International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 Issue: V Month of publication: May 2019

DOI: https://doi.org/10.22214/ijraset.2019.5239

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue V, May 2019- Available at www.ijraset.com

Design and Development of Air Gesture Algorithm based on Video Processing

Kusuma Bai R[1], R Jaya[2]
[1, 2]*Department of Computer Science and Engineering, New Horizon College of Engineering, Bangalore-560103 Karnataka, India*

Abstract: The increasing demand for Human Computer Interaction (HCI) has witnessed amazing innovations and prototypes that could be used in our day to day lives. Be it controlling TVs and mobile phones with hand gestures and even using gestures to operate systems, HCI has made interfacing with systems easy to operate. With such an intention of accessing systems and transportation making access easy to control and interface with motorised wheels, an air gesture algorithm has been developed. This paper focuses on developing the algorithm to identify hand from the background.
Keywords: Image Processing, Gesture, Python, Video processing, Open CV

I. INTRODUCTION

Paralysis is where your muscles in part of your body stop working due to many reasons. This loss of motion can affect the body in many ways. It can affect only certain parts of the body or it can affect the whole body, it can even affect one side of the face. This paralysis can be cured completely if it is recognised in earlier stages. If not, there are chances of not getting cured forever. [1].
Paralysis that affects the face, hand, and leg on only single side is termed as hemiplegia where hemi represents half and for the most part results from harm to the contrary side of the mind. Paralysis is mainly caused if the nerve to the spinal cord does not function properly. Loss of motion of both lower appendages is called paraplegia, and loss of motion of the two arms and the two legs is called quadriplegia. Loss of motion might be brief or perpetual, contingent upon the infection or damage. Since loss of motion can influence any muscle in the body, an individual may lose the capacity to move as well as the capacity to talk or to inhale unaided[2].
Most of the wheelchair wants patients either push the wheelchair manually or use a joystick based wheelchair. However, manual wheelchairs are difficult to control and electric wheelchairs are too expensive. Thus, in order to develop a cost effective interface for electric wheelchair, we have designed and developed a video processing algorithm and an embedded system prototype that can not only control direction of wheelchair motion, but also control or access home appliance.

A. Image Processing

Image can be defined as a pictorial representation of any information in the form of pixel. Each image is made up of many numbers of pixels. So the definition of pixel can be given as, it is a single unit of image. Another term that is used in terms of image processing is the resolution. It can be defined as the total number of pixels that are arranged in rows and columns. Hence image processing can be used to enhance the image. It is the processing of an image to get a useful information. In our project we are dealing with video processing where video is made up of many image frames.
The basic flowchart of any image processing or video processing is as follows:

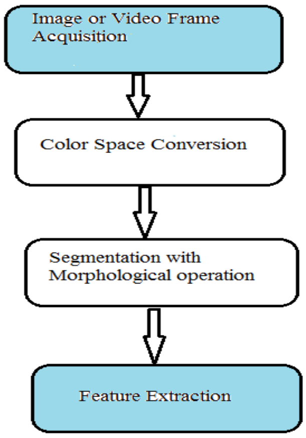


Fig 1.Flowchart of image processing

1) In the first step the images of the finger is captured using a web camera. Since this project is dealing with the video the camera captures the finger movement.
2) Once the required image is being captured, next step is to convert into suitable colour model such that the object of interest is easily differentiated from the background. In this project the object of interest will be the finger.
3) After the image has been converted into the suitable colour model it has to be segmented. The image has to be converted into black and white image. Here the finger will be converted into white colour and the background will be converted into black colour.
4) Noise can be present in an image at few cases. In such scenario it is necessary to remove the noise from the image for easier process. This is done in this stage using morphological operations.
5) At the end of this process, the extraction of the position of the finger from the image is done. Based on the position of the finger the video processing algorithm can be applied.

II. RELATED WORK

This paper discuss about the technique that is beneficial for the visually impaired people. Here they can draw the gestures based on the beep sound. This gesture was useful for such people to trace on the touch screens that were small in size[3].

There are various kinds of interface used to access and control the wheelchair. In this paper an interface was implemented in which the wheelchair was controlled by the eye movement or by simple hand gesture or by their voices. It was very complex to integrate all these methods in real life.[4] A technique which was device free was used. In this concept all the 26 alphabets were written in air for the recognition system which was cost effective and it did not include any type of sensors. This paper also discussed the drawbacks/limitations of sensor based, vision based and WiFi based system. To overcome the above limitations a technique known as Wri-fi was introduced.[5] A wearable readymade hand glouse was used for the figure gesture. All the gestures were captured by the sensors. It was very difficult to maintain it and failed to differentiate between the authorized user and the attacker. To overcome this, SVM classifier was utilized,[6] Recognition system that uses air swipe gesture was introduced. The user just made air swipes infront of the gadget's local camera. The camera captured the hand movement and made used of the library that is available in machine learning that is OpenCV for the calculation of the hand's position and apply the same for various applications.[7]

Sign language is very important for the people who cannot hear or talk. It includes concept such as irregular data and also complex structures. These irregular data is represented using air hand movement and complex structure is represented using graphs. This concept was limited because it recognises only the static pose of the hand. For the identification of dynamic hand pose a method was introduced. In this the dynamic hand movement was in the form of graph later extraction of required features was done.[8]

The user can sign in the air. Later this action is captured by the sensors that is present in the wrist wearable. This method do not restrict for different postures. A concept called gyroscope is used in this paper.[9]

III. EXISTING SYSTEM

In any image processing the first step is to read the image of our interest. Once the image is chosen it should be converted to the suitable colour model. The next step in this process will be differentiating the object of interest and the background. To do this calculation of the threshold is more important. This is done manually or by using histogram or by using mean strategy. This sometimes do not give an accurate results. To extract the position of the finger centroid based calculation is used. the following explains few limitations of the existing system: calculation of the centroid is troublesome and this may hinder the procedure in terms of manual wheelchair it requires someone to assist the patients for moving the wheelchair few of the electric wheelchair contains joysticks that is utilized to move the wheelchair different way however more often than not this can be awkward to utilize the existing strategies are high in expense

IV. PROPOSED DESIGN

The following figure shows the block diagram of the proposed system. A box is built in such a way that all the co-ordinates were visible on the screen. The camera is placed at the center of the box to capture the image of the finger. In our project since it is dealing with video processing images will be captured continuously. The captured image is then sent to the video processor where each image goes through the various steps involved in image processing. The box was divided into 9 squares and each square had a specific function to perform. Once the finger extraction was done, the finger's position was sent to the microcontroller to access the wheelchair. The data was sent in the form of 8 bits. The control was then transferred to the micro control board that was placed on the wheels of the wheelchair. Now based on the finger's position the wheelchair could do different operation.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue V, May 2019- Available at www.ijraset.com

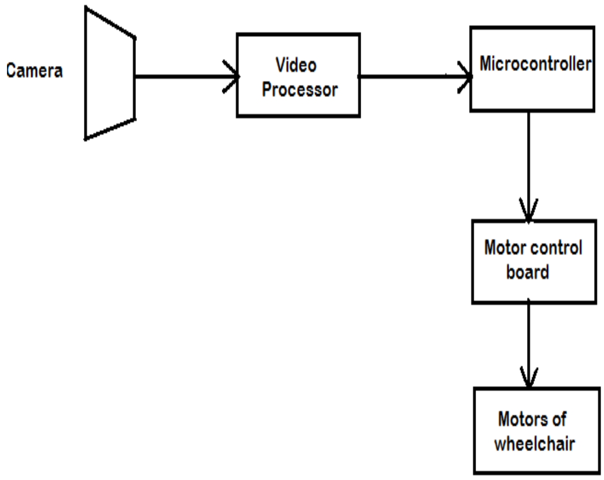


Fig 2. Proposed Block diagram

V. METHODOLOGY

The following flowchart will explain in detail the working of the entire project step by step:

A. Image Acquisition

Image acquisition is the first and foremost step in any image processing or video processing. The default colour model of any image is RGB.

B. Color Processing

The default colour model is he RGB model where R stands for red, G for green and B for blue. These are primary colours. Each of the colour value varies from 0 to 255 and hence it requires 8 bits for it's representation. If all the components are combined i.e. R, G and B then it requires 8*3=24 bits/pixel. This requires more memory to store and it is also time consuming. Therefore, it is efficient to select individual colour components rather than combining all the 3 colours together.

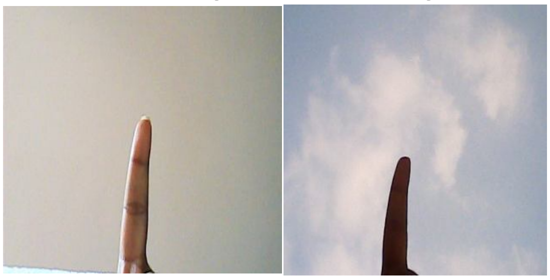


Fig. 3. Examples of RGB image

The images in the RGB format were converted to different colour models by using the following formulas:

<u>Grey-Scale</u>

Grey-scale=(R+G+B)/3

<u>YCbCr</u>

Y'=16+(65.481.R'+128.553.G'+24.966.B')

Cb=128+(-37.797.R'-74.203.G'+112.0.B')

Cr=128+(112.0.R'-93.786.G'-18.214.B')

<u>HSV</u>

M=max(R,G,B)

m=min(R,G,B)

C=M-m

$$H' = \begin{cases} \text{Undefined,} & \text{if } C=0 \\ G-B/C \bmod 6 & \text{if } M=R \\ B-R/C+2 & \text{if } M=G \\ R-G/C+4 & \text{if } M=B \end{cases}$$

$$H=60^0 * H'$$

$$S_{hsv} = \begin{cases} 0 & \text{,if } V=0 \\ C/V, & \text{otherwise} \end{cases}$$

$$V=M$$

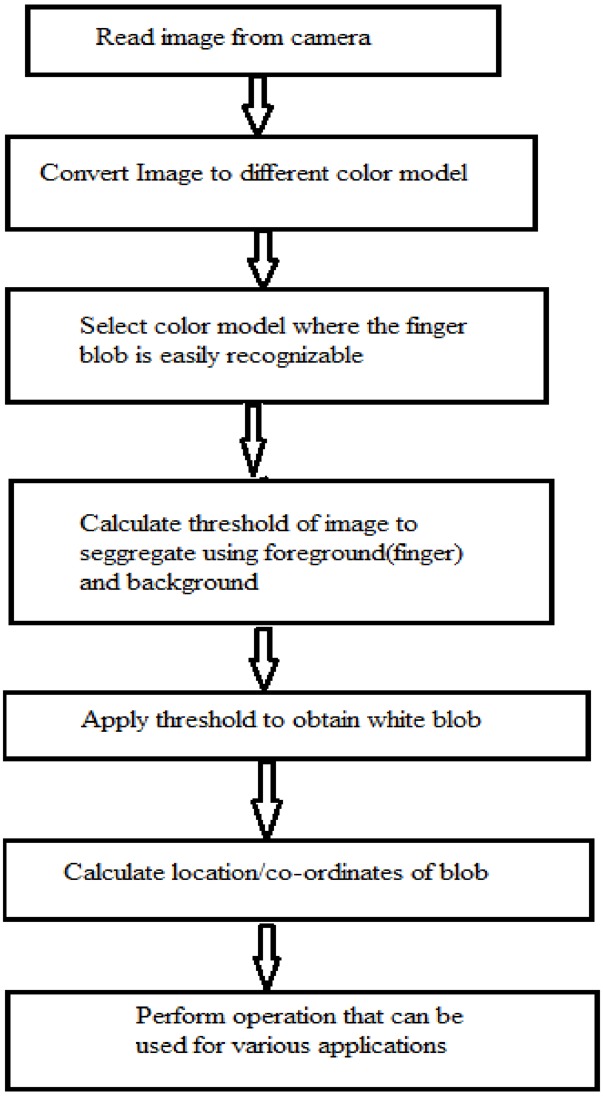


Fig 4.Flowchart for proposed system

Apart from these primary colour models there are few other colour models like YCbCr, grayscale, HSV ect. Each of these colour models have their own significances.

By closely examining the above images we can conclude that S component is the most suitable colour model. To get a clear clarification about how the S component was selected we can look at the chart below. Here most of the images look good in S component and it is also brightness independent.

Fig 5. Case study of various colour models

S component colour model is chosen based on the highest number of ticks obtained and it is also brightness independent. When this was tried with various background, it showed that S component was not suitable for the dark background. The following figure shows the example of such scenario.

For the dark background a different case study was done. Here we placed a torch light inside the box such that the finger illuminated and was easily differentiated from the background. In terms of selecting the suitable colour model Cb component was holding good for most of the images especially one with the dark background. The following figure explains this scenario.

So it can be concluded that the selection of the most suitable colour model was divided into two different categories i.e. S component was used for most of the images and for few images with the dark background Cb component was suitable. Even though we have considered two different scenarios and two different colour components, there are still difficulty in conversion of the image into suitable colour model since the back ground is not fixed. Since the normal background is concerned as of now the S component was chosen for the further processing such that the finger blob was easily recognizable.

C. Segmentation

Segmentation is mainly used to convert the colour image into white and black image. This is mainly done to track object of interest easily. In this project finger is the object of interest. In this experiment we have performed segmentation in two different ways. One is by calculating manually and other method by using machine learning. The object of interest is converted into white and the background is converted into black. This conversion is done by using the threshold. We can either use minimum threshold or maximum threshold. Firstly the threshold is calculated manually. This can be done by observing the Smin and Smax values of each

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue V, May 2019- Available at www.ijraset.com

traing images. Out of all the Smin value the minimum value is selected as Smin and the maximum value is selected as the Smax. So by using the manual method if the pixel value is greater than Smin and less than Smax. We convert that particular pixel into white and the remaining pixel into black. To calculate the minimum threshold using machine learning concept, first calculate the mean of each image and the corresponding intensity value of that particular image i.e. the Smin. Once these values are obtained then a graph is being plotted where X axis are mean value and Y axis are the intensity value. After the execution, a graph is generated along with the suitable equation. The machine learning equation chosen is:

F1(x)=p1*x^4+p2*x^3+p3*x^2+p4*x+p5

Once the threshold is obtained segmentation is performed using the following logic

$$Seg_{r,c} = \begin{cases} \text{White, } in_{r,c} >= \text{threshold} \\ \text{Black, } in_{r,c} < \text{threshold} \end{cases}$$

D. Noise Removal

Noise removal can be done by using the morphological operations. This mainly consists of two operation

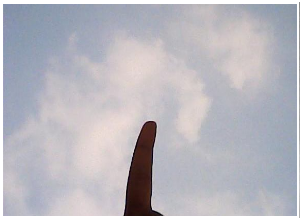
1) Erosion
2) Dilation

Erosion is done mainly to remove the unwanted white pixel from the image. Both erosion and dilation works only on center pixel. Structuring element is the main concept used here. In structuring element we try to divide the image into matrix. It is also known as sliding window. To perform erosion use the structuring element. The colour of the center pixel is checked, if it is black then it remains unchanged. If center pixel is white then search for the black pixel in the neighbour pixel.

Dilation is performed on erosion. It is mainly done for the enhancement of the object. The working of dilation is opposite to erosion. The following shows the erosion and dilation process.

VI. RESULTS

The following are the results obtained after each stage of the algorithm.


(a)


(b)

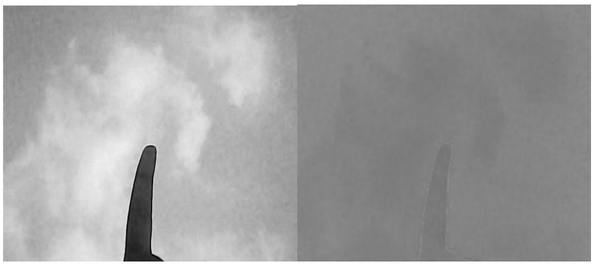

(c)


(d)

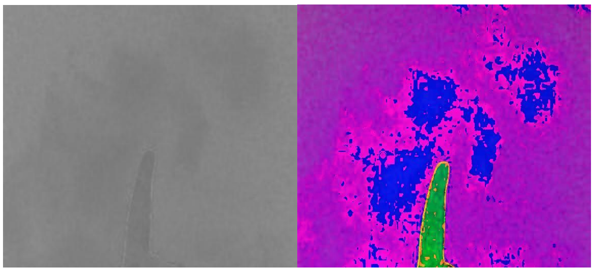

(e)


(f)

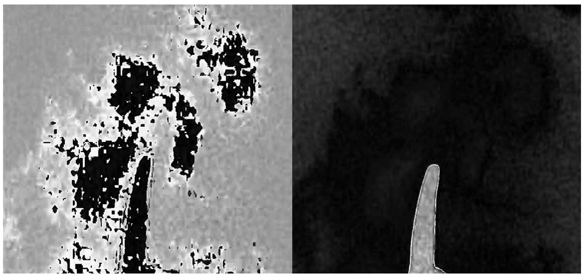
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue V, May 2019- Available at www.ijraset.com



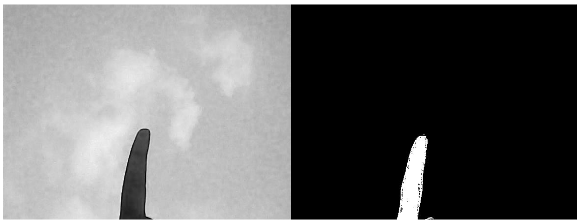
(g) (h)



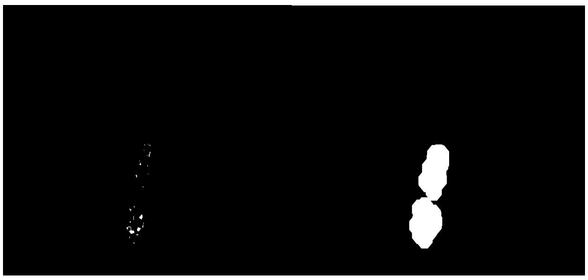
(i) (j)



(k) (l)



(m) (n)



(o) (p)

Fig 6. Results (a) RGB Image (b) R image (c) G image (d) B image (e) Grayscale image (f) YCbCr image (g) Y image (h) Cb image (i) Cr image (j) HSI image (k) H image (l) S image (m) I image (n) Segmented Image (o) Erosion image (p) Dilated image

VII. CONCLUSION

This project aims at developing a video processing algorithm for air gesture. As discussed earlier, a machine learning approach has been adopted for performing segmentation, which helps in segmenting hand from background irrespective of lighting conditions in the environment. This method can be further applied to control any HCI based system.

REFERENCES

[1] https://www.healthline.com/health/paralysis#types

[2] www.humanillnesses.com/original/pan-pre/paralysis.html

[3] BoYuGao, HyungSeok Kim, Hasup Lee, Jooyoung Lee, Jee-In Kim,*Effects of Continuous Auditory Feedback on Drawing Trajectory-Based Finger Gestures* , Year: 2018.

[4] D. H. De La Iglesia, G. Villarrubia, J.F. De Paz and J. Bajo,*Design And Implementation Of A Low-Cost Universal Control For Intelligent Electric*, , IEEE Latin America transactions, vol. 16, no. 5, Pages: 1328-1336, Year: May 2018

[5] Zhangjie Fu, JiashuangXu, Zhuangdi Zhu, Alex X. Liu and Xingming Sun, *Writing in the Air with WiFi Signals for Virtual Reality Devices*, Year: 2018

[6] Duo Lu, Kai Xu, Dijiang Huang,, *A Data Driven In-Air-Handwriting Biometric Authentication System*, , IEEE International Joint Conference on Biometrics, Pages: 531-537 ,Year: 2017

[7] Twinkle Sharma, Sachin Kumar, Naveen Yadav, Kritika Sharma, PiyushBhardwaj,*Air-Swipe Gesture Recognition Using OpenCV in Android Devices*

[8] BasheerAlwaely , CharithAbhayaratne,*Graph Spectral Domain Feature Representation for in-Air Drawn Number Recognition*, , 25th European Signal Processing Conference (EUSIPCO), Pages: 370-374, Year: 2017

[9] Chenyu Huang, Zhice Yang, Huangxun Chen, Qian Zhang, *Signing in the Air w/o Constraints: Robust Gesture-based Authentication for Wrist Wearables*, Year: 2017

[10] Rupanagudi S.R. et al. (2020) A Novel Air Gesture Based Wheelchair Control and Home Automation System. In: Abraham A., Cherukuri A., Melin P., Gandhi N. (eds) Intelligent Systems Design and Applications. ISDA 2018 2018. Advances in Intelligent Systems and Computing, vol 941. Springer, Cham

[11] S. R. Rupanagudi et al., "A novel and secure methodology for keyless ignition and controlling an automobile using air gestures," 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Jaipur, 2016, pp. 1416-1422

[12] S. R. Rupanagudi et al., "A high speed algorithm for identifying hand gestures for an ATM input system for the blind," 2015 IEEE Bombay Section Symposium (IBSS), Mumbai, 2015, pp. 1-6.