Hand Gesture Recognition
(Unveiling the Power of Nonverbal Communication through Literature Review)

Ms. Sherilyn Kevin
Department of IT Thakur College of Science and Commerce Mumbai, India

Mr. Yash Thakur
Department of DS Thakur College of Science and Commerce Mumbai, India

Mr. Vinay Chhotelal Gupta
Department of DS Thakur College of Science and Commerce Mumbai, India

Mr. Ezra Rajendran
Department of DS Thakur College of Science and Commerce Mumbai, India

Abstract: **In recent years, hand gesture recognition systems have gained significant attention due to their diverse applications and efficient interaction with machines through human-computer interaction. This paper presents a survey of recent systems for recognizing hand gestures, highlighting key issues and challenges faced in developing such systems. The review covers various methods for recognizing postures and gestures, as well as research results from different gesture recognition systems and databases. Furthermore, a comparison between the main phases of gesture recognition is provided. Finally, the advantages and drawbacks of the discussed systems are explained.**

Keywords – Hand gesture recognition, human-computer interaction, postures, gestures, neural networks, databases, comparison, advantages, drawbacks, key issues, challenges, surveys, recent systems, and research results.

1. INTRODUCTION

In recent years, advancements in neural networks have revolutionized the field of hand gesture recognition systems, enabling more natural and efficient interaction between humans and machines. The aim of these systems is to create an intuitive communication channel that allows recognized gestures to control robots or convey important information [1]. However, gesture interaction poses significant challenges such as segmentation, features extraction, and recognition. Human-computer interaction (HCI) [3][4] plays a vital role in ensuring system functionality and usability [3]. Gestures, whether static or dynamic, have numerous applications in communication between humans and machines and sign language [5]. With the recent advancements in neural networks, more sophisticated and accurate methods for acquiring necessary information have emerged to build gesture recognition systems. This paper presents a comprehensive review of recent hand gesture recognition systems, highlighting key issues, challenges, applications, and drawbacks. The paper is organized into different sections, starting with the key issues of hand gesture recognition systems, applications, challenges, literature review of recent systems, drawbacks, research results, and concluding remarks.

2. HAND GESTURE RECOGNITION: NAVIGATING

Challenges in Feature Extraction and Extraction Methods

Gesture recognition systems typically involve three main steps, which researchers have categorized following the acquisition of an input image from a camera, video, or data glove instrumented device. These steps are extraction method, feature estimation and extraction, and classification or recognition, as depicted in Figure 1.

Fig 1. Gesture Recognition System Steps

2.1.Extraction Method and image pre-processing:

The recognition of hand gestures requires several steps, with the first being the segmentation process, where the input image is divided into regions separated by boundaries [12]. The segmentation process varies based on the type of gesture being recognized, whether it is dynamic or static. The hand gesture needs to be located and tracked [12] for dynamic gestures, while for static gestures, only the input image needs to be segmented.

The common helpful cue used for segmenting the hand is the skin color [12][13], which is easy and invariant to scale, translation, and rotation changes [14].However, this method is affected by illumination condition changes and different races [6]. To overcome this problem, some researchers use data gloves and colored markers or infrared cameras [6][12]. The segmentation [12] process is an open issue due to factors such as complex backgrounds, illumination changes, and low video quality [12]. Preprocessing operations, such as subtraction, edge detection, and normalization, are applied to enhance the segmented hand image. Figure 2, provides examples of segmentation methods. Different tools and methods are used to model the hand, including parametric

and non-parametric techniques such as Gaussian Model (GM), Gaussian Mixture Model (GMM), and histogram-based techniques. The color space used in a specific application plays a vital role in the success of the segmentation process. Color spaces are sensitive to lighting changes, and researchers tend to use chrominance components only and neglect the luminance components, such as r-g and HS color spaces [6][15][18].
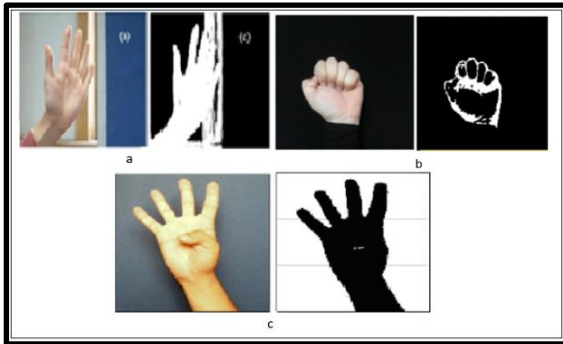

Fig. 2 Segmentation Method.

[a) [32], b) [15], c) [14]]

2.2. Features Extraction:

The process of feature extraction is crucial in achieving accurate recognition of hand gestures, and relies heavily on the quality of the preceding segmentation process [6]. A variety of methods can be used to extract features vectors from segmented images, depending on the specific application. Shape-based methods, such as hand contour and silhouette [6], are commonly used, while others may utilize information such as fingertips position or palm center.

For example, one study used a feature vector of 13 parameters [6], with the first parameter representing the aspect ratio of the hand's bounding box, and the remaining 12 parameters representing the mean brightness values of the image. Another approach utilized the [14] Self-Growing and Self-Organized Neural Gas (SGONG) neural algorithm to capture the shape of the hand, resulting in three features: palm region, palm center, and hand slope.

Another method involves calculating the [16] Center of Gravity (COG) of the segmented hand and the distance from the COG to the farthest point in the fingers, which is then used to estimate the number of fingers in the hand region. Additionally, some studies [15] divide the segmented image into blocks of varying sizes and extract the brightness factor for each block as the feature vector. However, experiments are needed to determine the optimal block size to achieve the best recognition rate.[15]
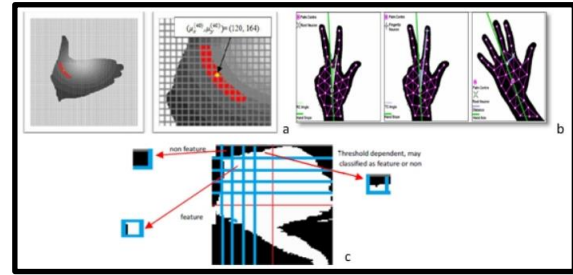

Fig. 3 Features Representation

[15][17][18]Some studies also utilize geometric central moment as local and global features, extracting them using Gaussian pdf. Figure 3, provides examples of different feature extraction methods, such as partitioning the segmented image into terraces and extracting local and global geometric central moments, or dividing the segmented hand [16][18] into blocks and using the brightness factor of each block as the feature vector (blocks with black area are discarded) [5][14].

2.3. Gestures Classification:

Recognizing hand gestures requires several steps, and the first one is to analyze the input hand image by selecting appropriate features and classification algorithms [7]. Although edge detection and contour operators [9] aren't the ideal solutions since they might lead to misclassification [9], there are other alternatives, such as the Euclidean [9][15][17] distance metric, that can be utilized for gesture categorization.

For the purpose of recognizing hand gestures, a variety of statistical techniques [20][13] have been employed, such as Hidden Markov Models (HMMs), Finite State Machines (FSMs) [21], [24][25][26]Learning Vector Quantization (LVQ), Principal Component Analysis (PCA), and neural networks. HMMs have shown particular success in recognizing dynamic gestures, while neural networks have been frequently used to extract hand form. Additionally, soft computing approaches like Fuzzy C-Means clustering (FCM) [6] and Genetic Algorithms (GAs) [27] have also shown success in identifying hand gestures.

The approach used is determined by the application and dataset, and recent advances in deep learning, notably Convolutional Neural Networks (CNNs), have demonstrated promising results for hand gesture detection. These developments indicate that hand gesture recognition systems will become increasingly accurate and popular in the future. Figure 4, explain the architecture of classification system.
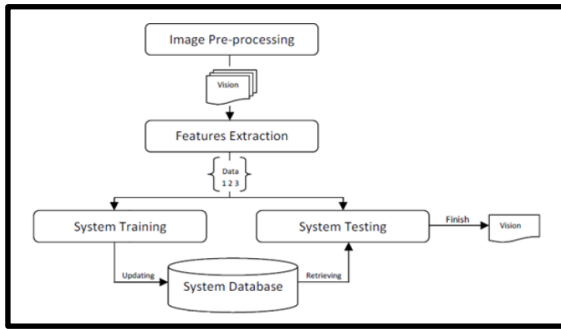
Fig. 4 Architecture of gesture recognition system

3. APPLICATION AREAS OF HAND GESTURES SYSTEM

Hand gesture recognition systems have been deployed in a range of applications across several sectors [7][9]. Sign language translation, virtual worlds, smart surveillance, robot control, and medical systems are some of these fields. Here's a rundown of some of the specific areas where hand motion detection is being used [8]:

A. Sign Language Recognition:
Sign language has been a focus of attention for gesture recognition systems, as it is used to interpret and explain various topics during conversations [7]. Several systems have been proposed for recognizing different types of sign languages [8]. For instance, American Sign Language (ASL) [8] has been recognized using boundary histogram, MLP neural network, and dynamic programming matching.[28] Japanese sign language (JSL) has been recognized using Recurrent Neural Network, 42 alphabet, and 10 words. [25] Arabic Sign Language (ArSL) has been recognized using two different types of neural networks, Partially and Fully Recurrent Neural Network.

B. Robot Control:
Controlling a robot using hand gestures is an intriguing application in the field of hand gesture recognition[6]. I recently came across a system that utilizes numbered hand poses to control a robot[16]. The system works by assigning specific meanings and functions to each numbered pose, for instance[16], the "one" pose represents "move forward" and the "five" pose represents "stop". By using these numbered poses, the user can give specific commands to the robot to perform different tasks.

C. Graphic Editor Control:
In a graphic editor control system, hand gestures are tracked and located as a preprocessing operation[7]. To draw and edit graphics,[20] 12 dynamic gestures are utilized. Drawing shapes include a triangle, rectangle, circle, arc, horizontal and vertical lines, while editing commands include copy, delete, move, swap, undo, and close[20].

D. Virtual Environments (VEs):
Virtual environments (VEs) are a popular area of application for gesture recognition systems, particularly in communication media systems[9]. A system proposed for [29] 3D pointing gesture recognition in real-time using binocular views, making it accurate and unaffected by changes in user characteristics or the environment[29].

E. Numbers Recognition:
Recently, hand gestures have been utilized in recognizing numbers as well. A new system has been [13] proposed that can isolate and identify a significant gesture from Arabic numerical hand motions in real-time using Hidden Markov Models (HMM).

F. Television Control:
People can use hand gestures and postures to control their television devices[9]. For instance, a specific set of gestures can be utilized to turn the TV on and off, adjust the volume, mute the sound, and change the channel[9]. These gestures typically involve opening and closing the hand in various ways[30].

G. 3D Modeling:
Generating 3D models of the hand necessitates knowing the various forms and angles of the hand in order to correctly create and display the 3D shape. Many systems that employ hand silhouettes[30][9] to build both 2D and 3D objects have been created. Moreover, 3D hand modelling is being investigated as a viable field of study for this reason. These models may be utilised to deliver more realistic and engaging experiences in a variety of applications such as virtual reality, gaming, and medical simulations. While there are still certain accuracy and efficiency problems to address, technological improvements are projected to increase the capabilities and possibilities of 3D hand modelling systems in the future[9].

4. Advantages that Hand Gesture Recognition Systems offer over other methods of Interaction and Communication:

1. Contactless Interaction: With the ongoing COVID-19 pandemic, contactless interaction has become more important than ever. Hand gesture recognition systems offer a completely contactless way of interacting with devices and interfaces, which can help reduce the spread of viruses and bacteria.

2. Natural and Intuitive: Hand gestures are a natural and intuitive way of communicating, and they require little to no learning curve. This makes hand gesture recognition systems particularly useful for applications where users may not be familiar with traditional input methods, such as children or elderly users.

3. Accessibility: Hand gesture recognition systems can be particularly useful for people with disabilities, such as those who have difficulty using traditional input methods like keyboards or mice. Hand gestures can be a more accessible and natural way of interacting with devices and interfaces.

4. Enhanced User Experience: Hand gesture recognition systems can enhance the user experience by providing more natural and intuitive

ways of interacting with devices and interfaces. For example, hand gestures can be used to control devices in a more fluid and seamless manner, which can improve the overall user experience.

5. Versatility: Hand gesture recognition systems can be used in a wide range of applications, from gaming and entertainment to healthcare and industrial automation. This versatility makes hand gesture recognition systems a valuable tool for many different industries and use cases.

5. SUMMARY OF RESEARCH RESULTS:

Table 1 compares the recognition methods used by various hand gesture recognition systems. Table 2 provides a summary of the application areas and invariant vectors of some hand gesture recognition systems. Lastly, Table 3 displays a summary of the extraction method, feature representation, and recognition techniques used in the selected hand gesture recognition systems, including hand extraction technique, feature vector representation, and recognition.

Method	# Recognized Gestures	# Total Gestures used For Training And Testing	Recognition Percentage	Database used
[8]	26	1040	DP 98.8%	American Sign Language (ASL)
			MLP 98.7%	
[5]	6	60	normal method 84%	Own Database
			Scaling normalization method 95%	
[31]	26	208	92.78%	American Sign Language (ASL)
[032]	0-9 numbers	298 video sequence for isolated gestures/ 270 video sequence for continuous gestures	90.45%	Recognize Arabic numbers from 0 to 9.
[20]	5 static/ 12 dynamic gestures	Totally 240 data are trained and then the trained are tested	98.3%	5 static gestures and 12 dynamic gestures.
[14]	31	130 for testing	90.45%	Own Database
[17]	6	60	100% for more than 4 gestures	Own Database
[18]	20	200	100% for 14 gestures, and >90 for 15-20 gestures	Own Database

Table 1. Comparison between recognition methods in hand gesture recognition methods used

Method	Application Area	Invariant factor
[19]	Real time system / control a computer graphic crane by hand gestures/ play games such as scissors/paper/stone	Lighting conditions / Translation
[14]	Sign Recognition	Lighting conditions / Translation
[8]	Sign Recognition	Rotation
[5]	Sign language	Rotation/ Translation/ Scaling
[16]	Robot control application	Translation/ Rotation / Scaling
[22]	Real-time system/ moderate computational resources devices e.g. netbooks	Rotation / Translation
[17]	Sign Recognition	Rotation/Translation/ Scaling
[18]	Sign Recognition	Rotation/Translation/ Scaling
[20]	Drawing graphical elements such as triangle, rectangular/ Editing graphical elements such as copy, paste, undo/ Mobile robot control/Virtual Reality.	-

Table 2. Summary of application areas and invariant vector of some hand gesture recognition systems

Method	Extraction method	Features Vector Representation	Classifier
[6]	HSV color space	13 parameters as a feature vector, the first parameters represents the ratio aspect of the bounding hand box and the rest 12 parameters are the mean values of brightness pixels in the image	Fuzzy C-Means (FCM) algorithm
[8]	-	Boundary Chord's size FFT	MLP Neural Network/ Dynamic Programming (DP) matching
		Boundary Chord's size	
		Boundary Chord's size histogram	
[14]	YCbCr color space	From Self-Growing and Self-Organized Neural obtained hand shape, then, three angles of the hand shape calculated; RC Angle, TC Angle, Distance from the palm center.	Gaussian distribution
[22]	Colored glove/ HSV and threshold based method	Nine numerical features formed by distances from palm to all fingers and four angles between the distances	Learning Vector Quantization (LVQ)
[032]	GMM for skin	Orientation quantization	HMM

	color Detection and YCbCr color space		
[28]	Thresholding technique	13 data item (10 for bending, 3 for coordinate angles)/ 16 data item (10 for bending, 3 for coordinate angles, 3 for positional data)	Back propagation network / Elman recurrent network
[5]	Thresholding	Divide the scaled normalized hand image into Blocks of intensity features.	Euclidean Distance Metric.
[18]	HSV color space/thresholding	178 features, for local and global features using moments	Euclidean Distance Metric.

Table 3. Summary of extraction method, features representation, and recognition of hand gesture recognition systems

6. CONCLUSION:

This paper has explored various methods for hand gesture recognition, including Neural Network, Hidden Markov Model (HMM), and fuzzy c-means clustering. Orientation histogram has also been used for feature representation. For dynamic gestures, HMM has proven to be highly efficient, particularly in robot control [20][16] applications. Neural Networks have been used as classifiers [8][25] to capture hand shape [14], while various methods and algorithms have been employed for feature extraction [15][17][18] and shape capturing, such as Gaussian bivariate function in [17]. The choice of recognition algorithm should depend on the specific application requirements[18].

Throughout this work, we have presented different application areas where gesture recognition systems can be used. We have discussed the challenges associated with gesture recognition and presented recent advancements in recognition systems. Finally, we have summarized some selected systems for hand gesture recognition.

Overall, hand gesture recognition systems offer a range of advantages over other methods of interaction and communication. They are intuitive and easy to use, and can be applied in various fields, including virtual reality, gaming, robotics, and healthcare. With advances in computer vision and machine learning, these systems are becoming more accurate and reliable, especially with the use of deep learning algorithms. As we move forward, we expect to see continued growth and innovation in this field, and we encourage researchers, developers, and businesses to explore the possibilities of hand gesture recognition systems to improve user experiences and increase overall efficiency.

7. REFERENCES:

[1] G. 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] Fakhreddine Karray, Milad Alemzadeh, Jamil Abou Saleh, Mo Nours Arab, (2008) ."HumanComputer Interaction: Overview on State of the Art", International Journal on Smart Sensing and Intelligent Systems, Vol. 1(1).

[4] Wikipedia Website.

[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 of Computer 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-Based Human Motion Capture," Elsevier, Computer Vision and Image Understanding, Vol. 81, pp. 231–268.

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

[13] 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.

[14] 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

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

[16] Malima, A., Özgür, E., Çetin, M. (2006). "A Fast Algorithm for Vision-Based Hand Gesture Recognition For Robot Control", IEEE 14th conference on Signal Processing and Communications Applications, pp. 1-4. doi: 10.1109/SIU.2006.1659822

[17] Mokhar M. Hasan, Pramod K. Mishra, (2012) "Features Fitting using Multivariate Gaussian Distribution for Hand Gesture Recognition", International Journal of Computer Science & Emerging Technologies IJCSET, Vol. 3(2).

[18] Mokhar M. Hasan, Pramod K. Mishra, (2012). "Robust Gesture Recognition Using Gaussian Distribution for Features Fitting', International Journal of Machine Learning and Computing, Vol. 2(3).

[19] W. T. Freeman and Michal R., (1995) "Orientation Histograms for Hand Gesture Recognition", IEEE International Workshop on Automatic Face and Gesture Recognition.

[20] Min B., Yoon, H., Soh, J., Yangc, Y., & Ejima, T. (1997). "Hand Gesture Recognition Using Hidden Markov Models". IEEE International Conference on computational cybernetics and simulation. Vol. 5, Doi: 10.1109/ICSMC.1997.637364

[21] Verma, R., Dev A. (2009)."Vision based hand gesture recognition using finite state machines and fuzzy logic". IEEE International Conference on Ultra-Modern Telecommunications & Workshops (ICUMT '09), pp. 1-6. doi: 10.1109/ICUMT.2009.5345425

[22] Luigi Lamberti, Francesco Camastra, (2011). "Real-Time Hand Gesture Recognition Using a Color Glove", Springer Proceedings of the 16th international conference on Image analysis and processing: Part I ICIAP.

[23] Minghai Y., Xinyu Q., Qinlong G., Taotao R., Zhongwang L., (2010). "Online PCA with Adaptive Subspace Method for Real-Time Hand Gesture Learning and Recognition", journal World Scientific and Engineering Academy and SocietWSEAN, Vol. 9(6).

[24] N. A. Ibraheem., R. Z. Khan, (2012). "Vision Based Gesture Recognition Using Neural Networks Approaches: A Review", International Journal of Human Computer Interaction (IJHCI), Malaysia, Vol. 3(1).

[25] Manar Maraqa, Raed Abu-Zaiter. (2008). "Recognition of Arabic Sign Language (ArSL) Using Recurrent Neural Networks," IEEE First International Conference on the Applications of Digital Information and Web Technologies, (ICADIWT), pp. 478-48. doi: 10.1109/ICADIWT.2008.4664396

[26] Tin Hninn H. Maung. (2009)."Real-Time Hand Tracking and Gesture Recognition System Using Neural Networks," World Academy of Science, Engineering and Technology 50, pp. 466-470.

[27] Cheng-Chang L. and Chung-Lin H., (1999)."The Model-Based Dynamic Hand Posture Identification Using Genetic Algorithm", Springer, Machine Vision and Applications Vol. 11.

[28] Kouichi M., Hitomi T. (1999) "Gesture Recognition using Recurrent Neural Networks" ACM conference on Human factors in computing systems: Reaching through technology (CHI '91), pp. 237-242. doi: 10.1145/108844.108900

[29] Guan, Y., Zheng, .M. (2008). "Real-time 3D pointing gesture recognition for natural HCI. IEEE Proceedings of the 7th World Congress on Intelligent Control and Automation WCICA 2008, doi: 10.1109/WCICA.2008.4593304

[30] Freeman, W. T., Weissman, C. D. (1995). " Television Control by Hand Gestures". IEEE International Workshop on Automatic Face and Gesture Recognition.

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

[32] Shuying Zhao, Wenjun Tan, Shiguang Wen, and Yuanyuan Liu, (2008). "An Improved Algorithm of Hand Gesture Recognition under Intricate Background", Springer the First International Conference on Intelligent Robotics and Applications (ICIRA 2008),: Part I. pp. 786–794, 2008. Doi:10.1007/978-3-540-88513-9_85