International Research Journal of Computer Science (IRJCS)
Issue 04, Volume 6 (April 2019)

ISSN: 2393-9842
www.irjcs.com

ALPHABET HAND GESTURE RECOGNITION USING HISTOGRAM OF ORIENTED GRADIENTS, SUPPORT VECTOR MACHINE AND K – NEAREST NEIGHBOR ALGORITHM

Anna Liza A.Ramos*
Saint Michael's College of Laguna
Binan City, Laguna, Philippines
annakingramos@yahoo.com.ph;

Glydle Dee M.Dalhag
Saint Michael's College of Laguna
Binan City, Laguna, Philippines
dalhagglydledee@gmail.com;

Mark Limuel D.Daygon
Saint Michael's College of Laguna
Binan City, Laguna, Philippines
marcoslimuel23@gmail.com;

Jhon Omar K.De La Cruz
Saint Michael's College of Laguna
Binan City, Laguna, Philippines
jhonomardelacruz@gmail.com

Abigail A.Macaranas
Saint Michael's College of Laguna
Binan City, Laguna, Philippines
macaranasabigail@gmail.com

Kim LJ V.Almodovar
Saint Michael's College of Laguna
Binan City, Laguna, Philippines
almodovarkimlj@gmail.com

Manuscript History
Number: **IRJCS/RS/Vol.06/Issue04/APCS10095**
Received: 13, March 2019
Final Correction: 21, April 2019
Final Accepted: 26, April 2019
Published: April 2019
Citation: A.Ramos, M.Dalhag, D.Daygon, Jhon, A.Macaranas & V.Almodovar (2019). Alphabet Hand Gesture Recognition Using Histogram of oriented gradients support Vector Machine and K-Nearest Neighbor Algorithm IRJCS:: International Research Journal of Computer Science, Volume VI, 200-205. **doi://10.26562/IRJCS.2019.APCS10095**
Editor: Dr.A.Arul L.S, Chief Editor, IRJCS, AM Publications, India

Abstract— Sign language is a tool used by deaf for communication. They use different gestures to express their thoughts and communicate to people. Each gestures or movement of the hand has a special assigned meaning. Previous studies were conducted on hand gestures using Kinect camera, detection gloves, and leap motion controller to improve the accuracy in detection of the implied meanings. This study aims to develop a model that would increase the accuracy rate of detection, using a customized camera that addresses the background and lighting conditions. Moreover, the study will provide learning to those people who are not familiar in sign languages. The model were tested to 28 gestures together with the forming of words, it acquired an accuracy rate of 94.49%.

Keywords— K- Nearest Neighbour Algorithm; Hand Gesture; Oriented Gradients; Vector Machine;

I. INTRODUCTION

A gesture is defined as a lively movement of hands and creating signs with them such as alphabets, numbers, words, and sentences. It can be classified into two types, the static and the dynamic gestures. Static gesture refers to a certain pattern of hand and finger orientation while dynamic gestures involve different movement and orientation of hands and facial expressions largely used to recognize a continuous stream of sentences [1]. The purpose of gesture recognition is to enable humans to interact with the human-made machines in a natural way without any mechanical devices and the mathematical equations [1]. The gesture recognition is a very difficult and complex task since the full recognition system should be able to identify the hand in different scales, positions, orientations, contrasts, luminosity, and others. Sign Language which is an effective tool for people who are not able to speak or hear anything [2]. Deaf people use these different gestures to represent their thoughts [3]. Each gesture or movement of hand, facial expression, and body movement has a special assigned meaning [4].

There are different sign language used in different part of the world. It depends upon the language they spoke and the culture of that particular place. For example in the USA, American Sign Language (ASL) is used while in England, British sign language is used (BSL). Similarly the Japanese sign language (JSL), and Filipino sign language (FSL). As of the year 2009, a total of about 517,000 people are part of the deaf and mute community in the Philippines. The deaf and mute communities more often than not lack education in the proper use of sign language. Looking into this specific community in the Philippines. The absence of proper training and knowledge on sign language results to individuals creating their own signs to communicate with their family. They exclaimed that the creation of their own sign language results from these factors: (1) illiteracy of both individuals with disabilities and their families, (2) lack of education and support, (3) poverty and all the problems that fall under this factor. As for those individuals who are able to overcome the factors mentioned above, they are educated with the most popular type of sign language, which is again the American Sign Language [5].

A. Background of the Study

The study reviewed various studies on hand gesture recognition accuracy results. A real time hand gesture recognition was proposed using Histogram of Oriented Gradients (HOG) descriptors and able to achieve a reasonable degree of accuracy of detection with performance degradation because of positional variation [6]. Developed a sign language application based on hand gesture recognition for deaf people using background separation algorithm for detection and K-NN method for classification with an 80 percent of accuracy and can further improved the tracking methods by using model based tracking [7]. Works on hand gesture recognition using leap motion controller and Kinect devices using SVM for classification method. With the use of leap motion controller and Kinect devices they achieve a high accuracy but the problem is it cost [8]. Utilize colour code gloves to track hand movement and achieve high detection accuracy. But the problem is the tracking method requires the user to wear coloured gloves which reduces the user experiences [9]. Developed a system that recognize alphabets characters using two feature vectors namely HOG and SIFT, and also used SVM to construct a training model which will further classify the tests image.

From 5 test image, 80% accuracy obtained in V alphabet and 100% accuracy from the image other than hand gesture. There is no full – fledged conclusion because there are no implemented algorithms for all the alphabets [10]. Have proposed the hybrid approach using SIFT and HOG combined as a feature extraction technique and SVM as a classifier and the recognition that is achieved for the 10 sets of gestures is 97 percent and can further improved by applying more deep learning methods and feature extraction methods for better accuracy [11]. Present a novel static gesture recognition method based on non-negative matrix factorization (NMF) and compressive sensing (CS) using local binary pattern and Histogram of Oriented Gradients with an accuracy rate of 98.13 percent but only works on 3 different background uniform light and dark, and as complex as their database [12]. The studies mentioned above provide insight to discover the best performing algorithms for specific used. Also to address some limitations and challenges of the existing works such as background restrictions, the problem in wearing of gloves and use of some Kinect devices to determine the detection. With this, the study will resolve all of the mentioned issues by using Logitech C270 web camera with film filter by removing an IR filter from the web camera then replace with IR illuminator to work in various type of background.

B. Research Objectives

1. To build a hand gesture database based on Filipino Sign Language (FSL).
2. To utilize best algorithm like Histogram of Oriented gradient for detection, SVM with localize contour for classifying the gestures and KNN for recognition and also to interpret sign language into written words.
3. To build a prototype that will recognize the hand gesture, interpret and transform into words.

C. Model of the Study

The figure 1 shows the model of the study with 6 Subjects, that has been captured their hand gesture with classifications- Big hand, Small hand and long fingers in 8 different angles using Python, Open CV and Logitech C270 camera with film filter and I.R Led Light that will help to reduce the noise of the background. These Hand gesture will be detected using the HOG (Histogram of Gradient) algorithm. HOG used in image processing for the purpose of object detection, for instant discussion the captured image will be resized to 200x200 and converted to grayscale, same filter to android phones camera. Here is the code: gray=cv2.cvtColor (img, cv2.COLOR_BGR2GRAY) for the conversion of RGB to grayscale. The first step in HOG is calculating the Hog descriptor and compute and normalize the gradient vector to get the pixel value of the image. After that present the value of x-direction and y-direction, to get the right value minus the right value into the left value and the answer will the x-direction. Then to get the left value(y-direction) minus the left value into upper value of the Gradient Vector. For the last process rather normalizing individually, cells are grouped into blocks and normalized it all. For the main discussion with formula of the HOG will be displayed later. Last, after the training or feature extraction the image will be classified into 8 hand gesture with different angle per letter using Support Vector Machine (SVM) ; with the help of K-nearest neighbors(KNN) algorithm that classify the closest training example in the problem space. These trained datasets will be stored in Hand gesture datasets.

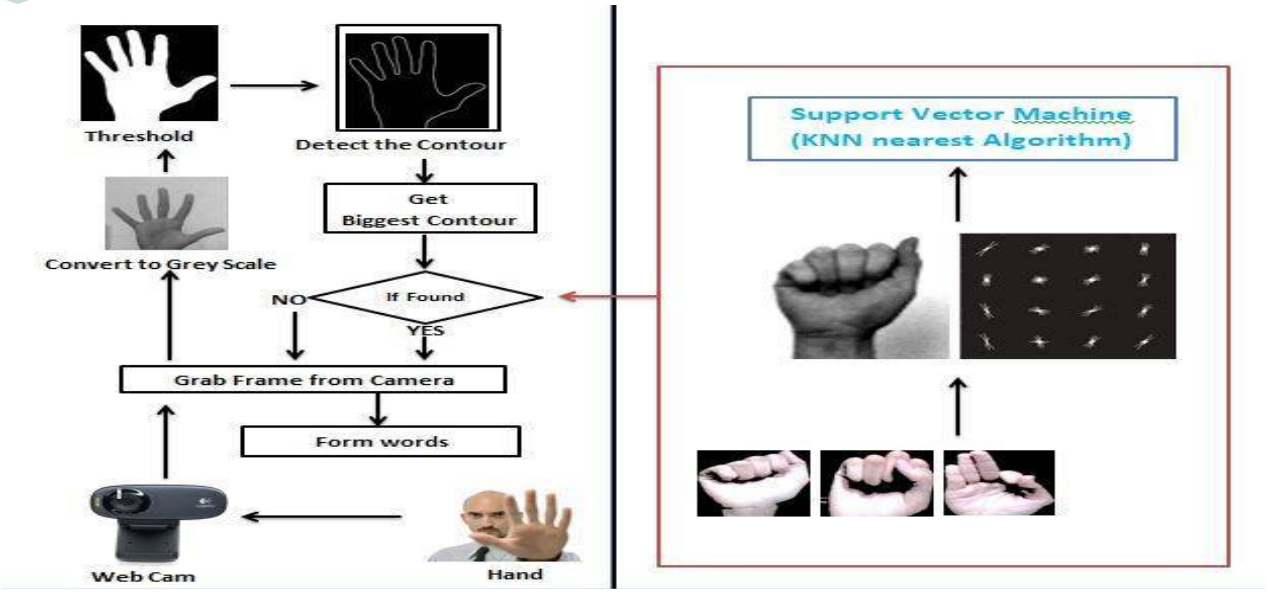

Fig. 1 Model of the Study

When the user start the real-time video streaming the hand gesture will be detected using the Histogram of Gradient Algorithm and convert into greyscale. Next step is Thresholding using the Otsu's method, in this process it identify the two classes of two pixels like foreground (hand detected) pixels and the background pixels. To identify these two classes get each variance value of the each classes by getting the total value of the Weight and Mean of each class. After getting the variance value of each classes it will compute the value of WCV (Within-Class Variance) that simplify the sum of two variance multiply by their associated weights. Then get the biggest contour which is the hand gesture using the Open CV function find Contours. After that get the convex Hull functions that calculate the center of the hand and the number of finger that can help the KNN classify the data. After classifying the detected hand gesture it will label the meaning of the hand gesture and form words.

II. METHODOLOGY

A. Building the hand gesture datasets

The study built the hand gesture standards by selecting six (6) subjects based on the size of the hand. These subjects were requested to project symbols from A to Z based from the Filipino Sign Language.

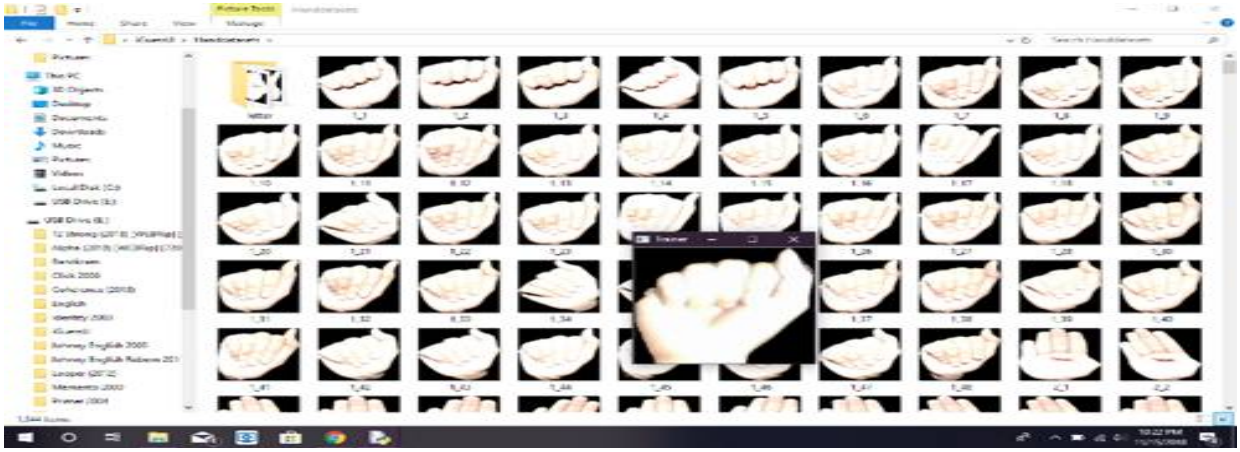

Fig. 2 The training datasets of the study

Every subject projected the 28 hand gestures twice with a total of 336 gestures. These gestures served as the training datasets for this study.

B. Hand Gesture Features Detection

The hand gesture will be detected using Histogram of Oriented Gradients and will be captured by 200 pixels wide and 200 pixels tall. The captured image will be converted to a grey scale that can help in calculating the Histogram of Oriented Gradients descriptor. In calculating the Histogram of Oriented Gradient to pixels, the image is divided into 8X8 cells. Within a cell, it will compute the gradient vector or edge orientation at each pixel. Then it will normalize the gradient vector and simply measures the changes of pixel values along x and y-direction around each pixel.

Rather normalizing each histogram individually, the cells are grouped into blocks and normalized it all. The first block will be 2X2 then the second block will be 50 percent overlap. The algorithm used in this system was proved to be the best-suited based on the environment and needed functions in this study.

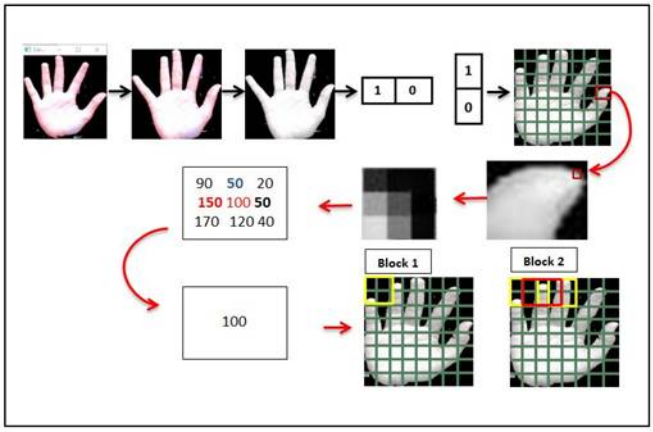

Fig. 3 Hand Gestures Features Detection

C. Hand Gesture Features Extraction

The greyscale image will be converted to monochrome by the process of threshold using the Otsu's method. Only 6 greyscale levels are used in the image. To identify the variance of the foreground, first we will compute the Weight and Mean. The total Weight of the foreground is 0.1667, and the total Mean is 1 then the total Variance is 3.

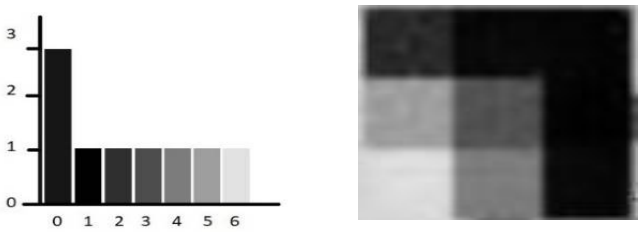

Fig. 4 Extraction of Hand Gestures

D. Computing the Contour and Convex Hull

In computing the contour and convex hull, the Open CV function "find contours" and "find convex hull" is used. From there it will get the biggest contour and construct the bounding rectangle of the convex hull. The rectangle is used to calculate the center of the finger and the center of the hand. From this it might help the KNN to easily classify the data.

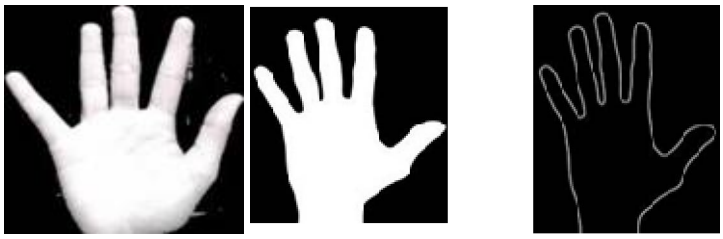

Fig. 5 Computing Contour and Convex Hull

E. Hand Gesture Classification

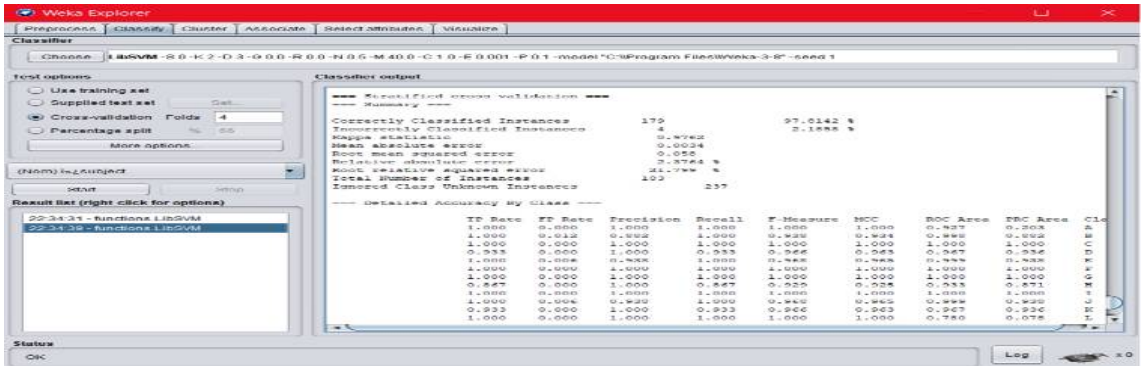

Fig. 6 Classification of hand gestures using Weka

The study classified the hand gestures using the Support Vector machine with help of KNN algorithm for recognition. Support vector machine classifies different gestures based on their feature vectors.

F. Hand Gesture Recognition

To recognize the hand gestures, the K – Nearest Neighbor algorithm was used by identifying the value of the K. The K is always a positive integer. The determined K or the number of nearest data were calculated using the distance between the query instance and all the training sample. It was then sorted for all the training samples and determined the nearest data based on K-TH.

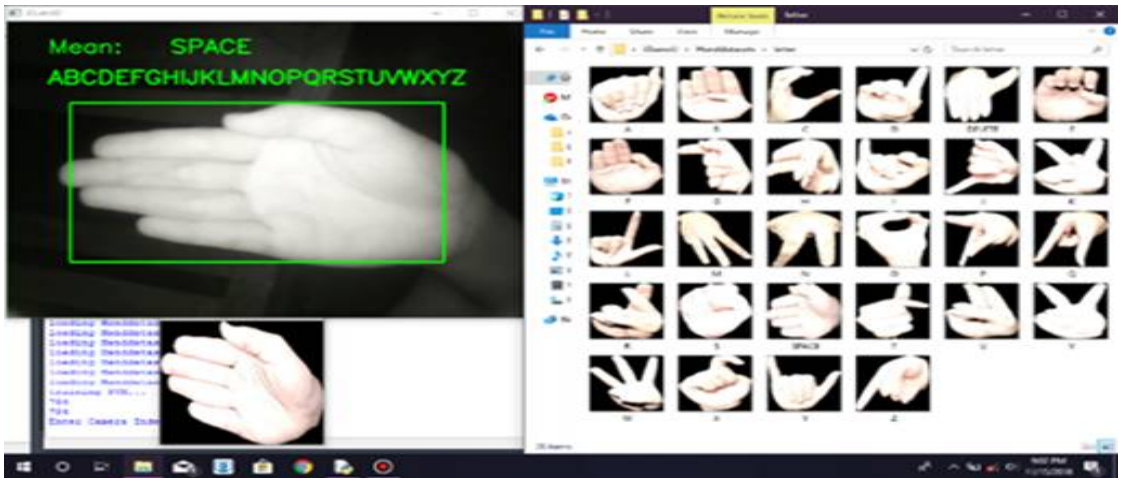

Fig. 7 Recognition of Hand gestures

G. Metrics of Evaluation

As for the recognition, the study used the accuracy formula that was equal to the number of detected gestures over the number of testing gestures. The study achieved a 94.49% accuracy rate by choosing the Leave One-Out Cross Validation – this is computed with help of Python's Scikit-Learn Library.

III. RESULTS AND DISCUSSION

The results show the accuracy performed by the six (6) subjects who projected the 28 hand gestures. It shows that the approach performed well and is fit for the real-time applications though there were some hand gestures with low accuracy results due to distance, positional variance and incorrect hand gesture projection of the users.

H ---- N, I ---- S, N----M/Q, E-----S/B, G---Z, U---L/S, R---D/U, O---E, K--W/V,

These were the hand gestures that were often detected incorrectly by the model. For example, the gesture letter N would often detect to be a gesture for letter H. However, the study resulted a rate of 94.49% accuracy rate. The figure above shows that Leave One-Out Cross Validation achieved better accuracy than the 10 folds cross validation due to the small number of datasets that the study had.

IV. CONCLUSIONS

In this paper the researchers presented a hand gesture recognition system based on sign language using Histogram of Oriented Gradients for detection, Support vector machine as a classifier and K Nearest Neighbor algorithm for recognition. The proposed model can recognize 26 letters from the Filipino sign language and with an additional 2 gestures for space and delete function and can form words, phrases, or even statements with a 94.49 percent of accuracy. In future work, the researchers would like to recommend

ACKNOWLEDGMENT

The researchers would like to thank the panelist Dr. Lorena W. Rabago, Dr. Meryli Rosas, Prof. Maria Art Antonette Clarino, and Prof. Raquel Bermudez. Also, Ms. Jazmin Tan for reviewing the grammar and sentence structures of this manuscript. Likewise, the researchers would also like to give thanks to their classmates and schoolmates who participated and appreciated this study – providing their valuable inputs along the entire process of this study.

REFERENCES

1. M. M. Hasan, and P. K. Mishra, "Robust gesture recognition using gaussian distribution for features fitting," International Journal of Machine Learning and Computing, vol. 2, no. 3, pp. 266-273, June 2012.
2. P. Pandey and V. Jain, "Hand gesture recognition for sign language recognition: A review" International Journal of Science, Engineering and Technology Research (IJSETR) Volume 4, Issue 3, March 2015
3. G Adithya V., Vinod P. R., Usha Gopalakrishnan , —Artificial Neural Network Based Method for Indian Sign Language Recognition||, IEEE Conference on Information and Communication Technologies (ICT) ,2013, pp. 10801085.

4. Rajam, P. Subha and Dr. G Balakrishnan, "Real Time Indian Sign Language Recognition System to aid Deaf and Dumb people", 13th International Conference on Communication Technology (ICCT), 2011, pp 737-742.
5. S. M. Pascua, PL. C. Espina, RP. E. Talag, L. N. Villegas, and L. A. De Guzman "Words in Vision: A Filipino Sign Language Thesaurus Management System Using Ren-py" International Federation of Library Associations and Institutions
6. A. Takashi "Hand Gesture Recognition using Histogram of Oriented Gradients and Partial Least Squares Regression" MVA2011 IAPR Conference on Machine Vision Applications, Nara, JAPAN
7. Shalini (September 2015) "Real Time Control System Based on Hand Gesture Detection and Recognition" International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064 Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438 Volume 4 Issue 9,
8. G. Martin, F. Dominio and P. Zanuttigh. (2015). "Hand gesture recognition with jointly calibrated leap motion and Depth sensor" Multimedia Tools and Applications, 1–25.
9. H. Cooper, B. Holt, and R, Bowden. (2011) "Sign language recognition in Visual Analysis of Humans" Eds. Springer, Oct., ch. 27, 539 – 562.
10. Parul, Hardeep, "Neural Network Based Static Sign
11. Gesture Recognition System" International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE), Vol. 2, Issue 2, pg. 3066 -3072, 20
12. F. J. Ansari (September 2017) "Hand Gesture Recognition using fusion of SIFT and HOG with SVM as a Classifier" International Journal of Engineering Technology Science and Research IJETSR ISSN 2394 – 3386 Volume 4, Issue 9
13. H. Zhuang, M. Yang, Z. Cui, Q. Zheng (Advance online publication: 22 February 2017) A Method for Static Hand Gesture Recognition Based on Non-Negative Matrix Factorization and Compressive Sensing, IAENG International Journal of Computer Science, 44:1, IJCS_44_1_07
14. Shariz, S. and Kulkarni, A. "Identifying Hand Configurations with Low-Resolution Depth Sensor Data", CS229 Final Project Paper, 2009.

