

Machine Learning Algorithm for Recognizing Numbers and Symbols

Chia Fatah Aziz
Software Engineering
Firat University
Elazig ,Turkey

Lutfu Sabansua
Economics and Administrative Science
Firat University
Elazig, Turkey

Abstract— Hand Gesture Recognition (HGR) is a system that have gained a great and more powerful attention in the recent years. This is due to its useful application and the ability to contact with machine effectively based on the concept of Human Computer Interaction (HCI). This paper presents an approach for HGR system using a software tool. The developed system reads the real time image as an input and then it compares it with the training set samples of hand signs. In this approach, for detecting the threshold regions, skin detection technique has been used and the used hand gestures are taken for recognizing numbers and symbols in Iraqi sing language using Kinect camera that have a depth sensor. The result of presented paper is to create communication between the dump and normal people, and to understand the meaning of Sign Language between countries.

Keywords—Human Computer Interaction (HCI); Hand Gesture; Sign Language; Kinect;

I. INTRODUCTION

The biometric is a process of assessment and factual inspection of people's behavior and physical characteristics. Biometrics are used for authentications for ID and access control which shows that everyone is unique and is identified by physical and behavioral characteristic. The Gesture is a mechanism used for communication between machine and people. This mechanism requires interface of a hand gesture recognition system for easy human interaction.

In human computer interaction and computer vision field, Hand gestures are much important, and with the objective of being to bring the performance of human machine interaction close to human-human interaction they become some active in the area of research. Hand gesture recognition has diverse area of applications, such as communication in video conference [1], using a finger as a pointer for selecting options from a menu and Sign-language recognition [2, 3].

Hand gesture recognition is an important research issue in the field of human-computer interaction, because of its extensive applications in virtual reality, sign language recognition, and computer games. Despite much previous work, building a robust hand gesture recognition system that is applicable in real-life applications remains a challenging problem. Existing vision-based approaches are greatly limited by the quality of the input image from optical cameras [4, 5]. Consequently,

these systems have not been able to provide satisfactory results for hand gesture recognition. Hand gesture recognition faces two challenging problems: hand detection and gesture recognition. The Hand gesture implementation involves significant usability challenges, including fast response time, high recognition accuracy, speed of learning, and user satisfaction, helping to explain why few vision-based gesture systems have matured beyond prototypes or made it to the commercial market for human computer devices [6, 7].

Kinect is a highly precise motion sensing input device by Microsoft for the Xbox 360 video game console and Windows PCs. The device features an RGB camera, depth sensor and multi-array microphone running proprietary software which provides full body 3D motion capture, facial recognition and voice recognition capabilities. The depth sensor captures video data in 3D under any ambient light conditions. Microsoft Kinect provides an inexpensive and easy way for real-time user interaction in this regard [8].

A. Literature review

Human Computer Interaction (HCI), also called Man-Machine Interaction (MMI) , refers to the contact between the machines, exactly the computer, and the human; or between people to understand each other, since the computer is unimpressive without appropriate utilize by the user or human [9].

The fundamental aim for creating or making a system for Hand Gesture Recognition (HGR) is to make a natural association between human and machine or computer. This is clearly essential especially when the recognized gesture and symbols can be more important and have a useful way for conveying the meanings of full information for Arabic Sign Language (ArSL) between human and machine, which can also be used for robot control [10].

There are two types of gesture, namely, the static hand gesture which is considered less computationally complex and the dynamic hand gesture which is a sequence of posters and computationally more complex compared to its static counterpart; however, it is great and suitable for real time surroundings [11, 12]. There are some methods for recognizing hand gesture, some of which need hardware devising and implementation, for example data glove devices and color marker, to easily extract extensive characterization of features of gesture [13]. While some other methods are

based on the presence of the hand by using the color of skin to segment the hand and extract the required features [13].

There are many skills that hand gesture is important for in our life, such as, games, HCI, robotics and many more which usually use different algorithms and tools in their implementation [14, 15].

In [16] Trigueiros *et al.* used a machine learning algorithm and attend to which machine learning algorithm is suitable for real time hand gesture recognition. In their research, they used NB, KNN, and NN, where they applied all the methods twice on two different sets of gestures sign.

In [17] Wang *et al.* to segment hand gesture utilize skin detection in RGB color space, combining achieved color extraction and clustering of YCbCr color space. After color image extracting, they continued to edge detection and with the morphology processing fill the binary holes.

In [18] Zhang *et al.* using circular gradient edge detection technology to hand gesture segmentation and skin detection technique in HSV color space. Their presented method transformed RGB color space to HSV color space and in the HSV color space set color threshold. After the filtering they finished the hand detection and found those points that makes the hand region by utilizing the technology of circular gradient edge detection that contained and filled the holes of binary hand segmentation.

II. BACKGROUND

In this section, the relevant explanation needed for system development and implementation is presented.

A. Real Time Arabic Sign Language Recognition

In real time recognition based on image-based ArSL systems, there are three categories: the Alphabet, isolated word, and the continuous recognition. Image-based recognition system can be composed of five steps, namely, image acquisition, preprocessing, segmentation, feature extraction, and classification. The input for image based sign language recognition for Arabic language applications is a static image or a sequence of signs in a video. Here, the signer point is the space between one sign and another. An important advantage of image-based ArSL recognition system is that the user is accepted as a signer without using the data glove device. This contains lighting condition, and the background of an image, hand and face segmentation along with additive noise with each of these types. However, the segmentation of faces and hands are computationally uneasy, yet, the latest algorithms and computing abilities to perform this segmentation have made it possible in real time [19].

B. The Mechanism of Computer Vision Techniques Performing Hand Gesture Recognition

There are three layers of hand interactive systems which consist of detection, tracking and recognition. The task of detection layer is to define and recognize the appearance of objects in the view of the

camera(s). After detecting the objects in the view of the camera, the tracking layer uses multi-view inputs to track the objects in the camera's viewpoint. This tracking layer helps the approximation of the positions and features of the objects in the view of the camera. Finally recognition, which is the last layer, recognizes the objects. This layer uses the task results processed in the detection and tracking layers and groups those task results to different gesture labels.

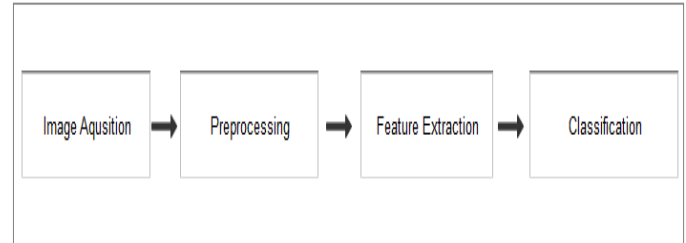


Fig. 1. System Overview

III. THE PROPOSED METHODOLOGY

The proposed method of this paper is to recognize digits from (1-9) and another symbols used in ArSL using supervised learning and libraries in depth data with Kinect sensor, shown in Figure 2, and a computer system with the standard PC, as well as a list box that allows the user to choose either PC camera or any USB camera. The used PC runs with 64-bit Windows 7 operating system and the used development environment is C# (Microsoft Visual Studio – 2015) provided with openCV 3.2.0 and EmguCV libraries.

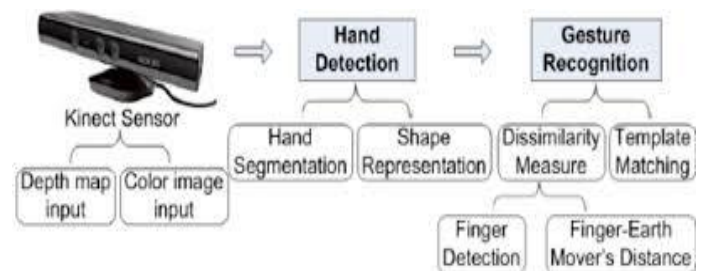


Fig. 2. Framework of the proposed gesture recognition system

The purpose of this paper is to explain the hand gesture learner machine for recognizing symbols. This aim can be elaborated in the following points:

- To recognize numeric and symbolic sign in Iraqi culture written with the hand in the air using machine learning algorithms and Kinect sensor.
- To observe how the used algorithms can be developed for hand gesture symbols recognition system.
- To use the hand gesture technique to recognize spoken messages in Iraqi culture with the depth sensor of Kinect camera.

A. The proposed system's implementation

The flowchart of the proposed system's implementation is shown in Figure 3.

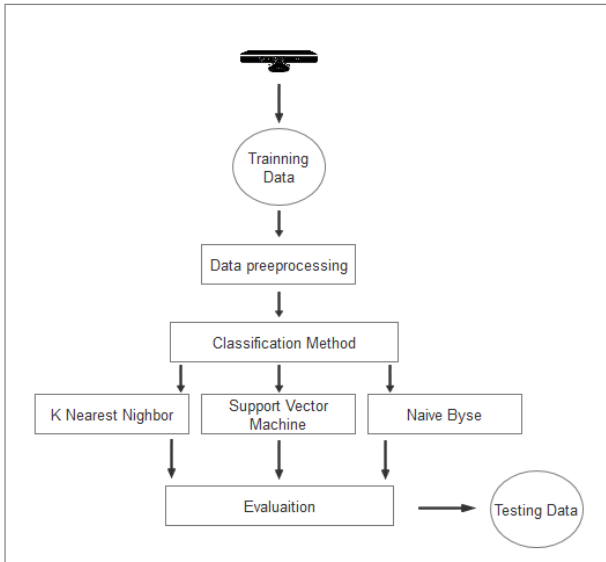


Fig. 3. The proposed system's implementation

According to the proposed method, the static image and a real video image are taken. These include a collection of captures of 200 body sign positions to recognize gesture of human using different classification method in data mining such as (SVM, NV, KNN) and then they are all compared to find the optimal classifier. Finally, the performance of each classification method in this study is computed.

The used signs and symbols in the study's data set computes (1-9) numbers and three other signs in the Iraqi culture as shown in Figure 4.

The setting is established as follows: first we use a single Kinect camera with the user stands in front of the camera and the distance between the camera and the floor is 1 meter and the distance between the human and the camera is around 1 meter as well.



Fig. 4. The tested gestures in the proposed system

B. Data Preprocessing

In this study, the preprocessing step uses a small neighborhood of a pixel in an input image to get a new brightness value in the output image. It consists of segmentation, and it is a process of converting or changing RGB image to binary image. This due to the generated noise, a Gaussian filtering is used to remove the undesired noise and also it help getting complete and filtered contour of gesture. Here, the output is a black and white image where its background is black and the hand is white which uses the steps like RGB to gray conversion thresholding and filtering, as shown in Figure 5.



Fig. 5. Converting RGB image to Gray

C. Data Training and Testing

In this experiment, all classification methods in data mining used in this application such as (KNN, SVM, NB) are used to train the system for recognition or classification of human gestures for (1-9) numbers and three other signs in the culture of Iraq. To obtain training data, the human hand should be kept stable in front of the camera to capture the hand image as illustrated in Figure 6.

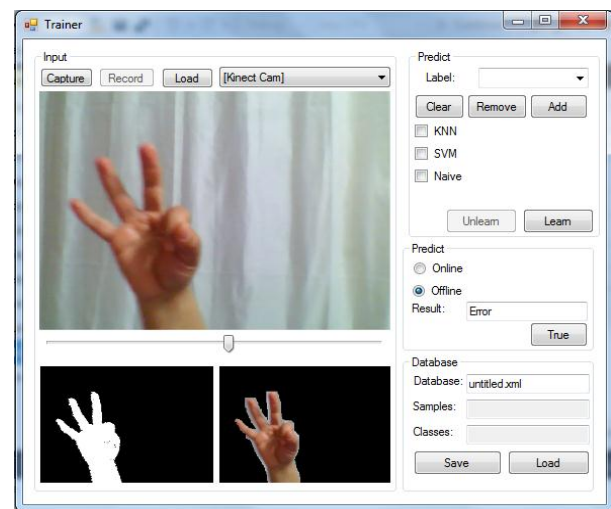


Fig. 6. Data training and testing procedure

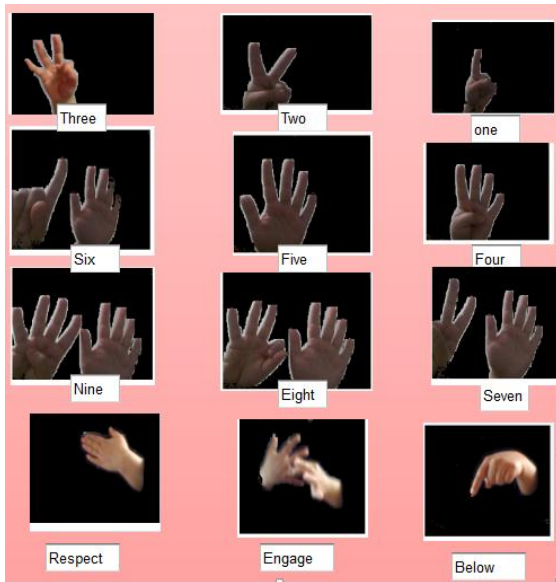


Fig. 7. Examples of the 12 different gestures to be recognized

After the training data is obtained, another process is started, that is, the recognition steps using the classifier method. All training imaged information is saved in the xml database, where the recorded and captured images are add to those which already in the database with proper labels. The images are recorded from different individuals in order to achieve the final recognition process with different data of hand gesture. Figure 7 shows the different training samples of the hand.

D. Hand detection

By using preprocessing steps and filtering techniques, the image become equal in intensity and more noticeable. The EmguCV library that is used to detect an image from a captured image found in the training data. From the first step of the recognition process, detected images are achieved in real time obtained from a Kinect camera. The detection of the hand is done automatically using the method of skin detector for image color space (RGB, HSV, YCbCr). In detection process, the hand has the white color and the background has the black color, as shown in Figure 8.



Fig. 8. The detected image

E. Hand recognition

In this study, three methods (SVM, KNN, NB) are chosen for all classification process. These methods are chosen based on the review of the literature which indicates that each of these selected methods for classification purpose is effective to hand gesture recognition and sign language recognition applications. The recognition process comes after the detection of hand gesture. After capturing the hand image in front of the camera, accomplishing the preprocessing and

running a complete detection, one of the classification methods such as SVM, KNN and NB is chosen; the learning process of the application for the hand sign recognition is done. Once a new capture is taken again, the hand sign that is captured is then processed in the application with the machine learning algorithm in order to find any match in the database. A true sign is predicted in online and offline basis. In the basis of offline predict, you have an error first, and if your sign matches to the image learned by the machine learning methods then write the gesture name in the predict textbox, and the true sign added to the XML database and don't match then type error instead of the sign name that show to you in the figure (9).

The accuracy and ability of each method of KNN, SVM, NV are vary between %90 - %95, under the same condition of Kinect camera and distance between human and camera, and good light ambience condition. In this application, a USB camera can be used instead of Kinect camera and different distance, any conversion in the ambient light, but changes of that points changes the percentage of accuracy recognition of each method. Meaning that they're negatively affected by the conversion with the normal conditions. Table I shows the accuracy results of each method for two numbers and two sign gestures of hand in the culture of Iraq.

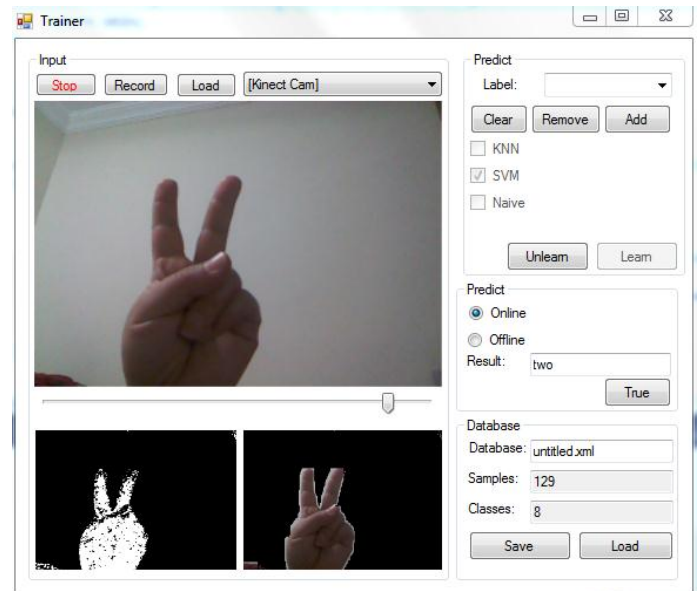


Fig. 9. Hand recognition process

Methods	The Percentage Accuracy of Tested Gestures			
	<i>Gesture one</i>	<i>Gesture Two</i>	<i>Gesture Below</i>	<i>Gesture Respect</i>
KNN	%92	%90	%93	%90
SVM	%90	%92	%90	%93
NB	%90	%94	%92	%94

Table [1] Accuracy rate of tested Gestures

IV. CONCLUSION

This paper presented the vision based automatic sign language recognition for Arabic numbers and symbols in the culture of Iraq. It strived to have a methodology that can help in making an effective software for recognizing Arabic sign,

and helping all other people in the world to learn what are the meaning of Arabic signs and letters by your hands. If all propose features and methods of machine learning like SVM, KNN, NB goes well, we could expect that numeric and symbolic sign shown by the hand in the air using Kinect sensor and this sensor can be recognized accurately with machine learning algorithm using hand gesture. Several hand gestures were properly classified with the concepts of the proposed software. The purposed application can be further developed to enable the communication between people with special needs and normal people as well as to further implement more signs from Iraqi culture and/or other cultures.

REFERENCES

- [1] S. Askar, Y Kondratyuk, K. Elazouzi, P. Kauff, O. Schreer, "Vision-Based Skin-Colour Segmentation of Moving Hands for Real-Time Application," 1st European CVMP, pp. 79-85, 2004.
- [2] X. Zhu, J. Yang, A. Waibel, "Segmenting Hands of Arbitrary Color," Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 446-453, March 2000.
- [3] N. Soontranon, S. Aramvith, T. H. Chalidabhongse, "Improved Face and Hand Tracking for Sing Language Recognition". ITCC 2005, pp. 141-146 Vol. 2, April 2005.
- [4] G. R. S. Murthy and R. S. Jadon, "A review of vision based hand gestures recognition," Int. Journal of Information Technology and Knowledge Management, vol. 2, no. 2, pp. 405-410, 2009.
- [5] C. Chua, H. Guan, and Y. Ho, "Model-based 3D hand posture estimation from a single 2D image," Image and Vision Computing, vol. 20, no. 3, pp. 191-202, 2002.
- [6] B. Stenger, A. Thayananthan, P. H. S. Torr, and R. Cipolla, "Filtering using a tree-based estimator," in Proceedings of the IEEE International Conference on Computer Vision, vol. 2, pp. 1063-1070, Nice, France, October 2003.
- [7] chs, M. Kolsch, H. Stern, and Y. Edan, "Vision-based " hand-gesture applications," Communications of the ACM, vol. 54, no. 2, pp. 60-71, 2011.
- [8] and A. Sugimoto, "Image categorization and semantic segmentation using scale-optimized textons," IT CoNvergence PRActice (INPRA), vol. 2, no. 1, pp. 2-14, 2014.
- [9] Fakhreddine Karray, Milad Alemzadeh, Jamil Abou Saleh, Mo Nours Arab, (2008) ."Human- Computer Interaction: Overview on State of the Art", International Journal on Smart Sensing and Intelligent Systems, Vol. 1(1).
- [10] 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-
- [11] 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). Department of Computer Science. The University of Tennessee Knoxville.
- [12] Simei G. Wysoski, Marcus V. Lamar, Susumu Kuroyanagi, Akira Iwata, (2002). "A Rotation Invariant Approach On Static-Gesture Recognition Using Boundary Histograms And Neural International Journal of Artificial Intelligence & Applications (IJAA), Vol.1.3, No.4, July 2012 173 Networks," IEEE Proceedings of the 9th International Conference on Neural Information Processing, Singapura.
- [13] 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.
- [14] 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.
- [15] R. Azad, B. Azad, and I. T. Kazerooni, "Real-time and robust method for hand gesture recognition system based on cross-correlation coefficient," Adv. Comput. Sci., Int. J., vol. 2, no. 5/6, pp. 121-125, Nov. 2013
- [16] P. Trigueiros, F. Ribeiro, and L.P. Reis. A comparison of machine learning algorithms applied to hand gesture recognition. Information Systems and Technologies (CISTI), 2012 7th Iberian Conference on, pages 1-6, 2012.
- [17] X. J. Wang, G. Z. Bai and Y. M. Yang, "Hand Gesture Recognition Based on BP Neural Network in Complex Background", Computer Applications and Software (in Chinese), vol. 30, no. 3, (2013), pp. 247-249.
- [18] G. J. Zhang, D. W. Zuo, X. F. Li and C. H. Shi, "Design and Development of a New Gesture Segmentation Algorithm Based on Circular Gradient", Machine Design and Manufacturing Engineering (in Chinese), vol. 42, no. 9, (2013), pp. 1-6.
- [19] P. Vamplew, A. Adams, Recognition of sign language gestures using neural networks, Australian Journal of Intelligent Information Processing Systems 5 (2) 94-102, 1998.