Comparative Study for Vision Based and Data Based Hand Gesture Recognition Technique

Oinam Robita Chanu[1], Anushree Pillai[2], Spandan Sinha[3], Piyanka Das[4]
Department of Biomedical Engineering
SRM University, Kattankulathur, Chennai, India
Email: rubi.chanu@gmail.com[1],anushreepillai@gmail.com[2], spandan.20.d@gmail.com[3],bmepiya@gmail.com[4]

***Abstract*- Communication for mute people has always been very difficult. Mute people use hand gestures, also known as sign language to communicate with normal people. Hands gestures have their assigned meanings which may differ from person to person and hence cannot be understood by normal people. To overcome this difficulty many vision-based hand gesture recognition systems and data glove based hand gesture recognition systems have been proposed. This paper illustrates about two different techniques of vision-based hand gesture recognition and one data glove based technique. The vision - based techniques are static hand gesture recognition technique and real-time hand gesture recognition technique. In both the techniques, MATLAB software is used for processing input images and no dataset is used for decision making which makes this system more accurate than the existing system designs. In data glove based technique, the glove consists of five flex sensors. The change in resistance of flex sensor is used to recognize the hand gesture. All the three techniques are performed on 10 subjects and compared to find the most accurate technique. It was found that both the vision based techniques showed 100% accuracy in bright lighting condition with a white background while the data glove based technique showed an accuracy of 86%. This which clearly shows that mentioned vision based is more stable and reliable compared to the data glove based technique.**

Keywords*- *Vision-Based Hand Gesture Recognition, Flex Sensors, Data glove based Hand Gesture Recognition, MATLAB software

I. INTRODUCTION

Speech disorder is a communication disorder in which a person's ability to produce sound is disrupted. Speech impairment can happen in the area of articulation, language or voice. People with speech disorder use gestures to communicate with others. The gesture is a physical movement which mute people use to convey their feelings and emotions. A gesture can be static or dynamic. In the last few years, many techniques have been introduced to perform the gesture recognition. Hand gestures are mainly used as sign language for communication. The Hand gesture recognition method is classified into two categories, such as – Data glove based hand gesture recognition and Vision-based hand gesture recognition [1] [6] [7] [8] [9].

In this paper, a comparison is made between the vision based technique and data glove based technique. Two types of vision based techniques are used. The first technique is the static hand gesture technique and the second technique is a real-time algorithm which takes video as an input. The data glove used in this paper is a flex sensor based glove. All the three techniques are performed and compared to find an accurate technique for hand gesture recognition.

II. METHODOLOGY AND RESULT

The proposed model contains three techniques of hand gesture recognition. This is done to find out the accurate method of hand gesture recognition techniques. The first two methods are based on image processing technique and the third technique is based on flex sensors [13].

A. Static Hand Gesture Recognition

System Overview

In the static hand gesture recognition technique, the trained hand gestures are captured using a camera and saved. These images undergo image processing techniques to give a corresponding contextual message and voice output for each finger counts which is shown in figure 1.

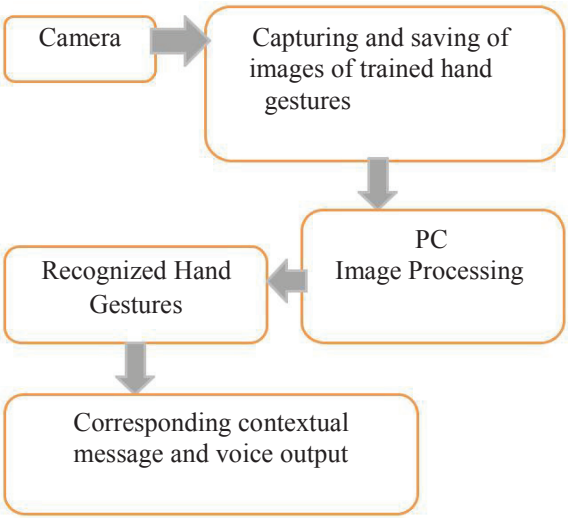


Figure 1: System Overview of static hand gesture recognition technique.

Retrieving the saved images and performing Gray-scale conversion

The saved images are retrieved from the modal dialog box which lists the images saved. The image needed is chosen from the dialog box which undergoes gray scale conversion where a colour 24 bit pixel image is converted to 8 bit pixel

image. Thus the image after gray-scale conversion gives information about intensity by removing the hue and saturation information. The RGB values of the image are converted to gray-scale values by forming a weighted sum of the *R*, *G*, and *B* components [2] [11]:

Gray=0.2989 * R + 0.5870 * G + 0.1140 * B

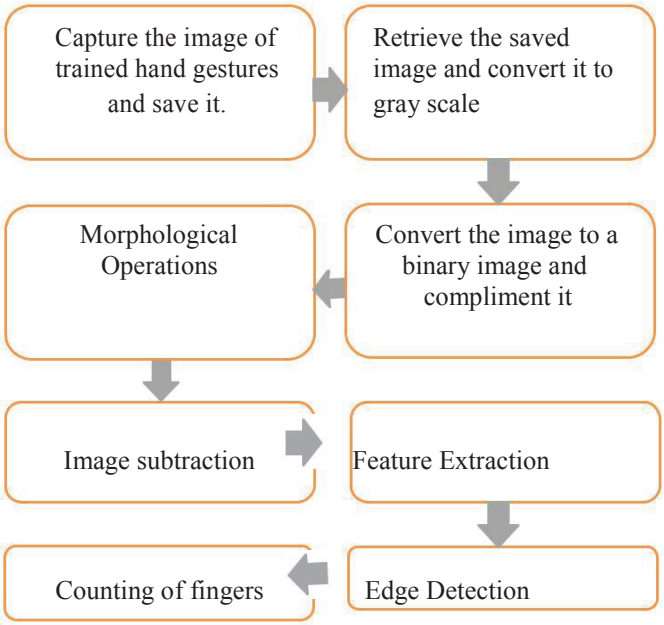


Figure 2: Procedure for static hand gesture recognition technique.

Converting to binary image

The gray image is converted into binary image. The output image replaces all pixels in the input image with luminance greater than level with the value 1 (white) and replaces all other pixels with the value 0 (black). The default level is 0.5. This helps in segmentation of the skin pixels and background pixels. Compliment the image.

Morphological Operations

For detection of the fingers the image is eroded and then dilated. Erosion removes finger pixels only. Dilation is used to get a clear image after the fingers objects are removed from the image. This gives an image of the hand without the fingers. Since the structuring element is a disk shape the formula for erosion and dilation is:

Erosion:

$$(A \ominus B)(x,y) = \min\{A(x+x', y+y') - B(x', y') \mid (x', y') \in DB\},$$

Where, A is the image and B is the structuring element and DE is the domain of the structuring element [4] [10].

Dilation:

$$(A \oplus B)(x, y) = \max\{A(x - x', y - y') - B(x', y') \mid (x', y') \in DB\}\},$$

Where A is the image and B is the structuring element and DE is the domain of the structuring element [3]

Image subtraction

Eroded and dilated image is subtracted from the image binary image which gives an output image that contains only fingers.

Subtracted image = Binary image – Image after morphological operations.

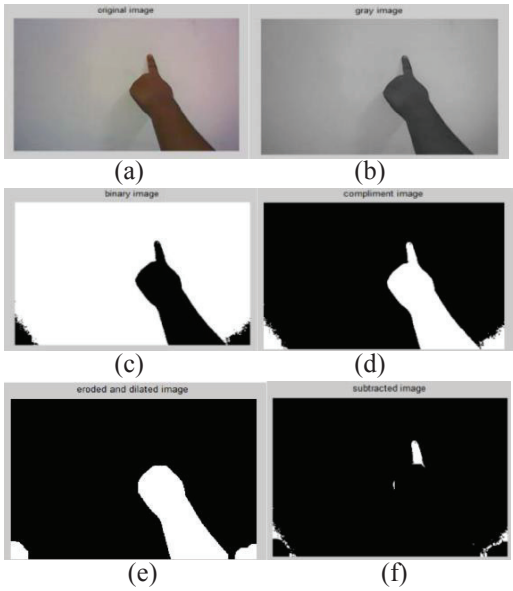


Figure 3: (a) Original Image (b) Gray Image (c) Binary Image (d) Complimented Image(e) Image after erosion and dilation (f) Subtracted Image

Feature extraction

Shape of the finger is detected using feature extraction method. It removes all the redundant pixels (if any shadows) other than the finger pixels thus it gives defined shape of the fingers. Feature extraction is done by removing the white pixels below a threshold and also by removing the pixels in the object connected to the image border. Thus after several process of removing redundant pixels only finger pixels remains in the image.

Figure 4: Several processes of removing redundant pixels **Edge Detection**

Edges of the fingers are detected using canny edge detection method. Canny edge detection is used because it is not affected by noise and can easily detect true weak edges [10]. Canny edge detection has a series of process i.e.

(1) Applying Gaussian filter to smoothen the image.

$$H_{ij} = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{(i-(k+1))^2 + (j-(k+1))^2}{2\sigma^2}\right); 1 \le i, j \le (2k+1)$$

For Gaussian filter kernel of size $(2k+1) \times (2k+1)$ [4] [11].

(2) Finding the intensity gradients

$$\mathbf{G} = \sqrt{\mathbf{G}_x^2 + \mathbf{G}_y^2}$$

$$\mathbf{\Theta} = \mathbf{atan2}(\mathbf{G}_y, \mathbf{G}_x),$$

Where, Gx is the horizontal direction and Gy is the vertical direction [5] [12].

(3) Non maximum suppression

(4) Double threshold

(5) Edge tracking by hysteresis

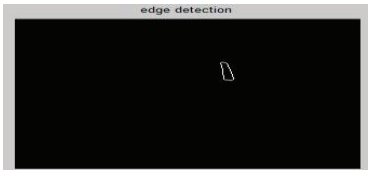


Figure 5: Edge Detection.

Counting of fingers

Change from background pixels to foreground pixels and vice -versa is detected. The number of changes are recorded which gives the count of the fingers. The count of the fingers is used to show and play the corresponding message.

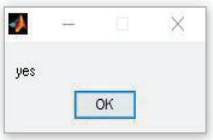


Figure 6: Contextual output for finger count 1

B. Real time Hand Gesture Recognition Technique

The real time hand gesture recognition technique is same as static hand gesture recognition technique only the difference is that the image is acquired in real time using a Logitech Webcam C170. The image processing technique is same as the static technique.

System Overview

In the real time hand gesture recognition technique, the trained hand gestures are captured using a webcam in real time. These images undergo image processing techniques to give a corresponding contextual message and voice output for each finger counts

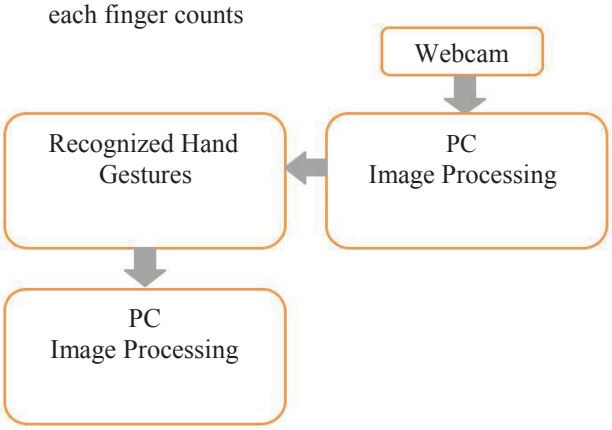


Figure 7: System Overview of real time hand gesture recognition technique

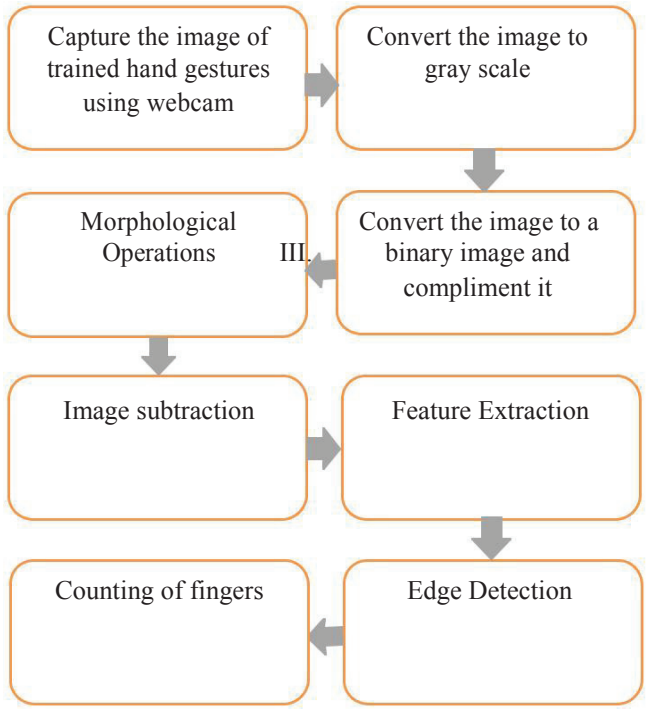


Figure 8: Procedure for real time hand gesture recognition technique.

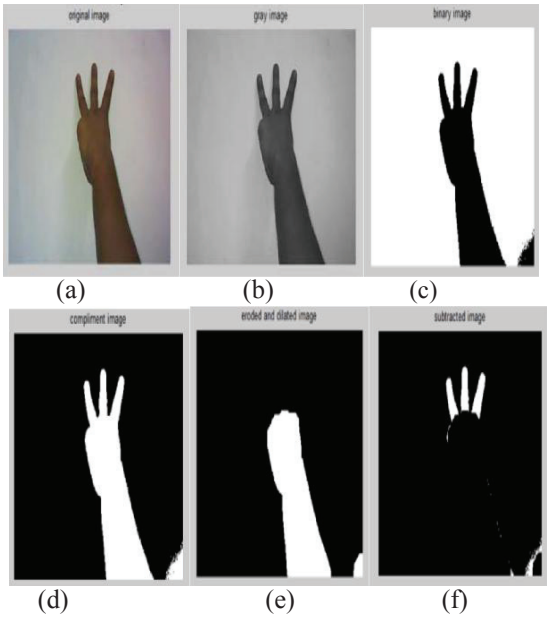
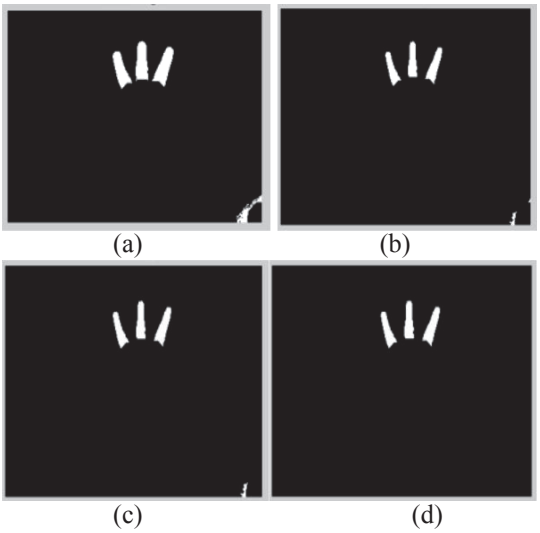
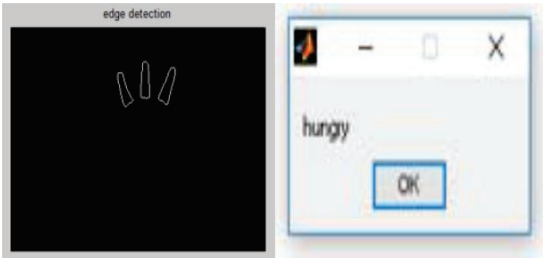


Figure 9: Output for real time hand gesture recognition technique (a) Original Image (b) Gray Image (c) Binary Image (d) Compliment Image (e) Image after erosion and dilation (f) Subtracted Image.



(a) (b) (c) (d)

Figure 10: Several processes to remove redundant pixels.



(a) (b)

Figure 11: (a) Edge Detection (b) Contextual output of real time hand gesture technique for finger count 3.

ANALYSIS TABLE FOR HAND GESTURE RECOGNITION USING IMAGE PROCESS TECHNIQUE

TABLE 1: ANALYSIS TABLE FOR HAND GESTURE RECOGNITION USING IMAGE PROCESS TECHNIQUE

Gesture	Number of samples	Successful cases	Recognition rate (%)	Error rate (%)
Yes	10	10	100	0
No	10	10	100	0
Hungry	10	10	100	0
Thirsty	10	10	100	0
Restroom	10	10	100	0
Cold	10	10	100	0
Hot	10	10	100	0
Go	10	10	100	0
Come	10	10	100	0
Pain	10	10	100	0

C. *Hand Gesture Recognition Technique using Flex Sensors* System Overview

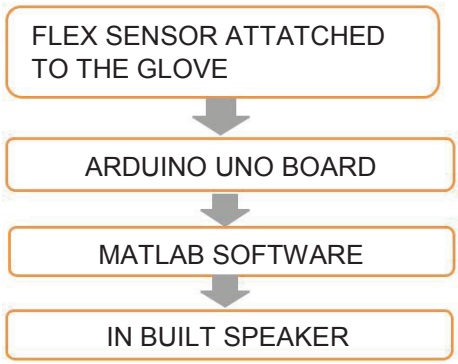


Figure 12: System overview of flex sensor based hand gesture recognition technique

Reading the analog inputs

Analog voltages of the flex sensors are given as inputs to the analog inputs of the Arduino Uno Board. The analog ports are read using the function analog read.

Write If-else condition code using output voltage of flex sensors

Output voltage of the flex sensors were recorded first and then the if-else condition is used to check which fingers are straight and which are bending. Thus the fingers counted by the Arduino is incorporated using MATLAB static or real time model of hand gesture recognition system to give an corresponding output to the fingers counted.

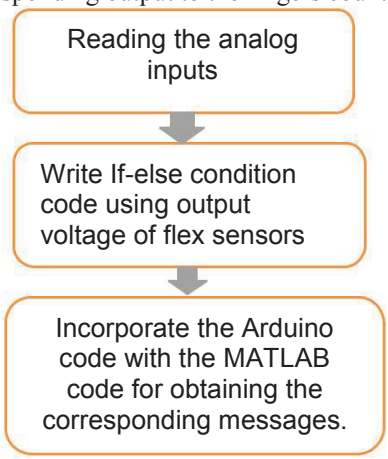
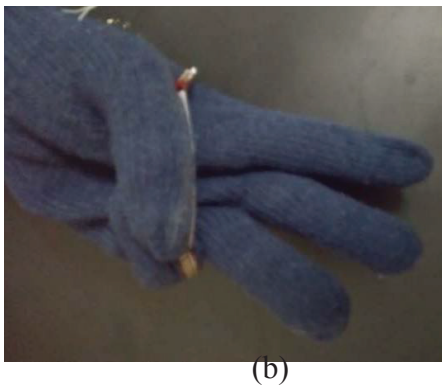


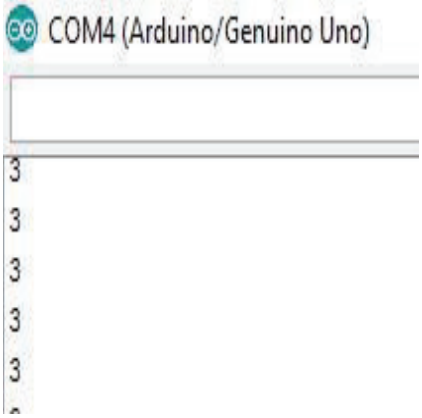
Figure 13: Procedure for flex sensor based hand gesture recognition.

(a)

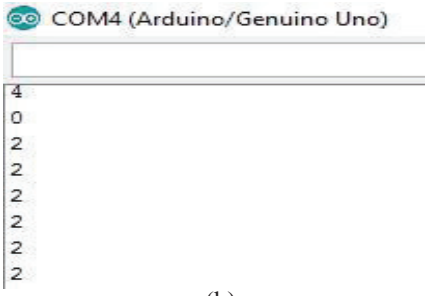


(b)

Figure 14: Hand Gestures using flex sensor shower finger count of (a) 2 (b) 3

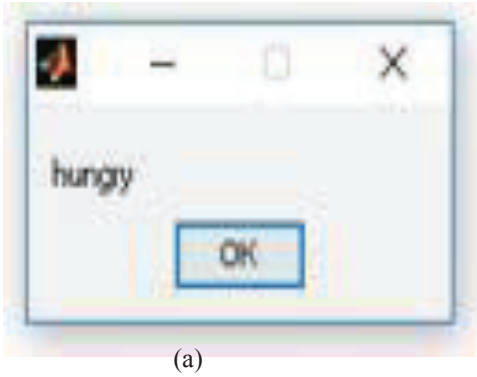


(a)



(b)

Figure 15: Output of hand gestures using flex sensors of finger count (a) 3 (b) 2 in Arduino



(a)

(b)

Figure 15: Output of hand gestures using flex sensors of finger count (a) 3 (b) 2 in MATLAB

III. COMPARATIVE STUDIES

Hand gestures recognition was performed for 10 subjects by fixing the camera at 40 cm-42 cm height. Good lighting was provided using a flash light with a clear white background for both the image processing techniques. Subjects were instructed to show fingers resembling the count of 1, 2, 3, 4 and 5. For flex sensors based hand gesture recognition model, subjects were instructed to show fingers resembling the count of 1, 2, 3, 4 and 5 by wearing a glove to which the flex sensors are attached. Output for all the three models i.e. for the static model, real-time model, and flex sensor-based model were taken and compared. For all the subjects, the conditions were kept same. The hand gestures were recognized efficiently with minimum error rate by the image processing method, while the flex sensor technique showed a considerable number of errors. The Flex sensor based technique depends on the resistance change of flex sensor which can be affected by the shear misfit of the glove or the high sensitivity issues of the flex sensor. The recognition rate of gesture recognition using flex sensors is found to be 86% whereas the recognition rate of other two gesture recognition techniques is Highly Accurate (provided with good lighting condition and clear white background).

TABLE 2: COMPARISON TABLE FOR THE THREE TECHNIQUES OF HAND GESTURE RECOGNITION

GESTURE	**RECOGNITION RATE OF ALL THREE TECHNIQUES (%)**		
	Static hand gesture recognition	Real-time hand gesture recognition	Flex sensor based hand gesture recognition
YES	100	100	70
NO	100	100	80
HUNGRY	100	100	100
THIRSTY	100	100	90
RESTROOM	100	100	90

IV. CONCLUSION AND FUTURE WORK

From the above experiment performed it was found that the flex sensor based technique have least accuracy compare to the other two vision based technique. But in the real -time recognition technique of the vision based technique's can give an accurate result only and only if good lighting condition is provided with a stationary white background. Out of all the three techniques, the static gesture recognition technique is found to be the most accurate technique.

Since the system is based on a Laptop for its working, thus it is difficult to use for practical purposes. Thus, in future, a standalone or android based system has to be made for the commercial use.

REFERENCES

[1] S. Mitra, and T. Acharya, "Gesture recognition: A survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews),* vol. 37, no. 3, pp. 311-324, 2007.

[2] W. Gonzalez, and R. E. Woods, "Eddins, Digital Image Processing Using MATLAB," *Third New Jersey: Prentice Hall*, 2004.

[3] G. Amalorpavam, H. T. Naik, J. Kumari, and M.Suresha, "Analysis of Digital Images Using Morphlogical Operations," International Journal of Computer Science & Information Technology, vol. 5, no. 1, pp. 145, 2013.

[4] J. Canny, "A computational approach to edge detection," *IEEE Transactions on pattern analysis and machine intelligence*, no. 6, pp. 679-698, 1986.

[5] P. R. V. Chowdary, M. N. Babu, T. V.Subbareddy, B. M. Reddy, and V. Elamaran, "Image processing algorithms for gesture recognition using MATLAB." pp. 1511-1514, 2014.

[6] Anushree Pillai, Spandan Sinha, Piyanka Das,Oinam Robita Chanu,"Contrivance Of Recognised Hand Gestures Into Voice And Text Output,**"** Proceedings of 35th IRF International Conference, pp.41-45,2017

[7] Y. Zhu, Z. Yang, and B. Yuan, "Vision based hand gesture recognition." pp. 260-265, 2013.

[8] R. Pradipa, and S. Kavitha, "Hand gesture recognition–analysis of various techniques, methods and their algorithms." pp. 2003-2010.

[9] P. Garg, N. Aggarwal, and S. Sofat, "Vision based hand gesture recognition," *World Academy of Science, Engineering and Technology,* vol. 49, no. 1, pp. 972-977, 2009.

[10] R. Muthukrishnan, and M. Radha, "Edge detection techniques for image segmentation," *International Journal of Computer Science & Information Technology,* vol. 3, no. 6, pp. 259, 2011.

[11] J. Babaud, A. P. Witkin, M. Baudin, and R. O. Duda, "Uniqueness of the Gaussian kernel for scale-space filtering," *IEEE transactions on pattern analysis and machine intelligence*, no. 1, pp. 26-33, 1986.

[12] A.-C. Tsai, A. Paul, J.-C. Wang, and J.-F. Wang, "Intensity gradient technique for efficient intra-prediction in H. 264/AVC," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 18, no. 5, pp. 694-698, 2008.

[13] A. Parveen, and R. UM, "Microcontroller based Hand Gesture Recognition System using Flex Sensor for Disabled People",pp.0975-8887,2015