International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 05 | May - 2021 ISSN: 2582-3930

HAND GESTURE RECOGNITION

Submitted by

PranavVerma Registration Number: 11713437
Vishal Yadav Registration Number: 11713272
AmanGangwar Registration Number: 11713176
Ramesh Prasai Registration Number: 11713115

Student, Department of Computer Science & Engineering Lovely Professional University, Phagwara, India

Abstract: A new approach that leads to hand gesture recognition basedon Computer Vision is proposed in this article. Hand gesture is a new form of technique that we used in computer vision Hand gestures are just a mode of nonverbal cues which can be applied to a variety of areas, including deaf-mute communication(physicalchallenge people), controlling machine or a robot, interaction between human and computer, automation, and medical purposes.Because of its wide range of applications and ability to effectively communicate with machines through human-computer interaction (HCI), the hand gesture recognition technology has recently gotten lot attention in recent times.Many various approaches have been used in gesture recognition research papers, many of them devoted to machine vision and instrumented sensor technologies. To put it another way, the hand sign may be categorized under a variety of headings, including stance and motion, or a combination between two of them static and dynamic.This paper examines the effect on hand gesture methods and discusses their benefits and drawbacks in a variety of situations.It also tabulates the results of these approaches, based on the techniques that are used in computer vision to interact with the differences as well as similarities, classification algorithms, number and movement's type, dataset used, and camera sort.

Keywords: hand gesture, computer vision, hand posture, human–computer interaction (HCI)

I. INTRODUCTION

When humans have established their footprint on the planet, it is critical that everyone understands what artificial intelligence (AI) can represent for the future of humanity in the twenty-first century technology era. Artificial intelligence (AI) seems to be an interesting topic in nearly all areas, including computing, science, education, medicine, industry, economics, marketing, and law.When we talk about the term A.I. i.e. Artificial Intelligence the first thing strike in our mind is Siri or Google's assistant to self-driving cars but that's not enough there are a lot more terms that come under Artificial Intelligence. There is a lot of improvement is done in computer vision from last decade, now we talk about hand. The key objective of implementing a hand gesture recognition system is to create a human with computer interaction interface in which a developer understood the movements can be used to guide a robot or relay useful information [1].

The challenge with gesture contact is how to make the resulting hand signals understandable and well received by the robot [2].Human-computer interaction (HCI), also known as Man-Machine Interaction (MMI)[3][4], is the relationship between a computer and a human, or more generally, a machine, which is meaningless without the human's proper use.As previously said, when developing an HCI device, two key characteristics should be considered[3]: usability and functionality. System flexibility applied to the extent and complexity to which the program can run and execute particular user functions effectively[3], while system functionality refers to the collection of features or resources that the system provides to the consumers[3].Influential output and powerful system are described as systems that achieve a reasonable match between these definitions[3]. Hand gestures that are used to communicate with humans and robots, and also between those people that use sign language[5].Static gestures (posture or specific pose) need dynamic movements are more versatile but suitable for real-world environments[6][7], while static gestures have a lower computational burden[6].For acquiring knowledge needed for a gesture recognition scheme, various approaches have been described [8][9]. To derive a concise overview of gesture characteristics, some approaches used external hardware equipment such as data glove equipment and color labels [8].Other techniques depend on the hand's appearance and the color of the skin to segment the hand and collect necessary features [8] are considered simple, normal, and less expensive than the methods previously described [8].Using various techniques and algorithms, several early findings described hand gesture device technologies and their increasing importance in our futures [10], especially in the field of Human Computer Interaction (HCI), Robotics, gaming, as well intelligence gathering [9][11].To determine hands using such various types of cameras, algorithms based on computer vision methodologies have been proposed. Skin color, shape, motion, structure, distance, 3D model, deep learn recognition, and other aspects are all attempted to be segmented and detected by the algorithms. These techniques present a number of difficulties, which are addressed in detail in the following parts of this article.The studies described above provide insight into certain hand gesture systems in different contexts, and they discuss issues including such scene context constraints, lighting conditions, algorithm precision for extracting features, dataset sort, knn classifier used, and implementation. However, there is no description of camera style, distance limits, or classification accuracy in any of the review papers.This paper shows how gesture recognition technologies have progressed, with a description of the various stages needed to create a stable solution that is less inconsistent by using various algorithms.

II. Methods of Hand gesture

Gesture recognition research's main goal is to create a system that can detect single human gestures and using them to convey messages or command structure. As a result, it encompasses not just the detection of human expression, but also the perception of the expression as meaningful orders.In most HCI implementations, two ways have been used to view movements. The first solution relies on statistical gloves (customizable or physical contact), while the second relies on machine vision and does not require the use of any sensors.

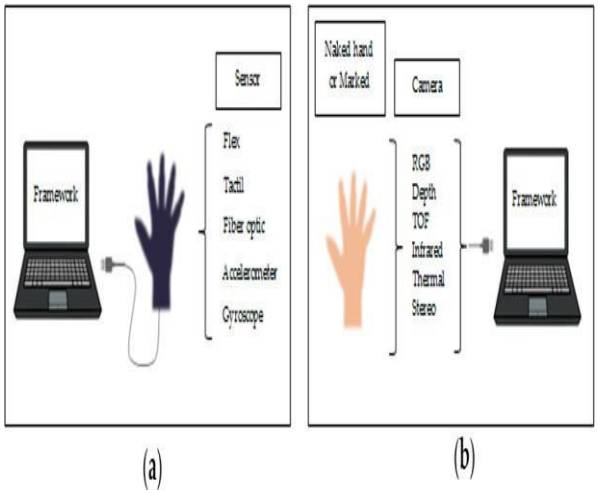

Figure 1: Different hand gesture tactics. (a) A glove-based attached sensor that can be wired to a device or carried around; (b) A computer vision–based camera that uses a marked glove or only a bare hand.

III. Instrumented Glove Approach to Hand Gestures

Hand motion and location can be captured using the portable glove-based sensors. Furthermore, using devices connected to the glove, they will precisely include the directions, positions, and dimensions of palm and finger locations. [16][17].However, this technique necessitates the user being physically attached to the device [17], which hinders user-computer interaction. Furthermore, these instruments are very expensive [17][18]. The new glove-based solution, on the other hand, employs contact technology, which is a more viable idea that is classified as Industrial-grade artificial intelligence.Using microscopy engineering, the glove provides input devices that allow the user to feel the form, texture, orientation, and weight of a simulated entity. Figure 2 depicts a glove which comprises of sensor for being converted sensor signal into hand signals.

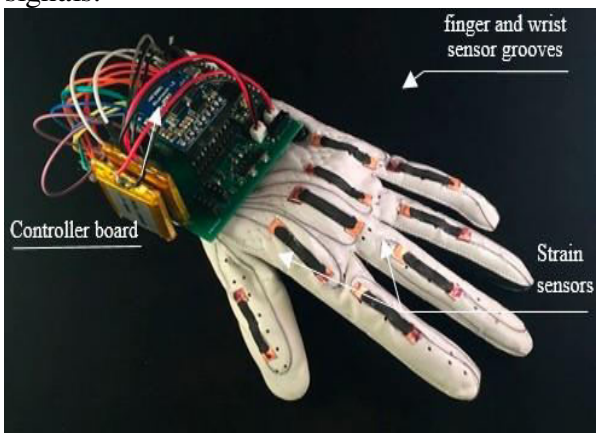

Figure 2. Sensor-based data glove

IV. Computer Vision Approach to Hand Gestures

Since it allows for contactless interaction between computers and humans, the webcam vision-based detector is a common, viable, and appropriate methodology [19]. Different viewing setups, such as sensor, super telephoto, TOF, and IR, can also be used[21].However, this methodology faces a number of difficulties, including illumination variations, backdrop problems, the impact of occlusions, dynamic backgrounds, processing time vs. quality[20][16].The following articles would go through these difficulties. Figure 3 shows a basic schematic of the camera depth perception sensor for detecting and recognizing hand movements.

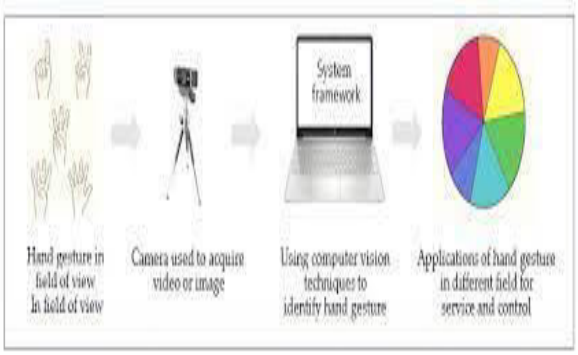

Figure 3.Computer vision methods are used to recognize movements. Where the user makes a simple gesture using one or both hands in front of the camera, which is connected to a device interface that uses a variety of techniques to retrieve features and distinguish hand gestures in order to manage a variety of applications.

V. Hand Gesture Recognition Issues: Features Extractions and Extraction Methods

After obtaining the image as input from camera(s), photographs, or a data glove optimized unit, most researchers divided hand gesture systems into three stages. These are the steps: Figure 4 shows the extraction process, attribute estimate and classification or identification.

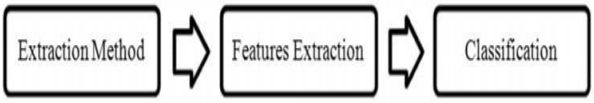

Figure 4.Hand gesture recognition steps

VI. Image pre-processing and Extraction Method

The first step in understanding hand signals is to segment them. It is the division of an input data (in this circumstance, hand motion pictures) across sectors [22] that are separated by parameters. The segmentation process depends vary according to the type of sign: if it's a dynamic motion, the hand signal must be tracked

[22], while if it's a static symbol (posture), only the input image must be segmented.To begin, the hand should always be positioned; typically, a feature extraction is used to decide the hand's location based on skin color[23], and then the hand must be monitored; there are two key ways for monitoring the hand: either as the video is separated into frames and each frame must be processed separately, in which cases the hand frame is viewed as a pose and segmented [22], in which scenario the hand frames is interpreted as a pose and segmented.The skin color [22] is a typical helpful signal for categorizing the hand because it is simple and symmetric to size, translation, and rotation adjustments [24]. To model the palm, various instruments and methods used skin and – anti pixels.The Gaussian Model (GM) and Gaussian Mixture Model (GMM) are parameterized approaches, whereas histogram-based methodologies are non-parameterized methodologies. It is, however, impaired by variations in lighting conditions in various races [6].Some studies have shown that using a glove and a colored marker are used to give accurate information regarding the orientation and position of the fingers and palm. this problem can be solved [22].Some even used an infrared sensor [6] and range information generated by a professional device called a Time-of-Flight (ToF) camera [22]. Although these instruments can detect a variety of skin tones in a cluttered setting, they are influenced by temperature changes [6].Segmentation is regarded as a concern in and of itself [9]. The grayscale images that is used in a particular function is critical to the effectiveness of the image segmentation; however, extracted features are vulnerable to lighting changes, so most studies focus on original image components only, ignoring luminance components such HS color and as r-g spaces.However, some considerations, such as [22], complicated backgrounds, lighting shifts, and poor video quality, obstruct the segmentation method.

VII. Features Extraction

An efficient recognition process requires a good segmentation process, which leads to a perfect image segmentation [6]. The based segmentation picture's features vector can be derived in a variety of ways, depending on the program. A number of techniques have been used to represent the functionality that can be derived.. Hand coordinates and outline [6] were used in some techniques, whereas fingertips position, palm middle, and other factors were used in others.

[6] Produced a function vector with 13 parameters, the first of which reflects the proportion the remaining 12 criteria reflect the actually imply visibility vectors of a scene, or even the threshold value of the side. [24] Used a neural algorithm called Self-Growing and Self-Organized Neural Gas (SGONG) to catch the outline of the hand, and The palm field, palm foundation, and hand slope were then removed.Many trials were conducted in order to determine the best block size for achieving a high recognition rate [25]. Extracted geometric center moments as feature selection method using Gaussian depicts few examples of feature extraction approaches in operation.

VIII. Gestures Classification

To identify the gesture, the gesture classification system is utilized after that the input hand image has been modeled and interpreted. The proper set of function parameters and an appropriate class label have an effect on the recognition system [7]. Contour operators or Edge detection [9], since many different hand stances are made, it can't be used to monitor gestures, resulting in misidentification [9]. The movements were classified using the Euclidean distance metric [22].While in domain of sign identification and palm type abstraction [24], neural networks have been commonly used.

IX. Proposed Approach for Recognizing Hand Gestures

Figure 5 depicts a high-level outline of hand gesture identification. The hand will be first identified using the context subtraction process, and then the effect is converted to a binary image. The fingertips and palm are divided to make finger identification easier.Furthermore,

the fingers are recognized and detected. Finally, a basic rule classifier is used to recognize hand signals.

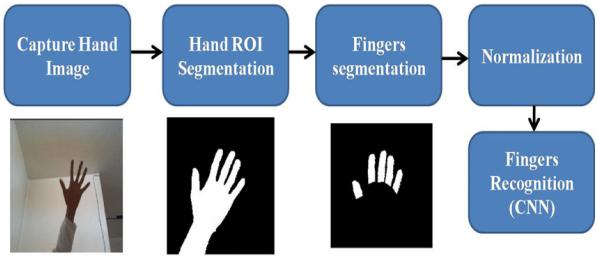

Figure 5 Proposed approaches for Recognizing Hand Gesture

The proposed approach for recognizinghand gesture scheme (shown in figure 6) is divided into three parts. We start with an RGB video of user, which is then converted to a binary video after threshold segmentation.Meanwhile, certain training set, such as corrosion, and expansion is needed. In the second section, we subtract all of the contour lines and use multiple tables to determine the contour of a palm. Finally, the focus system and pyramid - shaped merging system may be used to identify the gesture.

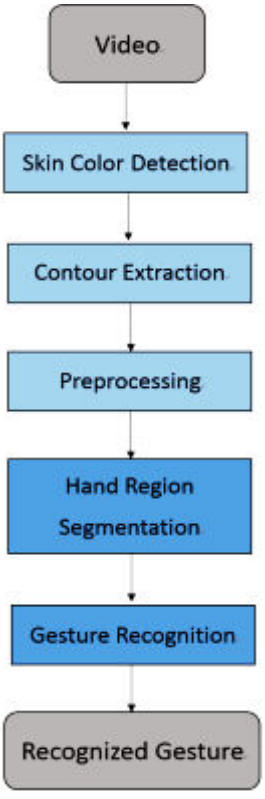

Figure 6 Hand Gesture recognition with the help of RGB video

X. Detection of skin color

We must first determine the character of the hand skin in order to identify it in the film. Ability to detect skin-colored images, while appearing to be a simple process, has proved to be very difficult in images obtained under complicated unrestrained imaging conditions, as per [12]. As a result, we devised a system based on the color function for the majority of humans. The formula is as follows:

R>85
R-B>10
R-G>10

We may accurately distinguish the skin as from surroundings, which may be referred to as dimension, using such a criterion. The image can now be converted to a binary frame.

XI. Preprocessing of data

The hand area may have cracks and holes due to the composition of the photograph, which would certainly impact the performance of the hand gesture. Since the activation function is usually blurry, classification method is used to fill in the gaps. The texture-based approach and the non-textured-based approach are the two major approaches for image data reconstruction, according to it [13].

An evolutionary algorithm-based exemplification image completing is suggested. The absolute variation method is a standard algorithm in the non-textured-based process. This paper proposes an optimized complete variance algorithm. The diffusion coefficients in the modified algorithm are affected by the difference and path between both the affected pixel and its neighbors.Although a novel in painting process achieves a successful outcome in other articles [14].

The processes, on the other hand, are often overly complex and time-consuming. To meet our order, we use some basic morphological operations like erosion and dilation. In fact, we dilate twice as much as we erode twice as much.

XII. Hand Region Segmentation and Contour Extraction

We need to isolate contours after we delete the image's noise. Each level cluster is treated as a contour. There is only one contour that reflects the hand area of these contours. Furthermore, the hand and face are the two contours.As a result of this fact, the challenge of separating palms from features is replaced by the challenge of locating the palm within the edges. As a result, we gather 100 samples from the face and hands. Then, in [15], we use VGGNet to identify them. VGGNet is a deep convolutional neural network created by the Visual Geometry Group and Google researchers.VGGNet studied the link between deep neural networks complexity and efficiency. VGG Net has recently designed deep learning models with 16-19 layers (convolutional layers and totally linked layers). by repeatedly stacking 33 small convolution cores and 22 overall pool layers. We use a 16-layer VGGNet in this article.

XIII. Gesture Recognition

To increase the feature vector and distinguish the information more effectively, the pyramidal bundling module and recognition system are used. To get all the size of 1/2 of the input images function diagram, the original input image is passed through the
3 X 3 convolutional layers and the full max pooling.The four different asset base pyramids are combined to produce feature maps with sizes of 1/4, 1/8, 1/16, and 1/32 each. As a result, the various scale characteristics can be detected. The values of global abstract elements are then used as channel dimensions at local concentrations using average global pooling.Finally, using completely linked layer and softmax, each class type final chance score is calculated.In comparison to stacked convolution and max-pooling structures, the designed methodology not only can easily obtain feature maps from various visual fields, and also aggregate the elevated image features globally, minimize the channel proportion by 1 x 1 convolutions, and efficiently use the weight of neighboring low-resolution networks.Assist in the creation of abstract functions. A significant number of studies using gesture data demonstrate that this framework can speed up network integration and increase recognition accuracy.

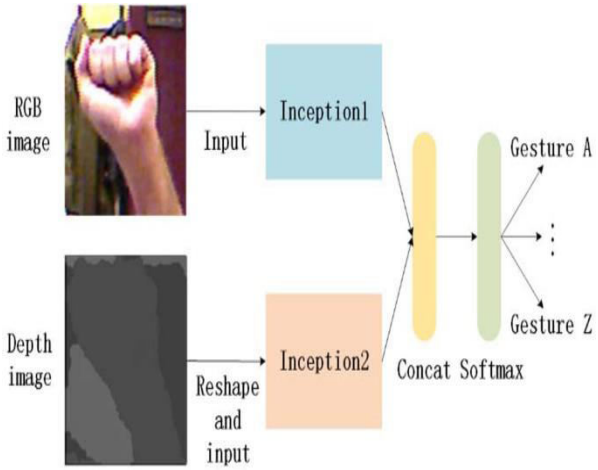

Figure 7 Training of a model

XIV. Result

We performed two experiments to test the approach suggested in this article. Hand area segmentation is the first experiment. We used 100 samples of the hand region with various gestures and 100 samples of the face region. 70 samples are used for measurement, while 30 samples are used for confirmation.The segmentation precision is 98.48 percent. To summarize, experimental precision should follow act.

The experiment involves the identification of gestures. A total of 100 research samples from 5 individuals were tested to detect 9 different movements (shown in figure 8). There are 70 evaluation data and 30 confirmation data for each gesture.

Figure 8Common gestures.

XV. Conclusion

Various approaches for gesture control are explored in this article, including Neural Networks, HMMs, k - means and clustering, and the use of an orientation graph for feature extraction. HMM tools are ideal for complex movements and have proven to be effective in a variety of applications, including robotics and for catching hand form in, NNs are used as classifiers.Even to obtain the outline of the hand, certain types of algorithms are needed for feature extraction, where used a Gaussian bivariate method for fit the based segmentation hand to reduce the rotating affection.The recognition algorithm that is chosen is determined by the application. The implementation areas for the gestures scheme are presented in this work. There is a detailed explanation of hand gesture problems and also a serious analysis of latest recognition schemes. A summary of a few chosen systems is also given.In Human-Computer Interaction, gesture is crucial. This paper proposes a system for recognizing gestures that is both reliable and effective. The representation of skin should be first studied through perspective rules, and then transformed to a feature vector. Extension and corrosion are seen after that. The hand edge is then found after all of the edges have been removed.

XVI. Acknowledgement

I'd like to express my gratitude to Gagandeep Singh, one of our academic staff, for exchanging his brilliant ideas on this topic with us. Thank you to all who helped with this research study, and please refrain from using powerful impressionistic words.

XVII. References

[1] G.N.R.S. Murthy, R. S. Jadon. (2009). “A Review of Vision Based Hand Gestures Recognition,” International Journal of Information Technology and Knowledge Management, vol. 2(2), pp. 405- 410.

[2] P. Garg, N. Aggarwal and S. Sofat. (2009). “Vision Based Hand Gesture Recognition,” World Academy of Science, Engineering and Technology, Vol. 49, pp. 972-977.

[3] FakhreddineKarray, MiladAlemzadeh, JamilAbouSaleh, Mo Nours Arab, (2008) .“Human- Computer Interaction: Overview on State of the Art”, International Journal on Smart Sensing and Intelligent Systems, Vol. 1(1).

[4] Kramer, R.K.; Majidi, C.; Sahai, R.; Wood, R.J. Soft curvature sensors for joint angle proprioception. In Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Francisco, CA, USA, 25–30 September 2011; pp. 1919–1926.

[5] Mokhtar M. Hasan, Pramoud K. Misra, (2011). “Brightness Factor Matching For Gesture Recognition System Using Scaled Normalization”, International Journal of Computer Science & Information Technology (IJCSIT), Vol. 3(2)

[6] Xingyan Li. (2003). “Gesture Recognition Based on Fuzzy C-Means Clustering Algorithm”,Department ofComputer Science. The University of Tennessee Knoxville.

[7] S. Mitra, and T. Acharya. (2007). “Gesture Recognition: A Survey” IEEE Transactions

on systems, Man and Cybernetics, Part C: Applications and reviews, vol. 37 (3), pp. 311- 324, doi: 10.1109/TSMCC.2007.893280.

[8] Simei G. Wysoski, Marcus V. Lamar, Susumu Kuroyanagi, Akira Iwata, (2002). "A Rotation Invariant Approach On Static-Gesture Recognition Using Boundary Histograms And Neural Networks," IEEE Proceedings of the 9th International Conference on Neural Information Processing, Singapura.

[9] Joseph J. LaViola Jr., (1999). "A Survey of Hand Posture and Gesture Recognition Techniques and Technology", Master Thesis, Science and Technology Center for Computer Graphics and Scientific Visualization,USA.

[10] Rafiqul Z. Khan, Noor A. Ibraheem, (2012). "Survey on Gesture Recognition for Hand Image Postures", International Journal of Computer And Information Science, Vol. 5(3), Doi: 10.5539/cis.v5n3p110.

[11] Thomas B. Moeslund and Erik Granum, (2001). "A Survey of Computer Vision-BasedHuman Motion Capture," Elsevier, Computer Vision and Image Understanding, Vol. 81, pp. 231–268.

[12] Chen W, Wang K, Jiang H, et al. Skin color modeling for face detection and segmentation: a review and a new approach[J]. Multimedia Tools & Applications, 2016, 75(2):1-24.

[13] Li K, Wei Y, Yang Z, et al. Image inpainting algorithm based on TV model and evolutionary algorithm[J]. Soft Computing - A Fusion of Foundations, Methodologies and Applications, 2016, 20(3):885-893.

[14] AlexandruTelea. An Image Inpainting Technique Based on the Fast Marching Method[J]. Journal of Graphics Tools, 2004,9(1):23-34.

[15] K. Simonyan and A. Zisserman, Very deep convolutional networks for large-scale image recognition, Proc. Int'l Conf. Learning Represent., 2015.

[16] Garg, P.; Aggarwal, N.; Sofat, S. Vision based hand gesture recognition. World Acad. Sci. Eng. Technol.2009,49,972–977.

[17] LaViola, J. A survey of hand posture and gesture recognition techniques and technology. Brown Univ. Provid. RI 1999, 29. Technical Report no.CS-99-11.

[18] Ibraheem, N.A.; Khan, R.Z. Survey on various gesture recognition technologies and techniques. Int. J. Comput. Appl. 2012, 50, 38–44.

[19] Kaur, H.; Rani, J. A review: Study of various techniques of Hand gesture recognition. In Proceedings of the 2016 IEEE 1st International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), Delhi, India, 4–6 July 2016; pp.1–5.

[20] Murthy, G.R.S.; Jadon, R.S. A review of vision based hand gestures recognition. Int. J. Inf. Technol. Knowl. Manag. 2009, 2, 405–410.

[21] Sonkusare, J.S.; Chopade, N.B.; Sor, R.; Tade, S.L. A review on hand gesture recognition system. In Proceedings of the 2015 International Conference on Computing Communication Control and Automation, Pune, India, 26–27 February 2015; pp. 790–794.

[22] N. Ibraheem, M. Hasan, R. Khan, P. Mishra, (2012). "comparative study of skin color basedsegmentation techniques", Aligarh Muslim University, A.M.U., Aligarh, India.

[23] Mahmoud E., Ayoub A., J¨org A., and Bernd M., (2008). "Hidden Markov Model-Based Isolated and Meaningful Hand Gesture Recognition", World Academy of Science, Engineering and Technology 41.

[24] E. Stergiopoulou, N. Papamarkos. (2009).

"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

[25] M. M. Hasan, P. K. Mishra, (2011). "HSV Brightness Factor Matching for Gesture Recognition System", International Journal of Image Processing (IJIP), Vol. 4(5).