A Survey on Sign Language Recognition

Sumaira Kausar
sum_satti@yahoo.com
Department of Computer Engineering,
College of Electrical and Mechanical Engineering,
National University of Sciences and Technology (NUST),
PAKISTAN

M. Younus Javed
myjaved@ceme.nust.edu.pk
Department of Computer Engineering,
College of Electrical and Mechanical Engineering,
National University of Sciences and Technology (NUST),
PAKISTAN

Abstract--- Sign Language (SL) recognition is getting more and more attention of the researchers due to its widespread applicability in many fields. This paper is based on the survey of the current research trends in the field of SL recognition to highlight the current status of different research aspects of the area. Paper also critically analyzed the current research to identify the problem areas and challenges faced by the researchers. This identification is aimed at providing guideline for the future advances in the field.

Keywords---- Sign Language, dynamic signs, static signs

I. INTRODUCTION

In daily life, human beings communicate with each other and interact with computers using gestures. As a kind of gesture, Sign Language (SL) is the primary communication media for deaf people. In taxonomies of communicative gestures, Sign Language (SL) is rightly graded as the most structured form of gesture. As non-SL gestures often consist of small limited vocabularies, they are not a useful benchmark to evaluate gesture recognition systems. However, SL on the other hand can offer a good benchmark to evaluate different gesture recognition systems because it consists of large and well-defined vocabularies.

In this paper latest researches conducted in the field of Sign Language recognition are reviewed and critically analyzed. This paper identifies the shortcomings in the research done in the area of sign language recognition. This identification of shortcomings is aimed at providing a guideline for the future possible work which can be done and what researchers should focus on while performing research in this particular area.

II. SIGN LANGUAGE RECOGNITION- A REVIEW

Sign language recognition is getting more and more attention of the researchers due to its applicability in many areas such as deaf people communication system, virtual reality, human computer interaction, machine control in the industrial field and many more.

While looking at the area, research can be divided into two broad categories on the basis of the signs recognized in these researches. These categories can be named as:

- Static Signs based recognition
- Dynamic/ Signs based recognition

In the following two subsections latest researches conducted in the field of sign language recognition are summarized. As paper is aimed at providing state of the art research trends and future guidelines, so only researches conducted in the recent most two years are quoted.

The most important considerations for a sign language recognition system are dictionary of signs, segmentation, feature selection and classification technique. Following summary is focused to cover latest trends of all these four aspects of the sign language recognition systems.

A. Static sign based recognition

First the researches conducted for the static signs are analyzed. Majority of the research till date is conducted for static signs. For analyzing the current research trends in the field, only the latest researches are quoted here.

Omer Rashid & et. al. [1] used Support Vector Machine to recognize static one hand finger spellings of American Sign Language (ASL) and taken geometrical descriptions of fingers as features along with Hu moment and Zernike moment for recognition. Only static alphabets and numbers in ASL are recognized. Features selected are invariant in terms of transformation because of using geometrical moments. Accuracy rate is 96.2% to 98.6% respectively for Hu moment and Zernike moment. Ali Karami and et. al [7] have recognized thirty two static alphabets of Persian Sign Language. Wavelet transforms is used for feature extraction and multi layered perceptron Neural Network for classification. Instead of applying segmentation technique, manually cropped image is used. Recognition is performed on bare hands with an accuracy rate of 94.06%. Ravikiran and et. al [8] have used boundary tracing and finger tip detection to recognize one handed static finger spellings of American Sign Language. No input gloves or colored markers are used. Subset of ASL alphabets is recognized. Classification is based on open fingers in a gesture. Technique does not require perfect alignment of hand to the camera. Kelly and et. al [9] have recognized static one handed postures using Support Vector Machine using a combination of eigen space size function and Hu moments features. Technique is robust and independent of particular signers. Two different data sets are used for this purpose. Chung huang and et. al [14] have used Support Vector Machine to recognize static signs in Taiwanese Sign Language. Aleem Khalid and et. Al [4] of have used data gloves to recognize one handed static urdu sign alphabets. They have used statistical Template Matching for classification purposes. Fuzzy Classifier used for the recognition of alphabets of PSL by Sumaira kausar and et. al [18] . Colored gloves were used for the recognition of PSL symbols.

B. Dynamic Signs based recognition

Dynamic/continuous sign based sign language recognition has a lot of potential in it for research. Some researchers have performed some research in the area but a lot of aspects yet to get more focused work. Some of the latest work done in this area is summarized here to highlight the techniques used for this purpose and problem areas and future guidelines.

Hong Li and et. al [2] have been able to recognize dynamic single hand and two arms signs using contour based features. The main focus of the research is end point localization. Dynamic programming is used for this purpose. Attained recognition rate is about 88% to 96%.Dynamic programming is used by Yang and et. al.[5] to recognize dynamic gestures of American Sign Language. Experiments are performed in a controlled environment. Special dressing scheme and background requirements are used for accurate recognition of signs. Results are compared with CRF and latent dynamic CRF and an improvement attained is 40% in terms of frame labeling rate. Dynamic Indian Sign Language of two handed signs is recognized by Anup Nandy and et. al [6]. They have used direction histograms, Euclidean distance and k-nearest neighbor metric for classification. The technique is illumination and orientation invariant. Eight two handed gestures are detected and recognized from videos of continuous signs, by Kelly and et. al [10] using movement epenthesis and Hidden Markov Model. The detection accuracy rate attained is 95.6 % and recognition accuracy rate is 93.2%. Combination of Confidence Measure (CM) obtained from aligning forward sign models (the conventional approach) to the input data with the CM's obtained from aligning reversed sign models to the same input, is used by Zafrullah and et. al [11] for the verification of continuous phrases of American Sign Language. Test results shown that this combination of CMs yields improvement in verification accuracy as compared to the traditional approach. Quantification of facial expressions along with tracking of hand shapes, based on skin color, is performed in a research conducted by Justus and et. al [12] for continuous sign language recognition. Zahoor Zafrullah and et. al [13] performed real-time American Sign Language (ASL) phrase verification using Hidden Markov Model for an educational game, CopyCat, which is designed to improve deaf children's signing skills. They achieved a phrase verification accuracy of 83%. Huang and et. al [14] have used Hidden Markov Model for recognizing Taiwanese dynamic signs, and Finite State Machine for verification of these signs. Kelly and et. al [16] implemented Conditional Random Field (CRF), Hidden CRF and Latent-Dynamic CRF based systems and compared these to a HMM based system when recognizing motion gestures and identifying inter gesture transitions. Nguyen and et. al [17]proposed to use a 2-layer Conditional Random Field model for recognizing continuously signed grammatical markers in ASL. Grammatical markers have facial feature movements and head movements. Accuracy rate of the proposed scheme is about 93.76%.

III. PROBLEM AREAS OF SL RECOGNITION

Critical analysis of the existing research has shown that there are many shortcomings and inherent challenges of the area of SL recognition. The shortcomings and challenges are generic in the sense that paper is not focused on the problems associated with a particular classification technique, because a lot of literature already available on it. Rather focus is on the general problems faced while developing an overall SL recognition system. These problem areas are:

- Segmentation
- Unrestricted environment
- Size of dictionary
- Invariance
- Variety of Gestures
- generality
- Start/end identification of gesture sequences
- Feature selection
- Feature extraction

A. Segmentation

Segmentation of hands for gesture recognition has many problems. Segmentation is achieved by two broad techniques. First one is skin color based and the second one is using some external aid.

Skin Color Based

Many researchers have used skin color based segmentation. But it has many inherent problems that negatively affect the accuracy of the segmentation. These problems are:

- Illumination variation
- Background colors
- Area of Interest segregation
- Wide range of skin colors

Variation in illumination is very difficult to accommodate for accurate segmentation. Similarly if there is complex background there is probability for false acceptance of objects near to human skin color other than hands, arms or face. It is difficult in such situations to correctly segregate the area of interest. There may be many other objects selected as objects of interest, which is definitely difficult to fully automate the problem. Another problem with the skin color based segmentation is that there is a large variety of skin colors, so process of selecting appropriate range for human skin color is very difficult.

External Aid

Skin color based techniques for segmentation have many problems. So many researchers have opted to use external aid to get rid of the problems associated with skin color based techniques. External aid used by researchers is of two types:

- Data gloves
- Color gloves

This external aid has made segmentation process easier but it is not appreciated by many. Because it is definitely not admired to first equip oneself to use a particular system and it definitely reduce its general applicability.

B. Unrestricted Environment

Due to problems in segmentation, feature selection and feature extraction many restrictions on the environment of the signers are forced by the researchers to get better accuracy. These restrictions may be are:

- Signer's Background
- Signer's Clothing
- Preprocessed images

Normally researchers put restrictions on background color to make the background simple to avoid problems in segmentation

and feature extraction. Similarly dressing scheme is also restricted by many researchers such as half sleeves, particular color of the cloth etc. sometime researcher prefer to manually crop or change colors of the image of the signer to avoid some of the challenges in classification problems.

C. Size of dictionary

Size of dictionary used in existing researches is usually very small. Majority researches have used only alphabets of some language that are usually static one handed signs. In case of continuous sign eight or ten phrases are used for this purpose. Because research would be of little applicability and usage till a larger dictionary of signs used.

D. Invariance

Another problem that is observed in many researchers' work is that feature selected is not invariant in terms of orientation of signer and illumination variations. So it is required for a good sign recognition system to possess invariance in terms of following:
- Rotation
- Translation
- Scaling
- Illumination

E. Variety of Gestures

Sign languages have gestures which normally use one handed signs, two handed gestures, arms position, facial expressions, eye gaze, head and body postures. It is very difficult to accommodate all these types of gestures, so normally limited set of gestures is used by researchers. Mostly one handed or two handed signs are used. Because as more and more types of gestures are included in the dictionary, it becomes very difficult to select an appropriate feature set that can accommodate all types of gestures.

F. Generality

It is observed in the current researches, feature selection is based on particular sign language and selected subset of the dictionary. So techniques used for recognition lack generality. As majority techniques are difficult to apply for broader dictionary of gestures. So generality can be another feature of the future techniques for SL recognition.

G. Motion gestures

Most of the researches done on sign language recognition are performed for static signs. Few researches performed for dynamic gestures. Researches performed on motion gestures have very limited applicability. Limited gesture types and very small dictionaries are used so far. Motion gestures have following challenges:
- Start/ end identification of gesture sequences
- Motion identification
- Isolated sign recognition

Motion gestures are sequence of signs. So it is important and yet very challenging to identify the individual/isolated sign in the sequence, motion identification of the signer and most importantly, start and end identification of isolated gesture and complete phrase. So it is very important to focus research on movement epenthesis.

H. Feature Selection

Feature selection is a very crucial step in the overall classification problem of sign language. There are many challenges that researchers have to face while designing a sign language recognition system. These challenges are:
- Minimizing Feature set
- Robust feature Selection
- Features based on local information

These are some the leading problems associated with the very critical phase of SL recognition system i.e. feature selection

I. Feature Extraction

When features are selected another important step is to extract these features from the image of the signer. Feature extraction process has to overcome following challenges:
- Complex background
- Colored and data gloves

Complex background definitely makes it difficult to extract required features accurately. For facilitating the process of feature extracting, many researchers have used data or colored gloves. This is definitely not required for the general applicability of the system.

IV. FUTURE GUIDELINES FOR SL RECOGNITION

After summarizing the latest researches in the field of sign language recognition and related problem areas, it is justified to present future guidelines depending upon the work already done. The challenges highlighted above are actually the future guideline for the research in the area of SL recognition. Researchers can improve future research in the field of SL recognition with the help of the following guidelines.
- It is a better approach to not use any external aid for segmentation, rather segmentation should be vision based. For coping up problems with skin color based segmentation, it is recommended to use a combination of skin color based features along with geometrical features. This can lead to the formulation of a better segmentation technique.
- Techniques for classification and segmentation should be adopted which are not putting much restrictions on the environment of the signer. e.g. if segmentation is using both geometrical and skin color based features, then restrictions on background colors and signer's clothes can be minimized.
- Dictionary used for future researches should be larger. For this purpose more focus is required on the feature selection process. Such features should be used which are general and can be applicable for a broader set of signs. So a larger set of dictionary of SL should be accommodated by the future researches.
- Transformation and illumination invariance is a required feature for the future researches in the field of sign language recognition systems. Incorporating classification technique with some invariant methodologies such as geometrical moments can be helpful in this respect.

- Variety of gestures should be accommodated in SL recognition systems. As this variety of gestures is an important aspect that should be catered for in the future work in the field. For this purpose special focus and extensive work is required on feature selection and feature extraction. Multilayered and multidimensional feature set is required to be developed. Special dimensionality reduction techniques then required to apply an efficient classification methodology.

- For an efficient classification solution, it is required to minimize the feature set while accommodating maximum possible set of sign dictionary and maximum types of gestures. It is also desirable yet challenging to select features which are robust to cover a larger set of dictionary and type of gestures. It is also required for a robust set of features that it relies on local information of signer without relying on some external cues or markers. So the feature sets defined in future techniques should accommodate these challenges.

- Feature extraction challenges should be addressed appropriately for the future applications. Such features should be selected that are not disturbed by signers background so these features can be correctly extracted.

V.   CONCLUSION

This paper analyzes latest research work done in the area of SL recognition. Future guidelines are provided while analyzing the shortcomings and challenges faced by the researchers in the area at the moment. The main aspect, which requires a lot of research effort, is dynamic continuous signs. Size of dictionary and variety of gestures should be maximized. This paper is just an effort to provide fresh researchers with a one point access to the current status of the research in the area along with their shortcomings anf future guidelines.

REFERENCES

1]Omer Rashid, Ayoub alHamadi,Bernd Michelis,"Utilizing invariant descriptors for finger spelling American Sign Language using SVM", ISVC 10, Proceedings of 6th international conference on advances of visual computing, Volume part 1.

[2] Hong Li and Michael Greenspan," Model-based segmentation and recognition of dynamic gestures in continuous video streams ", Science direct pattern recognition corrected proof 2010.

[3] Trmal Jan and Marek Hrúz : Evaluation of Feature Space Transforms for Czech Sign-Language Recognition . MetaCentrum Yearbook 2010, p. 145-150, 2010.

[4] Aleem Khalid Alvi, M. Yousuf Bin Azhar, Mehmood Usman, Suleman Mumtaz, Sameer Rafiq, RaziUr Rehman, Israr Ahmed, "Pakistan Sign Language Recognition Using Statistical Template Matching", World Academy of Science, Engineering and Technology 3 2005.

[5] Yang R., Sarkar S., Loeding B.," Handling Movement Epenthesis and Hand Segmentation Ambiguities in Continuous Sign Language Recognition Using Nested Dynamic Programming ",Pattern Analysis and Machine Intelligence, IEEE Transactions on March 2010,volume 32 Issue:3 pp 462 – 477.

[6] Anup Nandy, Jay Shankar Prasad, Soumik Mondal, Pavan Chakraborty and G. C. Nandi," Recognition of Isolated Indian Sign Language Gesture in Real Time   ",Information Processing and Management Communications in Computer and Information Science, 2010, Volume 70, 102-107, DOI: 10.1007/978-3-642-12214-9_18

[7] Ali Karami, ahman Zanj and Azadeh kiani, "Persian sign language (PSL) recognition using wavelet transform and neural networks ",Expert Systems with Applications Volume 38, Issue 3, March 2011, Pages 2661-2667

 8] J. Ravikiran, Kavi Mahesh, Suhas Mahishi, R. Dheeraj, S. Sudheender and Nitin V. Pujari," Automatic Recognition of Sign Language Images ",Intelligent Automation and Computer Engineering Lecture Notes in Electrical Engineering, 2010, Volume 52, 321-332, DOI: 10.1007/978-90-481-3517-2_25 Computer Science, 2011, Volume 6495/2011, 665-676,

 [9] Kelly D.,   McDonald J., Markham C.," A person independent system for recognition of hand postures used in sign language", Pattern Recognition Letters Volume 31, Issue 11, 1 August 2010, Pages 1359-1368

[10] Kelly D., McDonald J., Markham C.," Continuous recognition of motion based gestures in sign language " , Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on Sept. 27 2009-Oct. 4 2009, Kyoto,pp 1073 – 1080.

[11]   Zafrulla Z., Brashear H., Hamilton H.,   Starner T.,"A novel approach to American Sign Language (ASL) phrase verification using reversed signing",   Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference San Francisco, CA,, 13-18 June 2010 , pp 48 – 55.

[12] Justus Piater, Thomas Hoyoux, Wei Du, "Video Analysis for Continuous Sign Language Recognition",4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies, 2010 (Workshop at the 7th International Conference on Language Resources and Evaluation (LREC), Malta)

[13] Zahoor Zafrulla, Helene Brashear, Pei Yin , Peter Presti, Thad Starner,   Harley   Hamilton,   "American   Sign   Language   Phrase Verification in an Educational Game for Deaf Children ",20th International Conference on Pattern Recognition, Istanbul, Turkey ,August 23-August 26 2010.

[14] Chung-Lin Huang; Bo-Lin Tsai; ," A Vision-Based Taiwanese Sign Language Recognition", Pattern Recognition (ICPR), 2010 20th International Conference,Istanbul ,Turkey , 23-26 Aug. 2010, PP 3683 – 3686.

[16] Kelly D., McDonald J., Markham C., "Evaluation of threshold model HMMS and Conditional Random Fields for recognition of spatiotemporal gestures in sign language ",Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on Sept. 27 2009-Oct. 4 2009, Kyoto, pp 490 – 497.

[17] Tan Dat Nguyen and Surendra Ranganath, "Recognizing Continuous Grammatical Marker Facial Gestures in Sign Language Video ",Computer Vision – ACCV 2010 Lecture Notes in Computer Science, 2011, Volume 6495/2011, 665-676

[18] Sumaira Kausar, M. Younus,Javed, Shaleeza,Sohail: "Recognition of gestures in Pakistani sign language using fuzzy classifier" , 8th conference on Signal processing, computational geometry and artificial vision 2008, Rhodes, Greece, Pages 101-105