

DATASET FOR AMERICAN SIGN LANGUAGE

Aliya Gohar, Mrs. Shazia Saqib, Rabia Shakeel, Rimsha Naheed

Abstract: - For the deaf and dumb people, sign languages are the only path of communication. With the help of sign languages, physically disabled people can convey their feelings, emotions, and thoughts to other people. Since, for the common person, it is very complicated to understand these languages, these physically disabled persons are dependent on a, who interacts with the world to convey their thoughts and feelings. For the production of these sign languages, it was necessary to develop an efficient dataset. With 26 English alphabetical hand gesture images. Further, segmentation and classification are applied to datasets. This paper provides guidelines for the creation and selection of datasets.

Keywords: Datasets, hand gestures recognition, ASL, HCI and Feature Extraction

1. INTRODUCTION

Sign languages are the native languages for the deaf and dumb communities worldwide. The deaf community ratio is in minority in their world whereas dumb community is the majority in ratio. These two communities typically have any signing skills, there is an interest to build a communication path for both the communities. In sign language recognition, an automatic system extracts sign language from an image or video and represents sign in a written form, which then translated into a written text of a spoken language. Drawing on the emerging in the computer vision the simple method is found to be the

most robust one when tested on large datasets. We have chosen to use **datasets** to describe a collection of images used by researchers in some domain. The image dataset is based on **American Sign Languages** hand gestures. Although there are many hand gestures datasets available, there are some characteristics that differentiate our work from other datasets. Firstly, the images cover a variety of hand gestures using different illumination conditions. Secondly, the images are cropped and segmented from the original captured images, allowing researchers to test their own combinations of feature extraction method, (e.g. edge detection and

binary images). Thirdly publically-available datasets for hand gestures may be used, there is no need to use any special gloves or any other equipment while recording these gestures. In the data collection, the wrist cover was only used to improve the quality of the color segmentation but does not mean that wrist

Table 1: Publically Available Data Sets

| Description | URL | References |
|---|--|------------|
| Various Hand and Gesture Datasets | www.idiap.ch/resources/gestures | [5-8] |
| Gesture Recognition Database Image | www-prima.inrialpes.fr/FGnet/data/12-MoeslundGesture | [9] |
| Pointing 2004 Gesture Recognition Dataset | wwwprima.inrialpes.fr/Pointing04/datasets.html | [10, 11] |
| Two-handed datasets | http://www.idiap.ch/resources/twohanded/ | [12] |
| RWTHBOSTON-104 Database Video | www-16.informatik.rwthachen.de/~dreuw/dataset-rwth-boston-104.php | [13, 14] |
| Hand image dataset (ASL, rendered) | http://www.cs.bu.edu/groups/ivc/data.php | [4] |
| ASL video sequences | http://www.bu.edu/asllrp/ncslgr.html | [15] |

cover is considered to be compulsory for the development of recognition algorithm.

ASL and hand shape real-time recognition system is based on comparisons on images stored in the database. To create an effective **database** following are the requirements: naturality of gesture dataset, size, and shape of the set, an accurate analysis of the potential effects of the recording condition and an accurate description of an acquisition process.

2. LITERATURE REVIEW

In 2005, a former computer vision research group at IIMS created an image dataset [1]. The latest image dataset was produced containing all standard ASL (American Sign Language) gestures. These datasets have the limited number of gestures and postures, most of these do not have the full image available. The most common datasets are listed in table 1 [1].

Christopher Lee and Yangsheng Xu developed a glove based recognizing a minimal subset of a full sign language that hardly recognized the 14 of the letters from the hand gestures alphabetic datasets [8]. Since from the years, latest glove devices have been designed i.e. power glove, Sayre Glove, and Dexterous Hand Master. The most successful available glove is by far the VPL Data-Glove as represented in a fig.1.

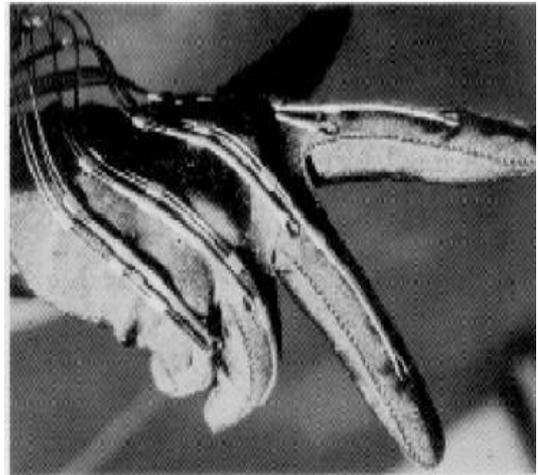


Figure 1: Data Gloves

VPL Data Glove was introduced by Zimmermen in the year 1970's [8]. Kalsh et

al specify one of the famous standard datasets for hand postures as hand sign languages [9]. Badi et al express human sign language as one of the convenient and effective communication paths for deaf and dumb people [9]. Sign language is one of the visual languages communicated using facial expressions, hands and arms gestures and using lip signs. Sign language understanding and identifying the process involved in translating these movements and action into meaningful speech and sentences. Sign languages are not global, these languages are region specific giving rise to different forms such as American Sign Language, Indian Sign Languages, Chinese Sign Language, German Sign Language, Argentinian Sign Language. **Sen et al** specify that all 26 alphabets in English vocabulary can be a unique hand gesture dataset [9]. Recognition process involves extracting the meaningful features from these gestures and interprets the specific meanings and finally generates tone or text message. **KTH** dataset is a dataset with six actions. These six actions are hand clapping, hand waving, walking, running, jogging, and boxing [9]. These six different actions are executed by 25 subjects in four different environments such as indoor, outdoors with scale variations, outdoors with different clothing and outdoors [9]. A research dataset for action recognition shows a domain at the

frontier between action and gestures. Gesture datasets are also moving away from research and new spread to the commercial market. For example, ARB lab has recently been running their company based on a gesture dataset and the related acquisitions software [11]. 3D hand tracking datasets along with hand shapes, movement trajectories are necessary component of signs, orientation, and computer-based recognition pattern for automatic sign recognition [12]. In observation, most of the datasets have been created for internal projects then released publicly. However, datasets carefully and specifically designed for comparison purposes of interest in the research community.

3. MATERIALS AND METHODS

Hand gesture recognition has four essential components i.e. hand detection, hand gestures, hand tracking and hand segmentation [10]. These four components of various processes as hand detection include background subtraction followed by skin segmentation and feature extraction to detect the user's hand for input. We differentiate three types depending on the problem. We target hand shape recognition, sign recognition and sentence recognition. The methodology of ASL recognition system comprised of stages i.e. Creation of datasets,

preprocessing, edge detection, feature extraction, gesture modeling and classification [13]. Since, gestures are considered as dynamic in nature that is more complex, but suitable for the real-time environment. Variety of methods has been proposed for acquiring information mandatory for gesture recognition system. A gesture can deliver information which is usually qualified as spatial symbolic information and effective information. The work on the dataset is to identify the symbolic information displayed as hand gestures. The ASL database used for identification consists of 26 English alphabets as shown in fig 2.



Figure 2 Static Alphabets of ASL

a. Description of Datasets

To make translator accessible through the simple web- app and a laptop with a camera, color images are selected. 26 English alphabetic images by using

different hand gestures that used to create an effective dataset. Images of 1920×2560 dimensions have been selected for the creation of datasets. The hand’s gestures images are cropped with no negative space. For further requirements, segmentation and classification could be applied to these color images. We have mainly focused on the static hand gestures datasets. Here are some multiple images of different gestures used for the creation of the effective datasets.



Figure 3 Multiple Datasets for ASL Gesture A

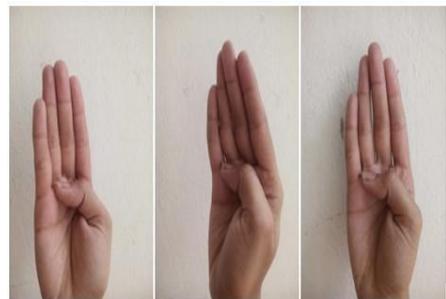


Figure 4 Multiple Datasets for ASL Gesture B

Table 2 represents the clear on video based research level datasets for recognition [6].

Table 2: - List of Datasets

| Name | Classes | Subjects | Samples | Language level | Availability |
|----------------------|---------|----------|---------|----------------|----------------|
| DGS Kinect 40 [1] | 40 | 15 | 3000 | Word | Contact Author |
| DGS RWTH-Weather [8] | 1200 | 9 | 45760 | Sentence | Public Website |
| DGS SIGNUM [8] | 450 | 25 | 33210 | Sentence | Contact Author |
| GSL 20 [1] | 20 | 6 | 840 | Word | Contact Author |
| Boston ASL LVD [3] | 3300+ | 6 | 9800 | Word | Public Website |
| PSL Kinect 30 [2] | 30 | 1 | 300 | Word | Public Website |
| PSL ToF 84 [2] | 84 | 1 | 1680 | Word | Public Website |

In general, the datasets that are video-based depend on segmentation and skin color tracking and are therefore not robust for background variations and interpreter cloths as well as hand-hand or hand face occlusions [6]. To store the datasets it is needed to create a database. Here I will represent the Massey’s image database containing a number of hand postures and hand gestures [7].

Table 3: Examples of Few Datasets

| Dataset | Lighting Condition | Background | Size | Number of Image | |
|---------|--------------------|----------------------------|-----------------|-----------------|-----|
| 1 | Hand gesture | Normal | Dark background | 640*480 | 169 |
| 2 | Hand gesture | Normal | RGB(0,0,0) | Varying/Clipped | 169 |
| 3 | Hand palm | Normal | Dark background | 640*480 | 145 |
| 4 | Hand palm | Normal | RGB(0,0,0) | Varying/Clipped | 145 |
| 5 | Hand palm | Artificial light/Dark room | Dark background | 640*480 | 498 |
| 6 | Hand palm | Artificial light/Dark room | Dark background | Varying/Clipped | 498 |

b. Creation of Database

The image database is giving rise to a new and challenging problem for the researchers. Since database technology from past forty years has played a vital role in the production of object-relational databases and relational databases [14]. Our database contains about 120 Images and dimension 1920×2560 composed of

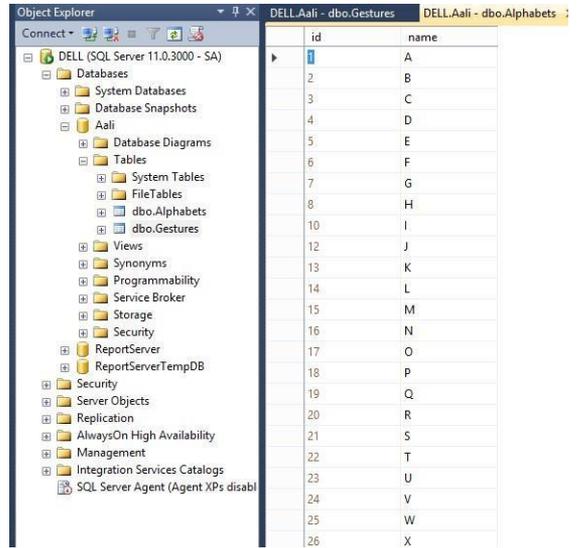


Figure 5: The American Sign Language snapshot

American Sign Language. It is necessary to mention here, that large size database provides more efficiency as it increases the recognition percentage. MS SQL is used as a tool server for the creation of the database. Multiple pictures of different gestures are stored in the database. The screenshots of the database are given

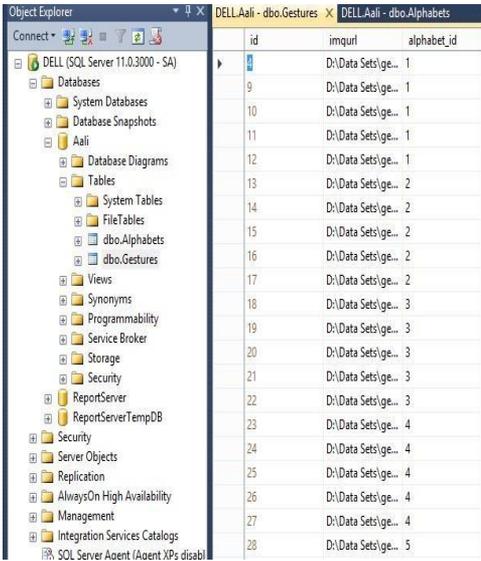


Figure 6: The American Sign Language

below. The snapshot (a) represents the 26 English alphabets stored in the SQL server database and Snap shot (b) represents the stored path of the different hand gestures buying the image URL technique, for the further matching of datasets.

After the successful creation of database, we connected the database on MATLAB for the matching purpose.

4. Complications to hand Gestures Recognition for datasets: Feature Extraction and Extraction Method

Feature extraction is the part of the data reduction process and is followed by

feature analysis. The basic and most important is analysis while determining exactly which features are important. The basic objective of feature extraction is to obtain the most discriminating information in the recorded images. The visual cues that are most commonly used are color, shape, texture, motion in video and spatial information. For example, color may represent the color information in an image, such as color binary sets or color coherent vector color histogram. Then parts of the feature may be obtained from one of cue or from the number of cues i.e. the merger of color and texture [2]. The hand gesture images are successfully examined from three stages i.e. preprocessing, **feature extraction** and classification.

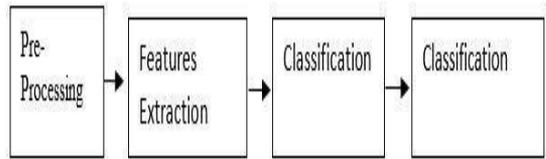


Figure7. Gesture Recognition Steps

a. Pre-processing:

In the **pre-processing** stage, some operations are applied to extract the hand gestures from to background and formulate the hand gestures for feature extraction.

The process starts with filtering noise, followed by an image adjustment segmentation, thresholding, edge detection, histogram equalization and normalization which have proved to be the good structure of image processing [4].

b. Segmentation

Segmentation is the first stage of recognition process in which acquired images are broken down into expressive segments. The main working of the segmentation process is a division of the image without considering region represents. If the binary images are chosen, only two regions exist, foreground region and background region [2]. For example, when a natural scene is segmented, a region of cloud, mountain, flowers, and the tree may exist [3]. Segmentation should be stopped when the object of interest in an application is isolated [5].

c. Thresholding:

There are many methods of feature extraction. Feature extraction step is usually followed by the classification method that uses the extracted feature vectors to classify the gesture image into the respective classes [2]. Only feature invariants are given as an example of feature extraction used with the dataset. A simple image segmentation problem occurs when an image contains an object

that has homogenous intensity and background with different intensity levels [2]. While using feature extraction, hands should be perfectly segmented with a simple color filter. When the environment changed, the image character will change along with the environment. It is harder to detect the image. Especially, when posture or shooting angle change, it will affect the image detection eventually. Therefore, there is need to search for a particular character, that is fixed structure feature posture. Especially when shooting angle varies, and use those features to detect the human hand.

To implement this method the researchers started to examine the image feature that does not change the environment. If the researcher found that skin color is a feature that does not change the environment after the experiment, after that it can be used. Different lighting conditions with skin color may affect the accuracy of hand detection. However, if the chosen image size is big, we can locate the smaller color space inside the skin color and will be able to make image detection simple and easy.

d. Gesture Classification

This process is used to recognize the hand gestures. The recognition process is affected by proper selection of features parameters and suitable algorithms. For example, edge detection or contour

operator cannot be used for hand gesture recognition; many hand gestures were generated before and could produce misclassification. Statistical tools are used for gesture classification and HHM has shown its ability regarding dynamic gestures.

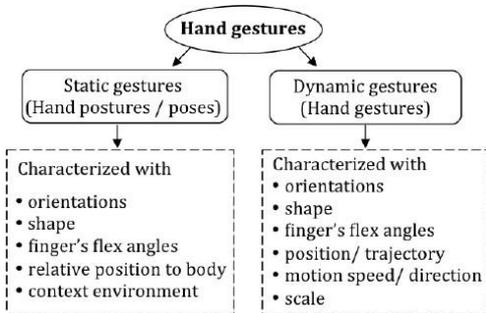


Figure 8: Gestures Classification 1

5. DEFINED CLASSES AND DISCRIMINATION BETWEEN ASL GESTURES:

In a study, there are 36 classes of datasets. However, depending on the extraction used, there are some similar gestures and postures difficult to classify. For example, the dissimilarity between “M” and “N” is the thumb appearing not between the fingers. Many simple feature extraction methods will not easily identify the difference between them. Other examples are “V”, “S” against “T” and “K” in the opposition of “2” and “I” against “J” etc. It is observed that some gestures are

identical “O” (letter O) and “0” (digit 0). The decision to coalesce classes does not have least impact on feature extraction and image acquisition method. For example, for a rotation invariant method “I” and “J” are too similar. However, for non-invariant features, they might belong to distinguish classes. Two ASL gestures “J” and “Z” would be dynamic in their original state. In order to standardized datasets, only the movements are not considered. Instead, we slightly rotated the gesture to create dissimilarity from one and other gestures. One can find many changes of ASL gestures alphabetic digits and letters [1]. Similar hand gestures datasets Table

Table 4: Similar hand gestures

| Similarity Gestures Observations | | | |
|--|--------------|--------------------------------|--|
| Identical | “0” and “O” | “V” and “2” “V” “2” and “K” | The difference between K in the position |
| Of thumb | Very Similar | “W” and “6” “Z” and | “1” |
| Different angles | | “M” and “N” “I” and | “J” |
| Different angles (J is dynamic in ASL) | | “D” and | “1” |
| Position of Thumb | | “S” and | “T” |

Position of Thumb

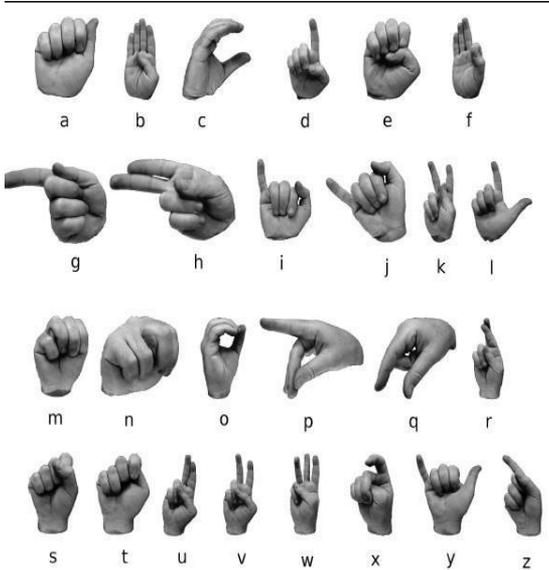


Fig 9: The ASL alphabetic hand gesture dataset 1

CONCLUSION

This paper concludes is about datasets based on American Sign Language in the form of static gestures. Multiple camera input pictures are used to create an effective dataset. These datasets contain only the 26 alphabetic hand gestures. These datasets are based on a single Sign language, it is needed to focus on many other languages such as Urdu Sign Languages for Urdu Speaking country, Arabic Sign Language for Arabic people, French Sign Languages and so on. We have only mapped the American Sign Language based on 26 English alphabetic characters and 0-9 digit based gestures to represents the similarity between these gestures.

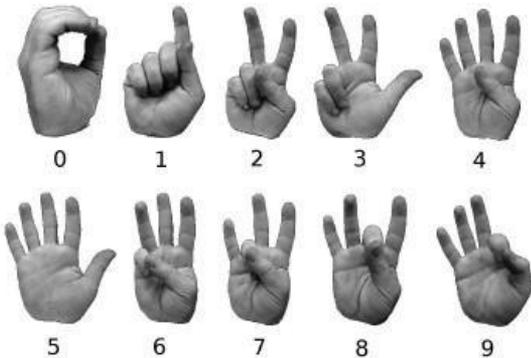


Fig 10: The complete ASL numeric datasets 1

REFERENCES:

- [1]A.L. BARCZAK, N. R. M. A. P. A. T. S., 2011. *Res. Lett. Inf. Math. Sci*, Volume 15, pp. 12-20.
- [2]Haitham Badi, S. A. K. S. H., 2015. *Science Gate Publishing- Available underCC BY-NC 4.0*, Volume 3.
- [3]Parvathy.D, P., 2016. Hand Gesture Recognition Using PreProcessing, Background Subtraction And Segmentation Techniques. *International Journal Of Applied Engineering Research ISSN*, Volume 11, pp. pp 3221-3228.
- [4]T.Awcock, G. a., 1995. Applied Image Processing. *Mc Graw Hill*.
- [5]R.E.Woods, R. a., 2002. Digital Image Processing. *Prentice Hall*.
- [6]Franco Rochentti, F. Q. E. L. L. a. A. R., n.d.LSA64: An Arigentinian Sign Language Dataset.
- [7]Farhad Dadgostar, A. A. S., n.d. A color hand gesture database for Evaluation and Improving Algorithm on Hand Gestures and Posture Recognition. *Massey University at Abaney Auckland, Newzeland*.
- [8]Sanjay Meena, D. S. A., 2011. Study of hand gestures recognition techniques. *Department of Electronics and Communication Engineering National*
- [9]Srinivas k, M. K. R., 2017. Study Of Hand Gestures And Recognition And Classification. *Asian journal of pharmcetutical and clinical Research*, v10s1.19540(April).
- [10]Nachamai.M, n.d. Alphabet Recognition Of American Sign Language:A Hand Gesture Recognition Approach Using Sift Algorithm. *International Journal Of Artificial Intelligence And Application (IJAIA)*, Vol 4,No 1(1 January 2013).
- [11]Simon Ruffieux, D. L. E. M. a. O. A. K., n.d. A Survey Of Datasets For Human Gesture.
- [12]Mark Dilsizian, Z. T. D. M. M. H. a. C. N., n.d. The Importance Of 3D Trajectories Of Computer Based Sign Languages. *7th Workshop on the Representation and Processing the Sign Language: Corpus Mining Language Resources and Evaluation Conference (LREC) Portoroz, Salovenia*, Issue 28 May 2016.
- [13]S.Nagarajan. T.S.Subrashini, M., n.d. Visual Interpretation of ASL Finger Spelling using Hough Transform and Support Vector Machine. *International Journal of Advanced Research in Computer and Communication and Engineering*, Vol.4(6 June 2015).

[14] Alhussain Akum, N. A., 2015. Hand Gesture Recognition Approach for ASL language using Hand Extraction Algorithm. *Journal of Software Engineering and Technology* , Issue 28 August 2015, pp. 419-430.