Towards Computer Assisted International Sign Language Recognition System: A Systematic Survey

Mikhaylyna Melnyk
Department of Information Technology
University of Telecommunications
Ukraine

Vira Shadrova
Department of Information Technology
University of Telecommunications
Ukraine

Borys Karwatsky
Department of Information Technology
University of Telecommunications
Ukraine

ABSTRACT
There is a number of automated sign language recognition systems proposed in the computer vision literature. The biggest drawback of all these systems is that every nation has their own culture oriented sign language. In other words, everyone needs to develop a specific sign language recognition system for their nation. Although the main building blocks of all signs are gestures and facial expressions in all sign languages, the nation specific requirements make it difficult to design a multinational recognition framework. In this paper, we focus on the advancements in computer assisted sign language recognition systems. More specifically, we discuss if the ongoing research may trigger the start of an international sign language design. We categorize and present a summary of the current sign language recognition systems. In addition, we present a list of publicly available databases that can be used for designing sign language recognition systems.

General Terms
Gesture Recognition, Computer Vision, Survey.

Keywords
International sign language, sign language recognition, deaf community, survey of sign language recognition.

1. INTRODUCTION
A sign language is a set of gestures and facial expressions which are used for the communication by the deaf community. It involves simultaneously combining the facial expressions, hand orientations and shapes, and body posture. Each sign language may have thousands of signs. We should emphasize that in this paper (or in the sign language world) the word "sign" refers to a meaningful unit element of a conversation in a sign language. A sign may correspond to a single word or a group of words in a spoken language. One can categorize the signs into two subgroups: hand performed signs and non-manual signs. In terms of grammar and syntax, a sign language may be completely different from its counterpart spoken language.

The research on computer assisted sign language recognition systems intend to develop methods and algorithms to recognize the meanings of the signs performed during a conversation. Sign languages pose the challenge that there is unfortunately no international sign language. Almost all of the sign languages can be considered as local spoken languages. Furthermore, there is a lack of written form of sign language rules and gestures.

The papers discussing computer assisted recognition of sign language have been presented in the literature since nineteen eighties. In this paper, we promote and highlight the idea that a future existence of a robust automated system that can efficiently recognize signs may lead to the standardization of sign languages and ultimately creation of an international sign language.

In the next sections, we present the key aspects of sign language recognition and categorize the current methods in terms of the data acquisition device being used. This is followed by a discussion on how a computer assisted international sign language recognition system can be developed.

2. HYBRID SYSTEMS
Early sign language recognition systems used hardware such as data gloves and accelerometers to acquire specific features of the signs. There are several reasons behind this trend. The processing power required for real time vision based methods were very high. It was (and some is still) very difficult to extract some features such as hand orientation from visual data. Furthermore, hardware such as a glove was thought of as being widely used without the need of environmental calibration.

In his patent, Grimes [1] proposed the idea that a glove equipped with sensors and associated electronic logic can be designed for manual sign recognition in American Sign Language. He described the overall framework as follows: a deaf or hearing impaired user wearing the data entry glove to input data performs a manual sign and the receiving device converts the signal into a single sign in the finder spelling alphabet. The early study of Hall [2] described an experimental system on finger postures to allow input by finger spelling using the LED glove. The proposed system was the pioneer for using hand gestures for sign language recognition. Kadous [3] employed instance-based learning and decision tree methods on Australian Sign Language data which were collected by by a PowerGlove. The user dependent system could recognize 95 signs. The accuracy of the proposed method was about 80%. Vogler and Metaxas [4] used DataGlove to obtain three dimensional coordinates, orientation, and velocity of the hands and arms. These parameters are used as features for Hidden Markov Models to recognize 486 American Sign Language sentences. Starner and Pentland [5] described a real time Hidden Markov Model based system for recognizing sentence level American Sign Language, where the signer wore solidly colored gloves. Brashear et al. [6] proposed American Sign Language
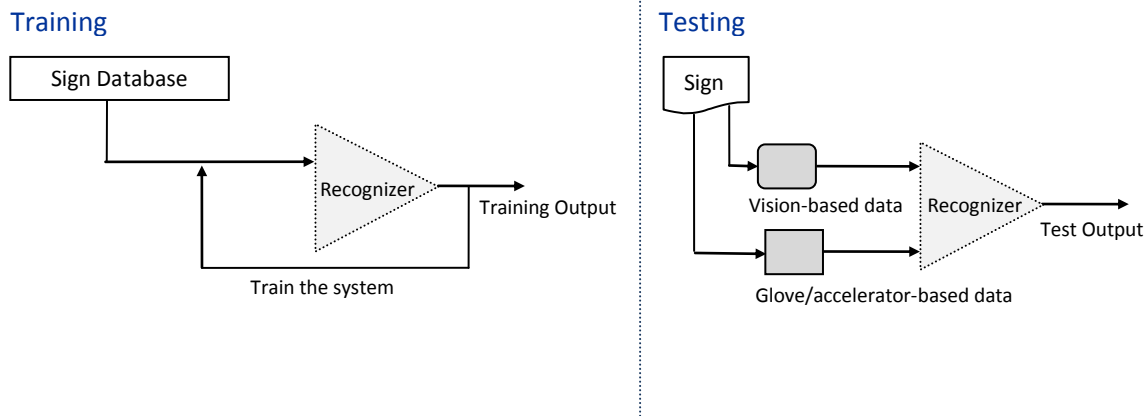
Training

Testing

Sign Database

Sign

Recognizer

Training Output

Train the system

Vision-based data

Recognizer

Test Output

Glove/accelerator-based data

Fig 1: A two-step framework for the development of a sign language recognition system

recognition system with the ultimate goal of designing a mobile assistive technology. They used multiple sensors to remove the noise and improve recognition accuracy. The proposed system consists of hat mounted camera, wearable computer, and accelerometers. They marked the left hand by a cyan band on the wrist and the right hand by a yellow band. Liang and Ouhyoung [7] proposed real-time continuous gesture recognition of Taiwanese Sign Language using DataGlove. This system used Hidden Markov Models for 51 fundamental postures, six orientations, and 8 motion primitives and achieved the average recognition rate of 80.4%. Gao et al. [8] described a system called HandTalker for the recognition of Chinese Sign Language. They embedded interfaces to translate the signs into text or vocal outputs. Hernandez-Rebollar et al. [9] used AcceleGlove for recognizing the 26 hand shapes of the American Sign Language alphabet. AcceleGlove was different from previous approaches in such a way that it did not require a tracking system to identify hand orientation. This allowed the system to identify postures that were not possible before. Mehdi and Khan [10] investigated the possibility of recognizing sign language gestures using sensor glove, which is made out of cloth with sensors fitted on it. The glove has 7 sensors, where 5 sensors are for each finger, one sensor is for measuring the tilt of the hand, and one sensor is for measuring the rotation of the hand. Their recognition algorithm is based on an Artificial Neural Network and yields a recognition accuracy of 88%. Akyol and Canzler [11] used color coded gloves for finding and tracking the hands reliably. They obtained a user independent recognition rate of 94% on a set of 16 signs of German Sign Language. Kuroda et al. presented [12] an intelligent data-glove called StrinGlove. StrinGlove can determine full degrees of freedom of human hand using 24 inductcoders and 9 contact sensors. Then, it encodes hand postures into posture codes on its hardware. This in-device encoding function decreases the load on the central CPU; thereby, StrinGlove eases to develop sign language recognition system Hernandez-Rebollar et al. [13] discussed an approach for capturing and translating isolated signs of American Sign Language using the system composed of an AcceleGlove and a two-link arm skeleton. Their classification method is based on Hidden Markov Models and Artificial Neural Networks. The proposed sign recognizer was tested using a 176-sign corpus achieving an accuracy of 95%.

In more recent studies, devices with more advanced technologies and combination of different sensors have been used. Oz and Leu [14] used a sensory glove called

CyberGlove and a 3-D motion tracker to extract the gesture features. Finger joint angles, which define the hand shape, obtained using the glove. The data from the tracker was used to estimate the trajectory of hand movements. Zafrulla et al. [15] presented an approach using Kinect depth-mapping camera for American Sign Language recognition. They compare the Kinect-based system to their CopyCat system, which is composed of colored gloves with embedded accelerometers. Assaleh et al. [16] proposed a low complexity classification method to recognize Arabic Sign Language using sensor-based gloves. The gloves have 5 bend sensors and a 3D accelerometer. Their system yielded recognition rates of 92.5% and 95.1% for user dependent and user independent cases, respectively. Jeong et al. [17] developed a finger-gesture recognition glove recognizing Korean Sign Language with a use of inexpensive and conductive materials. Velostat, proposed finger gesture recognition glove, is made of a film surfaced with carbon particles, used for anti-static package of electronic parts. They could measure bending angles of the fingers using voltage divided circuit and calculate the hand motion angle. Luzanin and Plancak [18] used a low-budget 5-sensor data glove for hand gesture recognition. Their probabilistic neural network-based gesture recognition method was able to process user independent simple and complex static gestures.

Due to inconvenience to the user and the high cost of the hardware-based systems, the trend is in favor of image and video based sign language recognition systems. In the next section, we will present approaches that use only vision based properties. In Table 1, we compare the different aspects of designing hardware and vision based computer assisted sign language recognition systems. Waldron and Kim [19] proposed a system based on a two-stage neural network that can recognize isolated American Sign Language. They used the hand shape and position data obtained from a DataGlove which is mounted with a Polhemus sensor. The first level of the method consists of four back-propagation neural networks recognizing the sign language phonology. Specifically, the data contains 36 hand shapes, 10 locations, 11 orientations, and 11 hand movements. The signs are fed to the second stage which recognizes the actual signs. They used both back-propagation and Kohonen's self-organizing neural work to compare the performance and the expandability of the learned vocabulary. When they use the back-propagation network for the second stage, they obtained that the network was able to recognize these signs with an overall accuracy of 86%. When they use a Kohonen's self-organizing network in the second

stage, the network could not only recognize the signs with 84% accuracy. However, they expand its learned vocabulary through relabeling.

Table 1. Comparison of the vision and hardware based methods based on the design concepts

Concept	Vision-based	Hardware-based
Calibration	*Depends environment*	*Required but stable*
User dependency	*Highly prone*	*Less prone*
Cost	*Low*	*High*
User experience	*Good*	*Inconvenient*
Feature extraction	*Challenging*	*Relatively easier*
World wide availability	*High likely*	*Low likely*

3. VISION-BASED SYSYEMS

Vision-based systems use appearance-based image features extracted from video sequence frames.

3.1 Isolated Sign Recognition

Most of the approaches focus on isolated sign recognition. A subset of those uses the finger spelling, where a word is signed letter by letter usually using hand shapes.

Fagiani et al. [20] proposed a system for automatic recognition of Italian Sign Language based on the position and the shape of head and hands. Their system is composed of a feature extraction stage and a sign recognition stage. Sequential forward selection method was used to obtain feature vectors with the minimum dimension and the best recognition performance. They represent each sign with a single Hidden Markov Model having parameters estimated through the re-substitution method. Akyol and Alvarado [21] proposed a recognition system for the mobile use. They extract and combine motion cues and skin color features from single-view image sequences. The proposed method employs Bayes' classification to generate a probability map based on the skin color. In addition, they used motion history to generate a similar motion probability map. By combining these two probability maps, they detect the signer's hands. Kishore and Kumar [22] presented a method to segment gestures of Indian Sign Language. Their algorithm extracts signs from video sequences under various dynamic backgrounds. The signs are segmented by minimizing the energy function of the level set fused by various image characteristics such as boundary and shape information, texture, and color. From RGB color video three color planes are extracted and one color plane is used based on the contrasting environments presented by the video background. Texture edge map provides spatial information which makes the color features more distinctive for video segmentation. Their experiments show that the method provides segmentation for user independent signer case under different imaging conditions. Haberdar and Albayrak [23] developed a series of recognition systems using initially global and then local features extracted from videos showing signs in Turkish Sign Language. Their method is a two-step algorithm, where they first classify the signs based on the movement of the hands. Then, shape of the hands is used to recognize the sign in the initially selected subset. Their system could recognize 172 isolated signs. Zhang et al. [24] presented a vision-based medium vocabulary Chinese Sign Language recognition. The

proposed recognition system is composed of two sub-modules. They combine techniques of robust hands detection, background subtraction and pupil detection to extract the feature information in unconstrained environment in the first stage. Their method employs principal component analysis to characterize the finger features more elaborately. In the second part, a Tied-Mixture Density Hidden Markov Models framework is used for sign language recognition. The goal of the second part is to speed up the recognition without significant loss of recognition accuracy. They achieved recognition accuracy is up to 92.5% based on 439 frequently used Chinese sign language words. Zieren and Kraiss et al. [25] presented a system that uses a background model to remove static areas from the input video on pixel level. They used Hidden Markov Models along with features normalized for person-independence and robustness. Their system performs person-dependent recognition of 232 isolated signs with an accuracy of 99.3% in a controlled environment. In the second test, person-independent recognition rates reach 44.1% for 221 signs. Overall performance of 87.8% is achieved for six signers in various uncontrolled outdoor and indoor environments but with a reduced vocabulary of 18 signs. Assan and Grobel [26] focused on the video-based signer dependent recognition of isolated signs using the manual parameters of sign language. They consider a sign as a doubly stochastic process to define Hidden Markov Models. The observations extracted from video frames are regarded as feature vectors. The system achieves recognition rates up to 94% on a 262-sign corpus. Zafrulla et al. [27] proposed a relatively novel approach for American Sign Language phrase verification by combining confidence measures obtained from aligned forward sign models and reversed sign models to the same input. They use the Normalized likelihood score and the Log-Likelihood Ratio in the proposed model. Interestingly, they perform leave-one-signer-out cross validation on a dataset of 420 American Sign Language phrases obtained from five deaf children playing an educational game called CopyCat. Their method has a significantly better match to the ground truth when compared to the traditional approach. Feris et al. [28] proposed a novel method for recognition of isolated fingerspelling gestures based on depth edge features. Their approach is based on a simple and inexpensive modification of the capture setup. They use a multi-flash camera with flashes strategically positioned to cast shadows along depth discontinuities in the scene, allowing efficient and accurate extraction of depth edges. The proposed method employs a shift and scale invariant shape descriptor for fingerspelling recognition. They observe an improvement over methods relying on features acquired by traditional edge detection and segmentation algorithms.

3.2 Continuous Sign Recognition

Vision based continuous sign recognition systems usually have an underlying framework that employ Hidden Markov models. Majority of the proposed methods in the literature are build up on a structure based on this foundation, while there are some exceptions.

Hienz et al. [29] describes the development of a signer dependent video-based continuous sign language recognition system using Hidden Markov Models. Their system is developed using a lexicon of 52 signs of German Sign Language and employs manual sign parameters. They used a single color video camera to capture images. The proposed system achieves a recognition accuracy of 95% using a bigram language model. Dreuw et al. [30] developed a system that is able to recognize sentences of continuous sign

language independent of the speaker. They employ pronunciation and language models in sign language. The recognition algorithm is based on the Bayes' decision rule. The system was tested on a publicly available benchmark database consisting of 201 sentences and 3 signers, and they could achieve a 17% word error rate. Holden et al. [31] presented an automatic Australian Sign Language recognition system that tracks the face and hands in videos. They perform the tracking using the corresponding features. The proposed systems could deal with the occlusion of the face and a hand by detecting the contour of the foreground moving object using a combination of motion cues and the snake algorithm. They use Hidden Markov Models as the base model and conduct experiments using 163 test sign phrases with varying grammatical formations. Using prior information about the grammar, the system achieved 97% recognition rate on a sentence level and 99% success rate at a word level. Dreuw et al. [32] investigated whether explicitly incorporating depth information could improve the recognition of American Sign Language. The depth information was extracted using a stereo-vision method on sequential frames. They demonstrated that the depth information helped disambiguate overlapping hands and improve the tracking of the hands; however, the tracking improvements did not have much influence on the final recognition results. Wang et al. [33] presented an approach to large vocabulary, continuous Chinese Sign Language recognition using phonemes instead of whole signs as the basic units. There are about 2400 phonemes in Chinese Sign Language. Because the number of phonemes is limited, Hidden Markov Model based training and recognition of the Chinese Sign Language signal has the potential of recognizing enlarged vocabularies. The proposed method can manage the finger-alphabet and gestures. They generate one Hidden Markov Model for each phoneme, and then the signs are encoded based on these phonemes. They used clustering of the Gaussians on the states, the language model, and N-best-pass to improve the performance of the system. They carried out experiments on a size of 5119 signs. Nayak et al. [34] proposed a continuous state space model, where the states are based on image-based features, without the use of special gloves. They present an unsupervised approach to both extract and learn models for continuous basic units of signs from continuous sentences. Co-articulation effects are very strong in sign language. Given a set of sentences with a common sign, they can automatically learn the model for part of the sign which is affected by co-articulation effects. Their model is based on traces in a space of relational distributions, and each point in this space represents a relational distribution, capturing the spatial relationships between low-level features, such as edge points. Bauer and Hienz [35] described the development of a video-based continuous German Sign Language recognition system based on Hidden Markov Models with one model for each sign. The goal of the proposed system is to design an automatic signer dependent recognition of sign language sentences, based on a lexicon of 97 signs. Their system achieves an accuracy of 91.7%. Kelly et al. [36] presented a multi-modal system for the recognition of manual signs and non-manual signals in continuous sign language sentences based on multichannel Hidden Markov Models. Manual signs and non-manual signals are handled independently using continuous multidimensional Hidden Markov Model threshold model. They use a single non-manual signal, head movement, to evaluate the proposed framework when recognizing non-manual signals. The system achieved a detection ratio of 95% and a reliability measure of 93%. Fang et al. [37] proposed a method for handle transition parts between two adjacent signs in large-vocabulary continuous sign language recognition. They develop a temporal clustering algorithm improved from k-means by using dynamic time warping as its distance measure for tackling mass transition movements arisen from a large vocabulary size. Then, they present an iterative segmentation algorithm for automatically segmenting transition parts from continuous sentences. They demonstrated experiments on continuous sign language recognition and achieved an average accuracy of 91.9% using a vocabulary of 5113 Chinese signs.

4. SIGN LANGUAGE DATABASES

The first challenge to be solved before even starting the design of a sign language recognition system is to obtain a comprehensive database. Without enough training or test data, it is not possible to develop a computer assisted automated recognition system. In this section, we report some of the notable databases from all around the world in alphabetic order. We try to cover as many sign language databases as possible from all around the world. These databases are specifically created for the development of computer assisted sign language recognition systems.

The Purdue RVL-SLLL Database [38] is an extensive database of *A*merican Sign Language gestures, movements, words and sentences. It was produced by 14 signers. This database consists of 2576 videos of 39 motion primitives, 62 hand shapes, and sentences.

The RWTH-BOSTON-400 Database [39] is collected for the development of isolated *A*merican Sign Language recognition systems. It consists of 843 sentences performed by 4 signers.

The *A*uslan Signbank [40] consists of the movies in the Auslan Archive together with linked linguistic annotation files. There are 7415 words in Australian Sign Language and the corpus consists of over one thousand separate video clips. The signs were performed by 100 signers.

Arabic Sign Language Database [16] consists of 80 signs and 40 sentences which were compiled from the signs. Each sentence was repeated 19 times. This corpus is the first fully labeled and segmented dataset for continuous Arabic Sign Language.

The *B*ritish Sign Language Corpus [41] is a collection of video clips showing conversations of 249 participants. The corpus includes annotations of 6330 signs from the signers in the conversational dataset.

LIBRAS-HC-RGBDS Database [42] is a corpus containing 61 hand configurations of the *B*razilian Sign Language. The data were acquired using the Kinect sensor. There are 610 videos of 5 signers in the database.

*D*icta-Sign Corpus [43] is European Union funded research project aiming to make online communications more accessible to deaf sign language users. The ultimate goal of Dicta-Sign is to develop the necessary technologies that make Web 2.0 interactions in sign language compatible. For example, a user can sign to a webcam using signs, or a computer recognizes the signed phrases, translates them into an internal representation of sign language, and then simulates it using an avatar. Dicta-Sign collects the world's first large parallel corpus of domain-specific utterances across different sign languages. Currently, British, French, German, and Greek sign languages are included. The goal is to cover a minimum of three hours of signing in each language, and a vocabulary size of more than 1500 signs. This database is going to be annotated and be publicly available. Dicta-Sign is

a great project aiming to integrate recognition, synthesis and linguistic modeling in one standalone system.

The Corpus *D*utch Sign Language [44] is a collection of video recordings of conversations in the Sign Language of the Netherlands. The corpus consists of recordings with multiple synchronized video cameras, accompanied by gloss and translation annotations. Data are collected from 100 native signers.

The SIGNUM Database [45] contains 450 basic signs in *G*erman Sign Language. Based on this vocabulary, 780 sentences were constructed by 25 signers.

*G*reek Sign Language Corpus [46] is a collection of sentence level sign samples and respective annotations. Video recordings of the database have been produced by 4 signers. The first part of the database is composed of a list of lemmata which are representative of the use of hand shapes as a primary sign formation component. The second part of it includes sets of controlled utterances.

Indian Sign Language Database [47] is prepared by nine signers, where each signer repeated the sign twice under different conditions with a total number of 1440 gesture videos for a total of 80 signs.

The ATIS Sign Language Corpus [48] is available for the following languages: *I*rish Sign Language, German Sign Language and South Sign Language. The corpus contains English phrases and sentences (595 signs) about booking flights and travel information.

The A3LIS-147 Database [49] contains 147 distinct signs from *I*talian Sign Language. 10 signers worked in the project. The signs are organized in six categories, based on different daily life scenarios.

*K*orean Sign Language Database [50] is a very small corpus containing twenty five fundamental gestures. Although there are about 6000 vocabulary words in Korean Sign Language, they are formed by combining a relatively small number of basic signs.

*J*apanese Sign Language Database [51] is a corpus of the face and hand posture images consisting of 516 frame sequences of 258 Japanese subjects. The hand postures represent 45 manual signs.

The MSL Database [52] has isolated and continuous signs in *M*alaysian Sign Language, where each sign is repeated 20 times. The weakest point of this corpus is that only one signer is employed.

*P*akistani Sign Language Database [53] is a relatively small corpus. It consists of 37 signs imported from Urdu, which is the spoken language. Signs in the corpus correspond to finger spelling alphabet.

*P*ersian Sign Language Database [54] is composed of 640 images of 32 signs that correspond to the letters in the Persian alphabet. Images are captured in controlled environment, where the background is solid black.

Spanish Sign Language Corpus [55] includes 4080 Spanish sentences translated into the sign language. The included sentences are from the situations about the renewal of the identity document and driver's license. The database also contains the signs for all the letters, numbers from 0 to 100, numbers for time specification, months and week days.

The BUHMAP Database [56] contains 132 videos of 8 non-manual signs, which are frequently used in *T*urkish Sign Language. The signs are performed by 11 signers. The database also includes ground truth data of 60 manually landmarked points of the face.

*T*aiwanese Sign Language Corpus [57] consists of about 3000 lexical items. The database is a collection of sentences, conversations, and narratives.

*V*ietnamese Sign Language Database [58] is relatively small corpus. It covers videos of 40 signs, where each one was recorded 10 times by two signers under the control of a sign language teacher.

5. HIDDEN MARKOV MODELS

Hidden Markov Models (HMMs) have been always in the center of the attention in the speech processing research [59, 60] for more than 30 years. It also plays a crucial role in sign language recognition research. The main idea is that an HMM is a finite model that represent a probability distribution function for an infinite number of possible input sequences. In continuous speech recognition, units, such as syllabus or phones, are characterized by HMMs whose parameters are different for each unit [61].

HMMs provide a conceptual toolset for constructing complicated models such as sign language recognition by simply drawing an intuitive picture. A sign is a combination of small manual or facial shapes and movements. During the literature review, we observe that many sign language researchers implement initial version of their recognition systems directly using HMMs [62-72] as either main module or a sub-module in the framework. When recognizing a sign, the framework visits a state of HMM by emitting a residue from the probability distribution function of the state [N-1]. Then, the framework selects the next state based on the state's transition probability distribution. For most of the signs, there may be many possible state sequences which is not possible to linearly explore. Therefore, a very efficient algorithm, called Viterbi algorithm, is used to optimize finding the most probable state path given a feature vector and an HMM for the sign [N-1]. Many researchers use HTK Toolkit to implement the Viterbi algorithm [N].

HTK is a toolkit developed by the Cambridge University for developing methods using Hidden Markov Models. HMMs is frequently used in sign language recognition applications to model the time varying features of a sign although HTK is initially and primarily designed for building HMM-based speech processing recognition systems. First of all, a distinct HMM is assigned for each sign. Then, HTK training tools are used to estimate the parameters of all HMMs corresponding to the signs in the database using training samples and their associated transcriptions. Finally, the trained HMMs are used along with the unknown samples and the optimal probabilities are estimated.

6. CONCLUDING REMARKS

In this paper, we present the current status of the research on computer assisted sign language recognition systems. There are many challenges need to be solved. First of all, publicly available data sets are limited both in quantity and quality. Signs are composed of hand shape, hand movement, facial expression, and body posture. Every nation has its own sign language with different rules. There is no well defined and commonly accepted set of rules for sign languages. Sign languages may not be compatible with their spoken language counterparts.

Considering the insufficient advancements in the last thirty years, it does not seem feasible to develop a computer assisted international sign language recognition system in near future. To accomplish such goal, we propose the following steps to the automatic sign language recognition researchers.

An international committee for defining and publishing a standard for an international sign language should be formed. The new international sign language should have features that are easy to model with current computer systems. Then, a publicly available database containing all isolated sign and examples of the conversations should be prepared. Finally, the researchers working on different sign languages should combine the current knowledge and experience in order to develop a computer assisted sign language recognition system.

We do not expect every nation to fully adapt the new international sign language. However, if at least an international standard for a subset of the spoken language (e.g., sentences frequently used during international travel) or the finger spelling is accepted, this will make it possible to design a system that can be deployed and used anywhere in the world.

Due to the recent advancements in the imaging technologies, we anticipate that a global sign language recognition system should employ only vision-based features.

Last but not the least, we should emphasize that the Dicta-Sign project [41] has a potential to create a well established international sign language and corresponding recognition system. With the increase of funding to the project, integrating sign language recognition systems into daily life would be possible. We expect that the availability of this database will significantly boost the productivity of sign language recognition studies. We believe that it will eventually result in a recognition system that can handle either multiple sign languages or an international sign language.

7. ACKNOWLEDGMENTS

8. REFERENCES

[1] Grimes, G.J. Digital data entry glove interface device. U.S. Patent and Trademark Office, 4,414,537, 1983.

[2] Hall, J.A. The human interface in three dimensional computer art space. Media Lab, Massachusetts Institute of Technology, MSVS Thesis, Cambridge, MA, 1985.

[3] Kadous, M.W. Machine recognition of Auslan signs using PowerGloves: Towards large-lexicon recognition of sign language. Workshop on the Integration of Gesture in Language and Speech, pp. 165-174, 1996.

[4] Vogler, C. and Metaxas, D., Adapting hidden Markov models for ASL recognition by using three-dimensional computer vision methods. In Proc. of Systems, Man, and Cybernetics, vol.1, no., vol.1, pp.156-161, 1997.

[5] Starner, T. and Pentland, A. Real-time american sign language recognition from video using hidden markov models. In Motion-Based Recognition, Springer Netherlands, pp. 227-243, 1997.

[6] Brashear, H., Junker, H., Starner, T., and Lukowicz, P. Using Multiple Sensors for Mobile Sign Language Recognition. 16th International Symposium on Wearable Computers, pp. 45, 2003.

[7] Liang, R.H. and Ouhyoung, M. A real-time continuous gesture recognition system for sign language. In Proc. of Automatic Face and Gesture Recognition, pp. 558-567, 14-16 Apr 1998.

[8] Gao, W., Ma, J., Shan, S., Chen, X., Zeng, W., Zhang, H., and Wang, J. HandTalker: A multimodal dialog system using sign language and 3-D virtual human. In Proc. of Advances in Multimodal Interfaces, pp. 564-571, 2000.

[9] Hernandez-Rebollar, J.L., Lindeman, R.W., Kyriakopoulos, N. A multi-class pattern recognition system for practical finger spelling translation. In Proc. of Multimodal Interfaces Conf., pp 185-190, 2002.

[10] Mehdi, S. A., and Khan, Y.N. Sign language recognition using sensor gloves. In Proc. of Neural Information Processing, vol. 5, pp. 2204-2206, 2002.

[11] Akyol, S., and Canzler, U. An information terminal using vision based sign language recognition. In Proc. of ITEA Workshop on Virtual Home Environments, VHE Middleware Consortium, vol. 12, pp. 61-68, 2002.

[12] Kuroda, T., Tabata, Y., Goto, A., Ikuta, H., and Murakami, M. Consumer price data-glove for sign language recognition. In Proc. of 5th Int. Con. Disability, Virtual Reality Assoc., Oxford, UK, pp. 253-258, 2004.

[13] Hernandez-Rebollar, J.L., Kyriakopoulos, N., Lindeman, R.W. A new instrumented approach for translating ASL into sound and text. In Proc. of Automatic Face and Gesture Recognition, pp. 547-552 2004.

[14] Oz, C., and Leu, M.C. American Sign Language word recognition with a sensory glove using artificial neural networks. Engineering Applications of Artificial Intelligence, vol. 24(7), pp. 1204-1213, 2011.

[15] Zafrulla, Z., Brashear, H., Starner, T., Hamilton, H., and Presti, P. American sign language recognition with the kinect. In Proc. of the 13th international conference on multimodal interfaces pp. 279-286, 2011.

[16] Assaleh, K., Shanableh, T., and Zourob, M. Low complexity classification system for glove-based Arabic sign language recognition. In Neural Information Processing, pp. 262-268, 2012.

[17] Jeong, E., Lee, J., and Kim, D. Finger-gesture Recognition Glove using Velostat. 11th Int. Conf. on Control, Automation and Systems, pp. 206-210, 2011.

[18] Luzanin, O. and Plancak, M. Hand gesture recognition using low-budget data glove and cluster-trained probabilistic neural network. Assembly Automation, vol. 34(1), pp.94-105, 2014.

[19] Waldron, M.B., and Kim, S. Isolated ASL sign recognition system for deaf persons. Transactions on Rehabilitation Engineering, vol. 3(3), pp. 261-271, 1995.

[20] Fagiani, M., Principi, E., Squartini, S., and Piazza, F. A New System for Automatic Recognition of Italian Sign Language. In Neural Nets and Surroundings, Springer, pp. 69-79, 2013.

[21] Akyol, S., and Alvarado, P. Finding relevant image content for mobile sign language recognition. International Conference-Signal Processing, Pattern Recognition and Applications, pp. 48-52, 2001.

[22] Kishore, P.V.V. and Kumar, P.R. Sign language video segmentation with level sets fusing color, texture, boundary and shape features. Journal of Signal and Image Processing, vol.3, no.3, 2012.

[23] Haberdar, H. and Albayrak S. A two-stage visual Turkish Sign Language recognition system based on global and local features. Lecture Notes in Artificial Intelligence vol. 4203, Springer-Verlag, pp. 29-37, 2006.

[24] Zhang, L.G., Chen, Y., Fang, G., Chen, X., and Gao, W. A vision-based sign language recognition system using tied-mixture density HMM. In Proc. of the 6th International Conference on Multimodal Interfaces pp. 198-204, 2004.

[25] Zieren, J. and Kraiss, K.F. Robust person-independent visual sign language recognition. Pattern recognition and image analysis, Springer, pp. 520-528, 2005.

[26] Assan, M., and Grobel, K. Video-based sign language recognition using hidden markov models. In Gesture and Sign Language in Human-Computer Interaction, pp. 97-109, Springer Berlin Heidelberg, 1998.

[27] Zafrulla, Z., Brashear, H., Hamilton, H., and Starner, T. A novel approach to American Sign Language phrase verification using reversed signing. Computer Vision and Pattern Recognition Workshops, pp.48-55, 2010.

[28] Feris, R., Turk, M., Raskar, R., Tan, K. H., and Ohashi, G. Recognition of isolated fingerspelling gestures using depth edges. In Real-Time Vision for Human-Computer Interaction, Springer, pp. 43-56, 2005.

[29] Hienz, H., Bauer, B., and Kraiss, K. F. Video-based continuous sign language recognition using statistical methods. In Proc. of . 15th International Conference on Pattern Recognition, vol. 2, pp. 463-466, 2000.

[30] Dreuw, P., Rybach, D., Deselaers, T., Zahedi, M., and Ney, H. Speech recognition techniques for a sign language recognition system. Hand, 2007.

[31] Holden, E.J., Lee, G. and Owens, R. Australian sign language recognition. Machine Vision and Applications, 16(5), pp. 312-320, 2005.

[32] Dreuw, P. Steingrube, P., Deselaers, T. and Ney, H. Smoothed Disparity Maps for Continuous American Sign Language Recognition. Iberian Conf. on Pattern Recognition and Image Analysis, pp. 24-31, 2009.

[33] Wang, H., Gao, W., and Shan, S. An approach based on phonemes to large vocabulary Chinese sign language recognition. 5th Conf. on Automatic Face and Gesture Recognition, pp. 411-416, 2002.

[34] Nayak, S., Sarkar, S., and Loeding, B. Unsupervised modelling of signs embedded in continuous sentences. In Computer Vision and Pattern Recognition-Workshops, pp. 81-81, 2005.

[35] Bauer, B. and Hienz, H. Relevant features for video-based continuous sign language recognition. Int. Conference on Automatic Face and Gesture Recognition, pp. 440-445, 2000.

[36] Kelly, D., Reilly Delannoy, J., Mc Donald, J., and Markham, C. A framework for continuous multimodal sign language recognition. In Proc. of Intl. Conference on Multimodal interfaces, pp. 351-358, 2009.

[37] Fang, G., Gao, W., and Zhao, D. Large-vocabulary continuous sign language recognition based on transition-movement models. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, vol. 37(1), pp. 1-9, 2007.

[38] Wilbur, R. and Kak, A., "Purdue RVL-SLLL American Sign Language Database", School of Electrical and Computer Eng. Tech. Report, Purdue University, 2006.

[39] Dreuw, P., Neidle, C., Athitsos, V., Sclaroff, S., and Ney, H. Benchmark Databases for Video-Based Automatic Sign Language Recognition. In LREC, 2008.

[40] Johnston, T., Schembri, A., Adam, R., Napier, J., Thornton, D. Auslan SignBank: the Auslan lexical database. http://www.auslan.org.au/

[41] Schembri, A., Fenlon, J., Rentelis, R., Reynolds, S., and Cormier, K. Building the British Sign Language Corpus. Language Documentation and Conservation, vol. 7, pp. 136-154, 2013.

[42] Porfirio, A., Wiggers, K., Oliveira, S., and Weingaertner, D. Libras sign language hand configuration recognition based on 3D meshes. IEEE Int. Conf. on Systems, Man, and Cybernetics, 2013.

[43] Efthimiou, E., Fotinea, S.E., Hanke, T., Glauert, J., Bowden, R., Braffort, A., and Goudenove, F. DICTA-SIGN: Sign language recognition, generation and modelling with application in Deaf communication. 4th Workshop on the Representation and Processing of Sign Languages, pp. 80-83, 2010.

[44] Crasborn, O. and Zwitserlood, I. The Corpus NGT: an online corpus for professionals and laymen. 3rd Workshop on the Representation and Processing of Sign Languages, 2008.

[45] von Agris, U. and Kraiss, K.F. Towards a video corpus for signer-independent continuous sign language recognition. In Proc. of Gesture in Human-Computer Interaction and Simulation. Int. Gesture Workshop, 2007.

[46] Efthimiou, E. and Fotinea, S. E. GSLC: creation and annotation of a Greek sign language corpus for HCI. In Universal Access in Human Computer Interaction, Springer Berlin Heidelberg, pp. 657-666, 2007.

[47] Kishore, P.V.V., Kumar, P.R., Kumar, E.K., and Kishore, S.R C. Video Audio Interface for Recognizing Gestures of Indian Sign. International Journal of Image Processing, vol. 5, no. 4, pp. 479, 2011.

[48] Bungerot, J., Stein, D., Dreuw, P., Ney, H., Morrissey, S., Way, A., and van Zijl, L. The ATIS sign language corpus. 2008.

[49] Fagiani, M., Principi, E., Squartini, S. and Piazza, F. A New Italian Sign Language Database. In Proc. of International Conference on Brain Inspired Cognitive Systems (BICS), Shenyang, China, Jul. 11-14 2012.

[50] Kim, J. S., Jang, W., and Bien, Z. A dynamic gesture recognition system for the Korean sign language (KSL). Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 26(2), pp. 354-359, 1996.

[51] Terrillon, J.C., Piplr, A., Niwa, Y., and Yamamoto, K. Robust Face Detection and Japanese Sign Language Hand Posture Recognition for Human-Computer Interaction in an Intelligent Room. In Proc. of Int'l Conf. Vision Interface, pp. 369-376, 2002.

[52] Al Qodri Maarif, H., Akmeliawati, R., Bilal, S. Malaysian Sign Language database for research. In Proc. of Computer and Communication Engineering, pp. 798-801, 3-5 July 2012.

[53] Kausar, S., Javed, M.Y., and Sohail, S. Recognition of gestures in Pakistani SL using fuzzy classifier. In Proc. of 8th Conf. on Signal Processing, Comp. Geometry and Artificial Vision, pp. 101-105, 2008.

[54] Karami, A., Zanj, B., and Sarkaleh, A.K. Persian sign language (PSL) recognition using wavelet transform and neural networks. Expert Systems with Applications, vol. 38(3), 2011.

[55] San-Segundo, R., Pardo, J.M., Ferreiros, J., Sama, V., Barra-Chicote, R., Lucas, J. M., and García, A. Spoken Spanish generation from sign language. Interacting with Computers, vol. 22(2), pp. 123-139, 2010.

[56] Aran, O., Ari, I., Guvensan, A., Haberdar, H., Kurt, Z., Turkmen, I., Uyar, A., and Akarun, L. A Database of Non-Manual Signs in Turkish Sign Language. Signal Processing and Communications Applications, 2007.

[57] Tai, J.H.Y. and Tsay, J.S. Taiwan Sign Language Corpus: Digital Dictionary and Database. TELDAP Intl. Conference, Taipei, Taiwan, pp 41-47, 2010.

[58] Hieu, D.V. Vietnamese sign language recognition for hearing impaired people using fuzzy hidden Markov models. Master of Science Thesis, Department of Information Technology, King Mongkut's Univesity of Technology, 2008.

[59] Picone, J.W. Signal modeling techniques in speech recognition. Proceedings of the IEEE, vol. 81, no. 9, pp. 1215-1247, 1993.

[60] Efthimiou, E., Fotinea, S. E., Vogler, C., Hanke, T., Glauert, J., Bowden, R., and Segouat, J. Sign language recognition, generation, and modelling: a research effort with applications in deaf communication. In Universal Access in Human-Computer Interaction. Addressing Diversity, Springer Berlin Heidelberg, pp. 21-30, 2009.

[61] Rabiner, L. A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE, 77(2), pp. 257-286, 1989.

[62] Starner, T.E. Visual recognition of American Sign Language using hidden Markov models. MIT, Cambridge, Dept of Brain and Cognitive Sciences, 1995.

[63] Ouhyoung, M., and Liang, R. H. A sign language recognition system using hidden Markov model and context sensitive search. In Proc. of the ACM Symp. on virtual reality software and technology, pp. 59-66, 1996.

[64] Wang, H., Leu, M.C., and Oz, C. American Sign Language Recognition Using Multi-dimensional Hidden Markov Models. Journal of Information Science and Engineering, 22(5), pp. 1109-1123, 2006.

[65] Al-Rousan, M., Assaleh, K., and Tala'a, A. Video-based signer-independent Arabic sign language recognition using hidden Markov models. Applied Soft Computing, 9(3), 990-999, 2009.

[66] Haberdar, H. and Albayrak, S. Real Time Isolated Turkish Sign Language Recognition from Video Using Hidden Markov Models with Global Features. ISCIS, Lecture Notes in Computer Science, Springer, pp. 677-687, 2005

[67] Grobel, K. and Assan, M. Isolated sign language recognition using hidden Markov models. IEEE Intl. Conf. on Systems, Man, and Cybernetics, vol.1, pp.162-167, 1997.

[68] Hienz, H., Bauer, B., and Kraiss, K.F., "HMM-based continuous sign language recognition using stochastic grammars", in Gesture-Based Communication in Human-Computer Interaction, Springer, pp. 185-196, 1999.

[69] Dreuw, P., Deselaers, T., Rybach, D., Keysers, D., and Ney, H. Tracking using dynamic programming for appearance-based sign language recognition. 7th International Conference on Automatic Face and Gesture Recognition, pp. 293-298, 2006.

[70] Fang, G., Gao, X., Gao, W., and Chen, Y. A novel approach to automatically extracting basic units from Chinese sign language. In 17th International Conference on Pattern Recognition,. vol. 4, pp. 454-457, 2004.

[71] Hieu, D.V. and Nitsuwat, S. Image preprocessing and trajectory feature extraction based on hidden markov models for sign language recognition. In Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, pp. 501-506, 2008.

[72] Holden, E. J., Lee, G., and Owens, R. Australian sign language recognition. Machine Vision and Applications, vol. 16(5), pp. 312-320, 2005.

[73] .Kelly, D., McDonald, J., and Markham, C. Evaluation of threshold model HMMS and Conditional Random Fields for recognition of spatiotemporal gestures in sign language. 12th Int. Computer Vision Workshops, 2009.

[74] Eddy, S. R. What is a hidden Markov model?. Nature biotechnology, 22 (10), pp. 1315-1316, 2004.

[75] Evermann, G., Kershaw, D., Moore, G., Odell, J., Ollason, D., Valtchev, V., and Woodland, P. The HTK book, vol. Cambridge: Entropic Cambridge Research Laboratory, 1997.