

Adaptive Hand Gesture Recognition System Using Machine Learning Approach

Rina Damdo^{1*}, Kanak Kalyani² and Jignyasa Sanghavi³

^{1, 2, 3}Department of Computer Science & Engineering, Shri Ramdeobaba College of Engineering and Management, Nagpur, India

ABSTRACT

One of the most important problem faced by our society is that deaf and hearing-impaired people are finding it hard to come up with the advanced moving technology. The sign language plays a vital role for these people as only through the sign language they can communicate with each other. But for other people it is very challenging to understand the sign language. But as we know, it is not feasible for everyone to learn sign language, need of a smart gesture recognition system for sign language is evolved. Sign language basically includes the various movements of palms, hands and other body parts. Gesture depicts an image of a specific body part. Gesture recognition is an advanced technology which is used to recognize the posture of body parts. Objective of this type of software is to provide medium between the human body movements with computer system. In gesture recognition system particularly hand gesture recognition provides a much effective and natural way for non-verbal data exchange when working with HCI system. This paper presents a real time vision approach for recognizing hand gesture based on machine learning using MATLAB. Information gathering is a first step in our proposed approach where we will develop dynamic hand gesture dataset of sign language. In next step we will represent our original proposed framework to depict hand gesture recognition with the help of hand shape and its various motions descriptors. This second step will recognize the sign once it has been captured and is much challenging trade, especially in a continuous stream which a main focus of our work. In this paper we have presented an adaptive hand gesture recognition system using machine learning approach which is adaptable to multiple applications, thus making the gesture recognition systems to be application adaptive.

KEY WORDS: ADAPTIVE HAND GESTURE RECOGNITION, MACHINE LEARNING, HUMAN COMPUTER INTERACTION, COMPUTER VISION.

INTRODUCTION

Recent study revealed that various hand gesture recognition is a trendiest area of research that attract lots of researchers in the field of machine learning and computer vision-based system. In this gesture recognition system specifically adaptive hand gesture recognition

approach is a most common way for interacting with HCI system where researchers are working mostly to develop application that can be operated with the help of only human being's gestures that include various parts of human body like hand, legs, facial expression, various musical facial expression etc. without taking help of any extra communication devices. HCI systems has been developed by combining various types of sub-domains like 3D/2D graphical interfaces systems, different types of multimedia based supported interactive systems or any virtual Advanced Environment based interactive system.

Including various advanced complex worldview to human being correspondence, cooperation, training & learning, Virtual environment frameworks additionally

ARTICLE INFORMATION

Received 7th Oct 2020 Accepted after revision 28th Dec 2020
Print ISSN: 0974-6455 Online ISSN: 2321-4007 CODEN: BBRBCA

Thomson Reuters ISI Web of Science Clarivate Analytics USA and Crossref Indexed Journal



NAAS Journal Score 2020 (4.31) SJIF: 2020 (7.728)
A Society of Science and Nature Publication,
Bhopal India 2020. All rights reserved.
Online Contents Available at: <http://www.bbrc.in/>
Doi: <http://dx.doi.org/10.21786/bbrc/13.14/26>

give different complexities as it has more advanced ways of data-presentation techniques and data exchange. For accomplishing common HCI collaboration for the virtual environment-based Systems, generally various hand gestures can work as a data input for system. Nevertheless, the excellent representation of different hand motions is still uncovered in every field of HCI based system-application. Basically, the various methods applied for solving the hand gesture-based system consist of determining the specific number of pixels in the captured image which covers the image of hand gesture, extracting the various features of captured picture for determining the pose and the occurrences of specific pose sequential data of the real time movements of the same. Various gesture analyses the deep interpretation of different human being behavior from different visual input device and is one of the latest research areas for researchers.

It is not just because of its ongoing and excited scientific challenges but also motivated by the increase of social requirement in terms of various software applications, for example modifications in real and virtual environments, monitoring system like CCTV security systems, medical support system, entertainment system etc. In real time application the accurate data should be fed so that the working of system should be evaluated very precisely. Hence accurate and complete gesture data should be obtained for any gesture recognition system. This paper proposes an approach for adaptive hand gesture recognition based on machine learning algorithm namely Artificial Neural Network (ANN). Here we have tried to represent the basic steps that should be carried out for extracting hand gestures from a sequential image frame of real time hand gesture. The rest of this paper is designed as follows: First we will review literature in this area. Next, we present the brief description of proposed methodology followed by the result obtained with two different light conditions and conclusion.

Literature Review: Recent study shows that many varieties of technology has been applied in gesture recognition system. Now a day's advanced machine learning algorithm has been applied successfully to various areas of research such as automatic face recognition system, automatic musical gesture recognition system, gesture recognition system, classification of various human's activity recognition system of automatic robotic road-detection system etc. Table 1 Compares various Gestures based techniques.

Hand Gesture Techniques: For accomplishing any type of work, initial step is to collect proper and complete information. For hand gestures recognition system, many techniques are available that can be used to determine input information for the system. Few prominent methods are listed below:

1. Vision based approach
2. Glove/Sensor Based approach

Vision Based Approach: In this type of approach [10], the web camera of computer system is used as an input

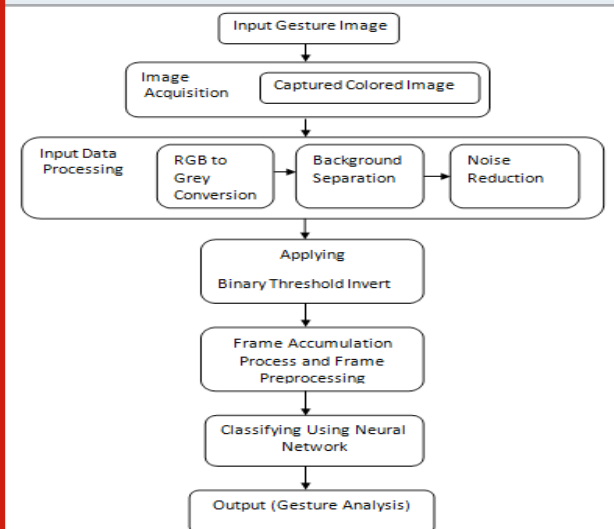
device to take input from the user for providing data as input to system. Here web camera is working like a vision system for the approach. This approach just needs a good quality of camera thus analyzing a natural vision of interaction between human and computers system.

Glove-based Approach: In this approach artificial hand gloves are generally used to provide the accurate position of hand gesture as an input to the system. These hand gloves basically consist of various sensor devices of digitizing hand and various finger positions accurately into non-parametric information. The add-on extra sensor device provides help to get more accurate hand configuration and various hand movements. This type of artificial hand glove is little expensive device and should be handled with lots of care, especially if a user is a non-technical person. This approach needs lot of set-up with the devices as the artificial hand gloves is connected externally with the system using different type of connecting cables and sensor devices.

Proposed Methodology: Now a days touch-free technology is developed everywhere. Now we don't need any device to connect with the computer system. We have wireless mouse to connect with system, just like keyboard, display device etc. For various technological applications gestures recognition plays a very important part. Specially in non-verbal communication which include hand movement, are considered as hand gesture recognition system. Non-verbal communications which include hand movement are considered as Gesture. Hand gesture recognition system are categorized into two types of gesture recognition as:

1. Static Hand gesture Recognition
2. Dynamic Hand gesture Recognition System.

Figure 1: Block Schematic of Proposed System



Our research work mainly focuses on dynamic hand gesture recognition system. Here in adaptive hand gesture recognition, real time movements of hands are included and analyzed by the system. Our proposed approach is based on convolutional neural network machine learning

algorithm. These types of network are mainly developed to recognize different visual patterns. Basically, set of image frames of hand position in every second is required for processing of our proposed approach. Fig 1 shows the block schematic of our proposed approach.

Our proposed approach will work in following three steps:

1. Capturing the Hand Gesture
2. Extracting the features
3. Recognizing the Hand Gesture.

Processing Steps: Step-1] Initially real time hand gesture is first fed to the system as an input as shown in Fig 2 a, b, c, d. Here the web camera or any image capturing device can be used as input device to the system. At a time at max 30 frames can be processed.

Figure 2: (a) Frame-1, (b) Frame -2, (c) Frame-3, (d) Frame-4



Step-2] Color Conversion Process: After the captured image is stored and processed by image capturing system it will do conversion of RGB colors in grey scale colors. This color conversion is called as RGB color presentation to HSV (Hue, Saturation, Value) image presentation. HSV color system has a color combination of Hue in "Radial Slice" manner with a center of point of each radial is a neutral color which comes mainly in color range of white to black in every top. The specific values are then used to find particular area of frames region:

$F = \max(G, R, B)$

RGB color function uses following formula to determine min and max range of different color of an image.

$RGB_{max} = \max(R, G, B)$

$RGB_{min} = \min(R, G, B)$

Following is the algorithm used for converting image from RGB color into HSV color model.

Step-1] $Y = (0.299 * R) + (0.587 * G) + (0.1115 * B)$

Step-2] $U = (B - Y) * 0.0565$

Step-3] $V = (R - Y) * 0.713$

Step-4] $UV = U + V$

Step-5] $R1 = R * 0.299, R2 = R * 0.0569, R3 = R * 0.114$

Step-6] $G1 = G * 0.299, G2 = G * 0.589, G3 = G * 0.115$

Step-7] $B1 = B * 0.288, B2 = B * 0.537, B3 = B * 0.114$

Step-8] $R4 = R1 + R2 + R3 / 3, G4 = G1 + G2 + G3 / 3, B4 = B1 + B2 + B3 / 3$

Step-9] End

Step-3] Background Separation Processing: In this step the identification of the moving object is done from the current scenario. As shown in Fig. 2 a, b, c, d the

If $\text{abs}(\text{Similarity_Prev_Similarity}) > T1$;

Progress += c;

Prev_similarity = similarity;

Figure 3: (a) Frame 1, (b) Frame b, (c) Frame 3, (d) Frame 4



background of image was having blue color which has been removed here in this step and the output is shown in Fig.3 a, b, c, d.

For removing the background from the captured image various methods are available which is elaborated by author Z. Vukovic [7] which is based on Gaussian calculative method for image segmentation in which every pixel is analyzed and used to subtract background from the moving hand in the captured image.

Step-4] Noise Reduction: In real time image capturing process, lot of noise is also captured with the input hence the noise must be removed from the input image. Here in this step captured noise is removed by applying median blur method with a particular kernel capacity of 12 to each frame for smoothing the image.

Step-5] Applying Bit/Binary Threshold Value: For obtaining the uniform binary image frameset binary thresholding process is applied where the inversion process of black and white color is done in the image. Typically, thresholding is needed to generate global threshold value. It should be noted that its value is varied in different light conditions is an image captured.

Step-6] Frame Accumulation & Processing: Here all the pre-processed frames will be accumulated for recognizing the gesture. The last frame will be in light color and more deviation time. The value of color intensity will completely depend on the movement of the hand motion occurred. Hence the brightness of the upcoming frame will depend on the deviation between the current and previous merged image and the time of the frame captured with image capturing device. The deviation calculation method is elaborated in [8] by author Zhou.

Step-7] Processing of Frames: For processing of frames following algorithm is used;

Step-1] Start

Step-2] Replace all 0's in "frames" having value of "PROGRESS"

Step-3] $\text{After_Mod} = \text{mod}(\text{frame}, 255)$;

Step-4] $\text{Merged_Frame} = \text{ADD}(\text{After_Mod}, \text{Merged_Frame})$

Step-5] $\text{Similarity} = \text{Struct_sim}(\text{init_Frame}, \text{Merged_Frame})$

If Similarity < T2;
 Classify (Merged Frame)
 Step-6] End

Step-8] Classification Process: Here in this step frames are continuously fed to Recursive Neural Network (RNN) for classification process & the gesture is analysed. Here RNN is used which consist of basic three layers. Every layer in this network contains 200 neurons. Gradient-Based optimization technique as described in paper has been selected as solver which is best suitable for large dataset and capable of handling too much noisy gradient data, needs very less memory for processing and proven excellent network for real-time application. As we had captured 2-D images and for giving them as an input to neural network, particularly 1-D array is generated using “FLATTEN” method. For training the RNN, hand movement details of every frame are generated and stored in memory. Afterwards these stored images will be transferred from memory with their details and given as input to RNN.

After the completion of training with selected dataset, this network is combined with the storage memory further for accomplishing classification processing and this has been done by us using python library functionalities. Gesture analysis is done by measuring structural value which is mentioned in Step-7. Similarity of frames value

is calculated using algorithm given in Step-7. Initially the measure value of T1=1, as soon as the system user will start moving his/her hand in any direction, it will start reducing and will continue till the deviate value from blank white image. If this measure value will get reduced lower than our decided threshold value T2, immediately the final frame is sent to recursive neural network to further classification process.

RESULTS AND DISCUSSION

For execution purpose Intel corei7 (2.8 GHz) Mac OSX with 8GB DDR3 system configuration is used. We have considered real time hand gesture & the frames have been captured and classified in real time style with the help of inbuilt web-camera of system. For execution process we have considered light condition to see the effect of the variation of light in performance of our system.

Table 2 depicts the adaptive real time hand gesture analysis in two different light conditions of two different system users. Roman numbers describe the hand gestures. As depicted in Table 2 with good light situation, our proposed system is having 89% success where as in poor light condition the success rate of our proposed system is 79%. Hence, we can conclude that in our proposed system light has a significant effect gesture recognition ability.

Table 2. Success Rate with two different User Hand Gesture Analysis

Gesture User	Good light Condition (% Success rate)					Poor light Condition (% Success rate)				
	i	Ii	iii	iv	%	i	ii	iii	iv	%
1	99	82	90	89	88.9	89	79	88	67	72
2	89	90	89	80	92	68	61	70	67	66.9
Total	94	86	89	85	90	78	70	79	67	69

CONCLUSION

From past years we analysed that dynamic hand gesture recognition has been proven a complex challenging issue, in so many areas and hence our research work is a bit ahead solution for these issue towards gaining the desired result in ongoing research field. We had tried to propose a vision based and RNN-based adaptive hand gesture recognition system. Result shows that light condition will have much effect in the performance of our recognition system. Good Light will provide better result than a poor light condition. In future this effect can be reduced by applying filtering method. Neural network classification provides excellent way for classification process and identifies a correct gesture with good probability. Real-time execution with RNN will gives considerably good results.

REFERENCES

Alpaydin, E., Introduction to Machine Learning2004: MIT

Press. Davies, E.R., Machine Vision - Theory, Algorithms, Practicalities. 3rd Edition ed, ed. Elsevier2005: Morgan Kaufmann.

Ben-Hur, A. and J. Weston, A User’s Guide to Support Vector Machines, in Data Mining Techniques for the Life Sciences 2008.

E. Ohn-Bar and M. M. Trivedi, "Hand Gesture Recognition in Real Time for Automotive Interfaces: A Multimodal Vision-Based Approach and Evaluations," in IEEE Transactions on Intelligent Transportation Systems, vol. 15, no. 6, pp. 2368-2377, Dec. 2014, doi: 10.1109/TITS.2014.2337331.

Hong R, H. M. Wang and Jian L, "Privacy-Preserving k- Nearest Neighbor Computation in Multiple Cloud Environments, IEEE Access, ISSN: 2169-3536, Volume 4, Page No’s: 9589-9603, December, 2016.

Ke, W., et al., Real-Time Hand Gesture Recognition for Service Robot. 2010: p. 976-979.

Roth, M., et al., Computer Vision for Interactive Computer Graphics, in IEEE Computer Graphics And Applications 1998, Mitsubishi Electric Research Laboratory. p. 42-53.

R. Xie and J. Cao, "Accelerometer-Based Hand Gesture Recognition by Neural Network and Similarity Matching," in IEEE Sensors Journal, vol. 16, no. 11, pp. 4537-4545, June 1, 2016, doi: 10.1109/JSEN.2016.2546942.

S. Y. Kim, H. G. Han, J. W. Kim, S. Lee and T. W. Kim, "A Hand Gesture Recognition Sensor Using Reflected Impulses," in IEEE Sensors Journal, vol. 17,

no. 10, pp. 2975-2976, 15 May 15, 2017, doi: 10.1109/JSEN.2017.2679220.

Snyder, W.E. and H. Qi, Machine Vision 2004: Cambridge University Press.

Stephan, J.J. and S. Khudayer, Gesture Recognition for Human- Computer Interaction (HCI). International Journal of Advancements in Computing Technology, 2010. 2(4): p. 30-35.

Vicen-Bueno, R., et al., Complexity Reduction in Neural Networks Applied to Traffic Sign Recognition Tasks, 2004.