees is along the negative diagonal. Depending on these four directions the edge orientation of any pixels is assigned to the edge it is closest. For e.g.if the orientation angle is 5 degrees, than it is considered as zero degrees).

Step 5:

Non-maximum suppression is applied to remove the unwanted pixels along the edge in the edge direction (sets it equal to 0) .This results in a thin line in the output image.

Step 6:

Due to noise the edge dips below the threshold value and similarly also extend above the threshold value resulting in a dashed line. It is called Streaking. To eliminate this problem hysteresis is applied where two thresholds are considered, high and low. All the pixels with a greater value then the higher threshold value or any pixel connected to that pixel and greater than the low value are assumed to be an edge pixel.

*E. Morphological operation*

Morphological operations apply a structuring element to an input image, creating an output image of the same size [10]. In this project we are using dilation and erosion to get a thin structure of the edge detected binary image. A structuring element is a matrix consisting of 0's and 1's with any arbitrary shape. The number of pixels added or removed from the objects in an image depends on the size and shape of the structuring element used to process the image. The center pixel of the structuring called the origin is responsible for identifying the pixel of interest. The neighborhood is defined by the pixels of the structuring element containing 1.These pixels are also considered in dilation or erosion processing.

*E .Result*

The proposed algorithm is implemented in a number of hand gestures. The images were captured in different environments, and various geometric variations. The color model applied was successful in detecting the skin pixels from the input image. Various hand shapes were captured for each sign and tested. For every alphabet a total 100 samples were considered and 65 images are perfectly detected resulting in 65 % success. The edge detection of the input sign language representation of American Sign Language character "B" is shown step by step In figure 4.

a) Input image which represents sign language "B",
b) Skin detected image, c) Image after edge detection,
d) Image after morphological operation

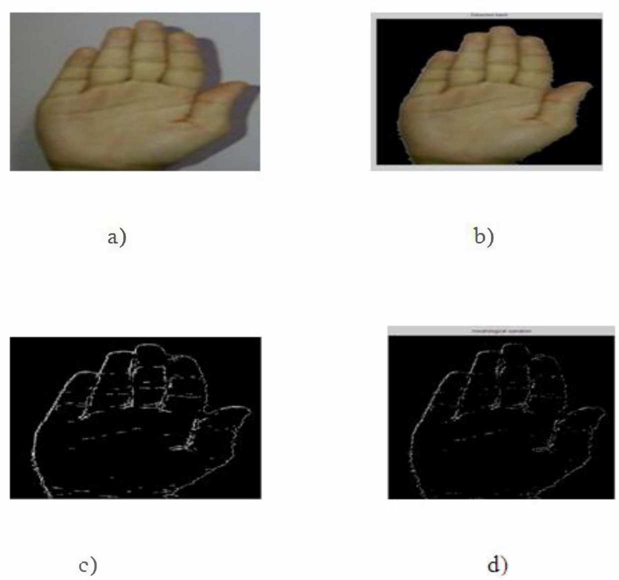


Figure 4: Step wise representation of the Sign language for character "B"

## III .Conclusion

In this paper we have presented an algorithm for hand gesture detection of American Sign Language. Using canny edge detection algorithm the hand gestures are detected successfully for the five alphabets that has been experimented. Some images are not detected successfully due to geometric variations, uneven background and light conditions. After detection of the hand gestures the next step is to extract the features and classify them for recognition. PCA and ANN will be used for feature extraction and recognition respectively.

## REFERENCES


[1] Sawant Pramada, Deshpande Saylee, Nale Pranita, Nerkar Samiksha, Mrs.Archana S. Vaidya ,"Intelligent Sign Language Recognition Using Image Processing" IOSR Journal of Engineering (IOSRJEN) e-ISSN: 2250-3021, p-ISSN: 2278-8719 Vol. 3, Issue 2, ||V2|| PP 45-51.

[2] V. S. Kulkarni, S.D.Lokhande, ,"Appearance Based Recognition of American Sign Language Using Gesture Segmentation", International Journal on Computer Science and Engineering (IJCSE), Vol. 2(3), pp. 560-565, 2010.

[3] Ravikiran J, Kavi Mahesh, Suhas Mahishi, Dheeraj R, Sudheender S, Nitin V Pujari," Finger Detection for Sign Language Recognition" Proceedings of the International Multi Conference of Engineers and Computer Scientists 2009 Vol I,IMECS 2009, March 18 - 20, 2009, Hong Kong.

[4] Jayashree R. Pansare, Shravan H. Gawande, Maya Ingle," Real-Time Static Hand Gesture Recognition for American Sign Language (ASL) in Complex Background" Journal of Signal and Information Processing, 2012, 3, 364-367 doi:10.4236/jsip.2012.33047 Published Online August 2012 (http://www.SciRP.org/journal/jsip)

[5] Fleck, M., Forsyth, D. A., Bregler, C.," Finding naked people". Proc. of the ECCV, vol. 2, 592–602, 1996.

[6] Shamik Sural, Gang Qian and Sakti Pramanik, "Segmentation and histogram generation using the hsv color space for Image retrieval"0-7803-7622-6/02/$17.00 ©2002 IEEE.

[7] Li Shuhua ,Guo Gaizh,"The application of improved HSV color space model in image processing,"978-1-4244-5824-0/$26.00 _c 2010 IEEE V2-10.

[8] Noor Adnan Ibraheem ,RafiqulZaman Khan, "Survey on Various Gesture Recognition Technologies and Techniques" International Journal of Computer Applications (0975 – 8887) Volume 50 – No.7, July 2012 38 .

[9] Canny, J., "A Computational Approach to Edge Detection", IEEE Trans.Pattern Analysis and Machine Intelligence, 8:679-714, November 1986.

[10] Gleb V. Tcheslavski," Morphological Image Processing: Basic Concepts" Spring 2009,ELEN 4304/5365 DIP 1.

[11] Darrin Cardani"Adventures in HSV Space".

[12] Archana S. Ghotkar & Gajanan K. Kharate, "Hand Segmentation Techniques to Hand Gesture Recognition for Natural Human Computer Interaction" International Journal of Human Computer Interaction (IJHCI), Volume (3): Issue (1) : 2012

[13] Yonas Fantahun Admasu,Kumudha Raimond,"Ethiopian Sign Language Recognition Using Artificial Neural Network", 978-1-4244-8136-1/10/$26.00_c 2010 IEEE.

[14] Lungociu,C., "Real time sign language recognition using Artificial neural networks" Studia univ. Babes bolyai, informatica, Volume LVI, Number 4, 2011.

[15] Simei G. Wysoski, Marcus V. Lamar, Susumu Kuroyanagi, Akira Iwata, "A Rotation Invariant Approach On Static-Gesture Recognition Using Boundary Histograms And Neural International Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.4, July 2012173 Networks," IEEE Proceedings of the 9th International Conference on Neural Information Processing, Singapura, 2002.

[16] E. Stergiopoulou, N. Papamarkos., "Hand gesture recognition using a neural network shape fitting technique," Elsevier Engineering Applications of Artificial Intelligence, vol. 22(8), pp. 1141– 1158, doi: 10.1016/j.engappai.2009.03.008,2009.