Review of Literature for the Development of Indian Sign Language Recognition System

Shweta Dour
Research Scholar,
Bhagwant University ,
Ajmer.

M.M. Sharma, PhD
Principal
Govt. Engineering College, Ajmer.

ABSTRACT

In recent years, sign language recognition has attracted much attention in computer vision . A sign language is a means of conveying the message by using hand, arm, body, and face to convey thoughts and meanings. Like spoken languages, sign languages emerge and evolve naturally within hearing-impaired communities. However, sign languages are not universal. There is no internationally recognized and standardized sign language for all deaf people. As is the case in spoken language, every country has got its own sign language with high degree of grammatical variations. The sign language used in India is commonly known as Indian Sign Language (henceforth called ISL).

General Terms

Computer vision , Wearable Computing , Spatio Temporal approach, Kinect Sensor.

Keywords

Sign Language Recognition, Isolated sign recognition and Continuous sign language recognition, 3D tracking system.

1. INTRODUCTION

Approaches to sign language recognition can be divided into isolated sign recognition and continuous sign language recognition. Isolated sign recognition is a special case of gesture recognition. Gesture recognition systems are typically designed to recognize artificial gestures. The user has to learn these gestures in order to communicate with the system. In order to have an idea about the best approach needed to build a sign language recognition system it becomes crucial to survey the already developed systems. For this purpose the research over the electronic databases and study of few papers that were most relevant to the system has been carried out. An analysis and comparison of the methods applied for development of the similar systems would help to apply the suitable approach to develop the system. The literature surveyed are summarized below. Several different approaches have been used in order to design Sign Language Recognition Systems.

2. WEARABLE COMPUTING BASED ACQUISITION

Wearable computing approaches[1] to sign language data acquisition offer accurate means of extracting information about the signers hand movements and hand shape. Each sensing technology varies along several dimensions, including accuracy, resolution, range of motion, user comfort, and cost . Glove based[2] approaches to gesture recognition generally known to be the approaches where it is required that the user has to wear   certain glove like device over their hands comprising of various kinds of sensors mounted on the devices.

Ruize *et al* [3] presented nonspecific person gesture recognition system by using MEMS accelerometers . The recognition system consisted of sensor data collection, segmentation and recognition. After receiving acceleration data from the sensing device, a segmentation algorithm was applied to determine the starting and end points of every input gesture automatically. They presented their work for seven different hand gestures such that included left, right, tick, cross, circle, up and down, and the input signals for which were obtained from MEMS 3-axes accelerometers. A total of 72 experiments were performed for a sequence of hand gestures ,and it was found that recognition accuracy of individual gesture was ranging from 91% to 100% and the overall accuracy was found to be of 95.6%.

Yun *et al* [4] proposed a combined ACC and sEMG approach for continuous SLR(Sign Language Recognition) based on the sign-component level analysis. They developed a combined ACC and sEMG approach for recognizing Chinese Sign Lnaguage at the subword level. The accelerometer (ACC) and the surface electromyographic (sEMG) sensors provide two potential techniques for gesture sensing. Multichannel sEMG signals,(Figure 2.1) recorded by a set of sEMG sensors placed on the arms, contain rich information for the coactivation and coordination of multiple muscles associated with different sign gestures. Yun was able to use the extracted features from sEMG and   ACC to generate three basic components related to sign. language,that were the hand shape, orientation, and movement.
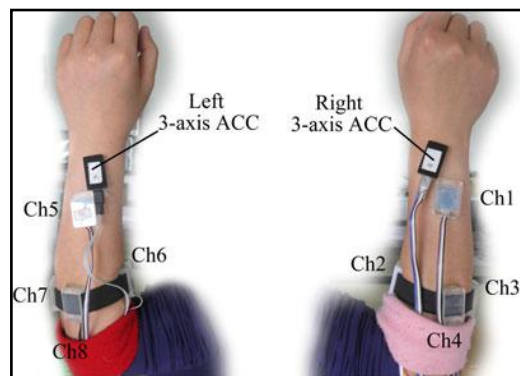

Figure 2.1 Positions of the ACC and sEMG sensors on the left and right arms.Here, "Ch" refers to the channel of EMG sensor proposed by Yun[4]

If recognition results considered then the system could recognize a sequence of subwords by combining the the similarities of the desired gesture at the component level. The recognition of 96.5% for a vocabulary of 120 signs and recognition of 86.7% for 200 sentences was obtained for the designed system . Vasiliki *et al* [5] developed a sign language recognition system by the use of hybrid adaptive weighting (HAW) process applied to sEMG and 3-D accelerometer data which related the individual way of signing with the signer's level of deafness (LoD). The resultant set of features, named as weighted intrinsic-mode entropy (wIMEn), help to increase accuracy of the recognition system. Zhang *et al* [6] was able to

obtain results for hand gesture recognition by utilizing the inputs obtained by two types of sensors namely multichannel electromyography (EMG) sensors and a three-axis accelerometer (ACC) . They were able to recognize the required gesture by determining the start point and the end point of the gesture.For classification of the gestures multistream hidden Markov models and decision tree were utilized together for obtaining better recognition. The experiments to find the classification for 72 CSL words depicted that showed that the approach was effective to merge EMG and ACC information , and also average accuracies of 96.3% and 95.3% was obtained for two different subjects.

Naik *et al* [7] was able to overcome the hurdles posed by unbalanced multicategory datasets by employing the Twin SVM. Data for which patterns are present in different classes and are from different distributions are to be considered as two separate entities. If it is not considered it may result in misclassification of the patterns. Naik achieved the gesture recognition using Twin SVM and found it as a better choice for the required task with multicategory datasets as compared to the sEMG sensors .
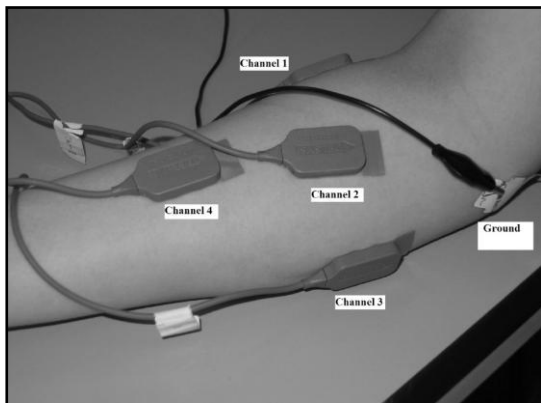

Fig2.2 Hand gesture experimental setup with four electrodes proposed by Naik [7]

The system was able to accurately identify actions by using sEMGs recorded from near the upper region of the elbow as show in Fig 2.2, making it suitable for applications such as transradial amputees.

An inexpensive visual motion data glove with high recognition accuracy was presented by Han *et al* [8]. The glove device used a single-channel video instead of commonly used motion-sensing fibers or multi-channel videos, with a reconstruction algorithm to compensate for the shortcomings of single-channel videos. A monocular camera was used to capture hand motion, and then a visual analyzer algorithm detected those optical indicators and recovered 3D positions of fingers and joints. Three classes of scenarios (left/right clicks, numbers, and the OK sign) were computed and reconstructed into 3D images in MATLAB.

Wagner *et al* [9] developed a 3D hand motion tracking and gesture recognition system using a data glove called KHU-l. The data glove interacted with a PC via a bluetooth device. It successfully performed hand motion tracking such as fist clenching, hand stretching and bending. Three gestures (scissor, rock, and paper) were tested with an accuracy of 100% for 50 trials each. The recognition in 3D and the wireless transmission were good improvements, but they incurred time delay.

Kim *et al* [10] proposed a system which combined sensor data from accelerometers and an Electromyogram (EMG) which was used to measure the electrical activity produced by the hand muscles. It was shown that the information added by the EMG greatly improved the recognition rate of signs. Figure2.3 shows a visualisation of the sensor setup for a single hand.
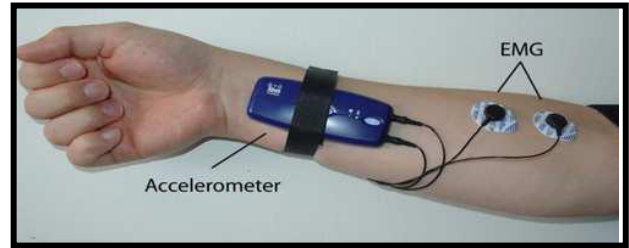

Figure 2.3 : Sensor Placement for Bi-Channel recognition system proposed by Kim [10]

Similarly, Oz *et al* [11] utilised a Cyberglove TM in association with a 3-D motion tracker that was called "Flock of Birds" to determine the features of hand posture . Figure 2.4 shows the CybergloveTM and "Flock of Birds" 3-D motion tracker.
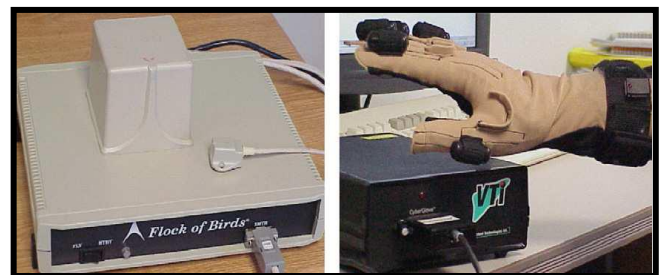

Figure 2.4: Left: Flock of Birds 3D motion tracker, Right: CyberGlove proposed by Oz [11]

Wang *et al* [12] proposed a viewpoint invariant data acquisition method. Their method was based on a virtual stereo vision system, using one camera and gloves with a specially designed colour pattern to indicate the 5 separate fingers, palm and back. Figure 2.5 shows a visualisation of the design of the gloves.
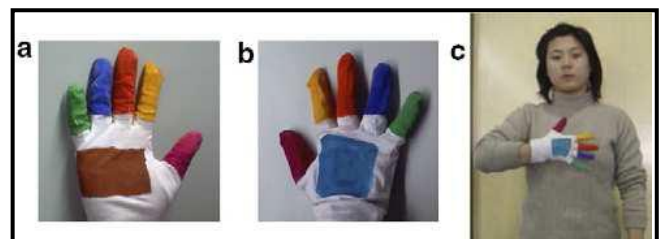

Figure 2.5: Gloves used by Wang [12]

Bui *et al* [13] developed a new SLR system for the Vietnamese Sign Language. They utilized an extra sensor which was mounted at the back side of the handin order to obtain better recognition . A different classification for letters was introduced, in which all letters were divided into three clusters according to the X-axis value of the palm, and three fuzzy rule-based systems were constructed for further classification without having to collect training samples. Bending or flexing of fingers were measured in five levels: Very-Low, Low, Medium, High, Very-High.The system consists of six ADXL202 accelerometers for sensing the hand posture,a BASIC Stamp microcontroller Fig 2.6. They concentrated over the 23 letters of Vietnamese sign Language which also included two basic postures of"punctuation" and "space" .

Kenn *et al* [14] came up with a way to integrate glove-based devices into multiple applications with the help of a context framework. The textile glove with integrated electronics device was able to perform in at least three applications as demonstrated: to move/zoom/select parts of a map, to navigate to a remote control in presentation, and to direct a toy robot to move left/right. forward/ backward.

Fig.2.6: Sensing glove with six accelerometers and a basic stamp microcontroller proposed by Bui [13]

One problem was that this device could only detect gestures in the X and Y axes, without the ability to detect motion in the Z axis such as the so-called "yaw". In addition. to achieve wearability, light weight, and cool appearance, the accuracy of recognition was sacrificed.

Another data glove based system was proposed by McGuire *et al*[15] where a mobile sign language translator was implemented using an Acceleglove (Fig 2.7) .
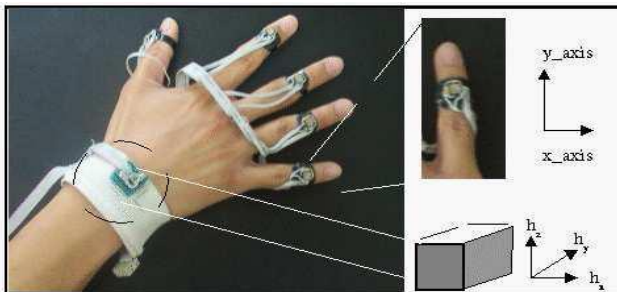

Figure 2.7: The Acceleglove proposed by McGuire [15]

Fang *et al* [16][17] developed a large vocabulary sign recognition systems using two CyberglovesTM and three "Pohelmus" 3SPACE-position trackers. Two trackers were positioned on the wrist of each hand and another positioned on the signers back and were used to collect orientation and position data. The CyberglovesTM collected 18-dimensional hand shape information for each hand.

Another novel approach to sign language data acquisition was taken by Brashear *et al* [18] where features from both a hat mounted camera and accelerometer data were used to classifying signs Figure 2.8. While wearable computing approaches to data acquisition can extract accurate features representing the signs being performed, some of these approaches require that the signer wears cumbersome devices which can hinder the ease and naturalness of signing.

Figure 2.8: Head mounted camera and accelerometer data collection implemented by Brashear [18]

An alternative approach is to acquire gesture data through a camera based input. In order to capture gesture based information from camera based inputs, the hands must be located in the image sequence and this is often carried out using color, motion and edge information .Number of works have proposed different techniques for the segmentation of hands from an image sequence .

3. SIGN LANGUAGE RECOGNITION: SPATIOTEMPORAL GESTURE RECOGNITION

Hossein et al [19] presented a model for conceptual imitation learning. Conceptual imitation tries to model true imitation in human which is said to be fulfilled by abstraction, generalization, recognition, and regeneration of action. They used , Hidden Markov models for the purpose of recognition of various modalities defined for their text. They also introduced a method to combine the perceptions of various modalities for the recognition of multimodal concept. One of them was imitation learning based on a few hand gestures by the interaction with the teachers. For the task, variety of experiments were performed on the basis of perceptions from different modalities, which included the vision, motor, and audition. Secondly it was required to learn a set of actions by recognizing their emotional effects. The results of the experiments conducted on a humanoid robot were able to show the efficacy of their proposed model for conceptual imitation learning.

Annamária et al [20] presented the concept and implementation of a hand posture and gesture modeling and recognition system . The system transformed the data of the hands obtained after processing to develop a feature model which was actually a fuzzy inference system of hands that was used as the input to a fuzzy neural network and based on the fuzzy inference model generated determines the actual hand posture . Then , determining the order of the input hand gestures the recognition of the gesture is achieved. As a feature model, it introduced a new fuzzy hand-posture model (FHPM). For hand-posture recognition, a modified Circular Fuzzy Neural Network (CFNN) architecture was proposed together with a reduced time training procedure. However, the correct identification rate proved to be in average above 96%, currently with a limitation of the need of using a homogenous background.

Dreuw et al [21] provided criteria to evaluate hand and head tracking algorithms; the proposed criteria were tracking error rate (TER) and word error rate (WER). Dynamic programming tracking (DPT) avoided taking possibly wrong local decisions by tracking back from the last frame of the sequence to the first frame to get a best path with the highest score given by a score function, which was calculated for every frame starting from the very beginning. By integrating into an HMM, this model allowed simultaneous tracking and recognition. However, it was

not robust enough to perform on the dataset due to the high number of near-profile images in the database.

Alon et al [22] proposed a framework where the system was able to recognize the gesture for even a location of hand that was very ambiguous and even when there was no data available for the start or the end point of the gesture. Hence the system could be utilized for continuous streams of images for which the gestures being performed even for complex backgrounds with moving hands. The performance of the system was estimated for two applications as to the digits recognition for the signs of hand gestures by users in front of a complex background, with short-sleeved shirts, and for the recognition of occurrences of signs in a video database containing continuous signing inputs for American Sign Language .

Buehler et al. [23] detected hands and arms at the same time. During the recognition, every pixel was assigned to either the limb model or the background model to avoid the ambiguity from hand, forearm, and upper arm occlusion. Distinctive, unambiguous frames were identified and linked together by tracking the arm configurations. This system was tested with BBC broadcast footage using a continuous sequence of 6,000 frames performed by three signers. The result accuracy was above 91 % when there was little arm overlap, and it was at least 59% when there was much overlap.

Shanabeh et al [24] proposed an isolated temporal gesture technique for the recognition of Arabic sign language. They proposed temporal features which were extracted through forward, backward, and bi-directional predictions. Experiments, based on a database of isolated signs, showed that their method achieved a classification performance ranging from 97% to 100% when classifying 23 different sign classes.

Wang et al [25] proposed a view invariant sign recognition system. In their proposed system, the recognition task was converted to a verification task based on the geometric constraint that the fundamental matrix associated with two views should be unique when the observation and template signs are obtained synchronously under virtual stereo vision and viceversa. Experiments conducted on a vocabulary of 100 signs, where 5 isolated samples of each sign were recorded, showed their method achieved an accuracy of 92%.

Infantino et al [26] proposed an approach for the recognition of Italian sign language and the system integrated the concept of commonsense engine for the purpose of detecting sentences in the context. The proposed architecture allowed the knowledge of the recognition process to be modeled and managed in a simple and robust way. Moreover, the introduction of the semantic context resolved the problem of the analysis and validation of a sentence. The experiments presented a test set of 30 videos of sentences using 20 Signs and 80 videos of sentences using 40 signs which showed that the system maintained a high recognition rate.

Cooper et al [27] implemented an isolated sign recognition system using 1st order Markov Chains. In their model, signs were broken down in visemes (equivalent to phonemes in speech) and a bank of Markov Chains were used to recognise the visemes as they were produced. The first layer had three types of classifiers that detected tab (placement), sig (movement) and ha (arrangement) from visemes, which are visual representations of speech sounds in sign languages. The experiments of 164 signs with 1,640 examples, the first layer achieved around 30% accuracy; when combined with the second layer, the accuracy was increased to 72.6%, a significant

improvement by using two levels of classifiers for a large lexicon.

Deselaers et al [28] performed Dynamic programming tracking (DPT) in hand tracking, which is a model-free and signer-independent tracking method. DPT avoided taking possibly wrong local decisions by tracking back from the last frame of the sequence to the first frame to get a best path with the highest score given by a score function, which was calculated for every frame starting from the very beginning. By integrating into an HMM, this model allowed simultaneous tracking and recognition.

Agris et al [29] proposed a system of isolated sign language which was developed by combining A Posteriori estimation and Maximum Likelihood Linear Regression method . they recognized single handed isolated signs by the above combinational approaches. They implemented selected adaptation methods from speech recognition to improve the performance of their system when performing user independent recognition. A recognition rate of 78.6% was reported when recognising 153 isolated signs.

Juang et al [30] proposed a method for temporal sequence processing by means of recurrent fuzzy network. They applied their method to a gesture recognition task and experiments showed a recognition rate of 92%.

Ong et al [31] provided another approach in which they used a boosted cascade of classifiers that detected hand shapes in grey scale images in order to track hands. These images were grouped by applying the K-means clustering algorithm. Then a tree structure was formed to contain two layers of "weak" hand detectors. The classifiers at the first layer chose all candidate image blocks that may contain hand shapes, and then passed them to the second layer to be compared to all the images in the corresponding clusters. The FloatBoost algorithm was used to perform boosting in finding weak classifiers. In an experiment of 5,013 hand images, the first layer achieved 98% accuracy in detecting the existence of hands, and the second layer achieved 97.4% accuracy in hand shape detection.

Fang et al [32] addressed the problem of large vocabulary sign recognition by proposing g a combination of self organising feature maps, HMMs a hierarchical decision tree, with low computational costs, for the recognition of isolated signs. Experiments were conducted on a data set of 61365 isolated samples of 5113 different signs. Results showed an average recognition rate of 91.6%.

Yang et al [33] extracted motion trajectories from American Sign Language (ASL) videos and classified signs using a time delay neural network. Experiments based on a vocabulary of 40 signs showed the average recognition rate of unseen test trajectories was 93.4%.

Starner et al [34] presented two vision-based SLR systems using hidden Markov models one used a second-person view with a desk mounted camera and the other was the first-person view with a camera mounted on a hat worn by the user. HMM was used for training and continuous motion tracking. Both systems used a skin color matching algorithm for hand tracking In comparison, the second-person view system had a word accuracy of 92% while the first-person view system had a word accuracy of 98%. The high accuracy indicated that HMM is good for the purpose of continuous motion tracking.

The research into spatiotemporal gesture and sign recognition has two main categories: isolated and continuous recognition. Isolated recognition focuses on the classification of a single

hand gesture. In continuous recognition, the user performs gestures one after the other and the aim is to spot and classify meaningful gesture segments from within the continuous stream of sign language. Isolated Gesture Recognition

4. GESTURE RECOGNITION USING DEPTH DATA

Maisto et al [35] proposed an algorithm with three different variations to recognize and correctly identify the fine gestures of hands, in particular fingertips, through the use of the Kinect motion sensor. The proposed methods do work in real time without any noticeable lag and they exhibit good performances in the tests that carried out in this regard. The presented system could be improved by the help of a neural network, which should be capable of learn new poses and store them in a database, allowing users to create their own customizable gesture database. In this way, it was possible to overtake the problems seen in the FT-CDB and FT-CMB algorithms and hence, achieve a better performance as well.

Vanden Bergh et al [36] used a ToF camera instead of stereoscopic cameras to enhance recognition. A ToF camera with a low resolution (176 x 144 pixels) was used to get depth image for segmentation, and it was paired with an RGB camera with a high resolution (640 x 480 pixels) for hand detection. The skin color used for detection was decided by a pre-trained, adaptive skin color model, which was updated with color information taken from the face. Three situations· were evaluated in hand detection: the hand was next to the face, the hand overlapped with the face, and a second person was behind the tester. Depth-based detection achieved more than 98% accuracy in all three situations, while the accuracy of color-based detection decreased dramatically from 92% of the first situation to 19.8% of the third situation.

Argyros et al. [37] proposed a 3D hand tracking method using a stereoscopic system of two video streams. It combined color-based tracking and 3D hand tracking, and it could run in real time. Images were captured by each camera; hand blobs were marked by 2D color-based hand trackers in both video streams, and then matched together by a calibration calculation. After this, hand contours were aligned and reconstructed in the 3D space.. The ability of processing in real time isan advantage, yet the estimated depth data using the stereoscopic system is rather noisy, and the need for calibration increases system complexity.

Nanda et al. [38] presented a method to track hands in highly cluttered environments using 3D depth data provided by a sensor that was based on the time-of-flight (ToF) method for the purpose of obtaining color data and depth data simultaneously . The potential fields of possible hands or face contour were calculated by three algorithms: getting potential fields by using distance transform, k-components based potential fields with weights. and basin of attraction. The system was tested in head tracking and hand tracking on ten people with good results. It was able to track hands with occlusion, and to recognize simple motion like "step back" and "stop".

5. GESTURE RECOGNITION USING KINECT

The recently developed depth sensors, e.g., the Kinect sensor, has gained popularity I the field of human-computer interaction (HCI). Tremendous amount of progress has been achieved by means of the the Kinect sensor ,in the field of human action recognition , tracking human body and face recognition etc. Hand Gesture recognition is still found to be a tedious task even if the features obtained by the Kinect Sensors. Hand seems to be smaller in ratio when compared to the complete human body , also having a complex geometry associated with it. Thus the task of recognizing gestures made by hand remains to be a problem to be solved . After the launch of Microsoft Kinect in November, 2010, several exciting recognition systems based on this device were developed in less than 18 months. The resolution of its HGB camera and depth camera are both 640 x 480 pixels, which are fairly sufficient under many situations.

Ren et al [39] presented a robust part-based hand gesture recognition system using the Kinect sensor. A novel distance metric, Finger-Earth Mover's Distance (FEMD), was used for dissimilarity measure, which represents the hand shape as a signature with each finger part as a cluster and penalizes the empty finger-holes. They focused on developing a hand gesture recognition that was based on parts of hands using Kinect sensor. To handle the noisy hand shapes obtained from the Kinect sensor, they proposed a novel distance metric, Finger-EarthMover's Distance (FEMD),that was used estimate the dissimilarity between different hand shapes. The proposed FEMD distance metric was based on a part-based representation which represents a hand shape as a signature with each finger part as a cluster. Such a representation enables the computation on the global features, thus it is robust to local distortions and it is robust to articulation, orientation and scale changes.

Sun et al [40] came up with a discriminative exemplar coding (DEC) approach, and also utilized the Kinect sensor, for the purpose of moldelling the different signs. The approach can be summarized as three steps. First, of input candidate exemplars were considered for learning based on the videos of sign language for every category of sign taking into consideration the discrimination among them. On one hand, they efficiently conducted background modeling to extract human body and locate the hand position in frames. On the other hand, they obtained more discriminative features, as well as sign description. Based on this descript on, the MIL was employed to learn the similarities between frames. Based on the learned exemplar based classifiers, each sign bag were described, then AdaBoost was employed to select the most discriminative features to form a strong classifier.

Yang et al [41] proposed a gesture recognition system using depth information provided by Kinect, and implemented in a media player application. It was able to recognize eight gestures to control the media player, with a maximum confusion rate of 8.0.This system demonstrated the applicability of using Kinect for gesture recognition in a contact-less user interface.

A method to track fingertips and the centers of palms using Kinect was presented by Raheja et al. [42]. It applied thresholding on the depth of hand regions for segmentation. Then the palm was filtered and subtracted from the hand, so that only the fingers were left in the image. Therefore, by determining the minimum depth, fingertips were found. The center of the palm was determined by finding the maximum of distance within the image of the hand. When fingers were extended, the accuracy of detecting fingertips was nearly 100% accuracy, and that of palm centers was around 90%. However this method did not attempt at gesture recognition.

Rebollar et al [43] proposed another approach using depth data provided by Kinect to detect fingertips. First, it found hand points by thresholding on depth data, and then generated the convex hull containing the hand by Graham Scan Fingertips were detected by calculating the angle between candidate points. After fingertips were found, the mouse clicking motion was recognized and tested on the popular game Angry Bird.

6. PROPOSED APPROACH FOR RECOGNITION OF INDIAN SIGN LANGUAGE

The investigation over the use of Spatio-Temporal Features to accuarately describe the motion of sign, using boosting to combine the most effective features into a single strong classifier per sign. Sign is a temporal medium and as such it is logical to include some temporal information in the symbols describing each frame. It is also required to add identifiers so that the learning mechanism can distinguish between the head and the 2 different hands. To this end cluster indexes are concatenated across 2 frames and an identifier added to the same. The integral image provides a rapid method to calculate the block features in an image. Taking this into the temporal domain, in a video stream one can imagine the images stacked one behind the other in temporal order.

The use of a double image buffer system is proposed that updates two image buffers simultaneously. While one buffer is being restarted the other is used to perform the calculations. A pictorial representation is shown in figure 6.1, as can be seen the buffers are offset by half their size, and therefore work out of step with each other, when the first is half full the other starts from zero and fills, when the first buffer is full the window switches into the second buffer which is only half full and the first buffer is restarted. This way, at any frame there is always a full window of history yet the total value of any point is limited since there is a maximum of two full windows calculated.
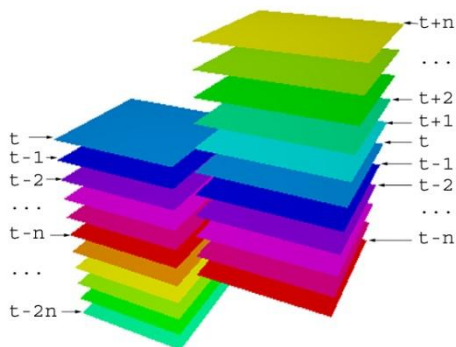

Figure 6.1 - The dual image buffers

7. CONCLUSION

The survey of various approaches related to the Sign Language Recognition Systems infers that the design of the sign language recognition systems started with the wearable computing approaches which included certain sensors associated with the hands or some coloured gloves wore over the hands and gradually with time the recognition systems were utilizing the spatiotemporal approaches for the purpose. The spatiotemporal approaches captured the videos for bare hands and achieved the sign language recognition by processing the data and making use of various machine learning algorithms. The latest approaches for the sign language recognition systems make use of Depth data and Kinect Sensors for the obtaining the features from the hands signs in the videos. Still aiming to design a system for bare hands in the videos will be cost effective and more user friendly for people.

8. REFERENCES

[1] Sigal Berman, and Helman Stern , "Sensors for Gesture Recognition Systems" IEEE Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews, Vol. 42, No. 3, MAY 2012 pg 277

[2] Laura Dipietro, Angelo M. Sabatini, and Paolo Dario, "A Survey of Glove-Based Systems and Their Applications" IEEE Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews, VOL. 38, NO. 4, JULY 2008 pg 461

[3] Ruize Xu, Shengli Zhou, and Wen J. Li, "MEMS Accelerometer Based Nonspecific-User Hand Gesture Recognition" IEEE Sensors Journal, Vol. 12, No. 5, MAY 2012

[4] Yun Li, Xiang Chen, Xu Zhang, , Kongqiao Wang, and Z. Jane Wang, "A Sign-Component-Based Framework for Chinese Sign Language Recognition Using Accelerometer and sEMG Data" IEEE Transactions On Biomedical Engineering, Vol. 59 No. 10, October 2012 pp 2695-2704

[5] Vasiliki E. Kosmidou, , Panagiotis C. Petrantonakis, and Leontios J. Hadjileontiadis, "Enhanced Sign Language Recognition Using Weighted Intrinsic-Mode Entropy and Signer's Level of Deafness" IEEE Transactions On Systems, Man, And Cybernetics—Part B: Cybernetics, Vol. 41, No. 6, December 2011 pp 1531-1542

[6] Xu Zhang, Xiang Chen, , Yun Li, Vuokko Lantz, Kongqiao Wang, and Jihai Yang "A Framework for Hand Gesture Recognition Based on Accelerometer and EMG Sensors" IEEE Transactions On Systems, Man, And Cybernetics—Part A: Systems And Humans, Vol. 41, No. 6, November 2011

[7] Ganesh R. Naik, Dinesh Kant Kumar, and Jayadeva, " Twin SVM for Gesture Classification Using the Surface Electromyogram" IEEE Transactions On Information Technology In Biomedicine, Vol. 14, No. 2, MARCH 2010 pg 301

[8] Y Han. "A low-cost visual motion data glove as an input device to interpret human hand gestures". IEEE Trans Conscum Electron, 56(2):501-509, 2010.

[9] Johannes Wagner, Jonghwa Kim, Matthias Rehm, and Elisabeth Andre. " Bi-channel sensor fusion for automatic sign language recognition."8th IEEE International Conference on Automatic Face & Gesture Recognition, pages 1–6, September 2008.

[10] Jonghwa Kim, Johannes Wagner, Matthias Rehm, and Elisabeth Andre. Bi-channel sensor fusion for automatic sign language recognition. 2008 8th IEEE International Conference on Automatic Face & Gesture Recognition, pages 1–6, September 2008.

[11] Coz and M Leu. "Linguistic properties based on American Sign Language isolated word recognition with artificial neural networks using a sensory glove and motion tracker." NEUROCOMPUTING, 70(16-18):2891–2901, 2007.

[12] Qi Wang, Xilin Chen, Liang-Guo Zhang, Chunli Wang, and Wen Gao. Viewpoint invariant sign language recognition. Computer Vision and Image Understanding, 108(1-2):87–97, 2007.

[13] Duy Bui and Long Thang Nguyen, Recognizing Postures in Vietnamese Sign LanguageWith MEMS Accelerometers IEEE Sensors Journal, VOL. 7, NO. 5, MAY 2007

[14] Holger Kenn, Friedrich Van Megen, and Robert Sugar. A glove-based gesture interface for wearable computing

applications. (IFAWC), 2007 4th International Forum on Applied Wearable Computing, pages 1 -10, 2007.

[15] R.M. McGuire, J. Hernandez-Rebollar, T. Starner, V. Henderson, H. Brashear, and D.S. Ross. Towards a one-way American Sign Language translator. Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Pro- ceedings., pages 620–625, 2004.

[16] Wen Gao, Gaolin Fang, Debin Zhao, and Yiqiang Chen. "Transition movement models for large vocabulary continuous sign language recognition." IEEE FG 2004, pages 553–558, May 2004.

[17] Gaolin Fang, Wen Gao, and Debin Zhao. "Large vocabulary sign language recognition based on hierarchical decision trees." Proceedings of the 5th international conference on Multimodal interfaces - ICMI '03, page 125, 2003.

[18] H. Brashear, T. Starner, P. Lukowicz, and H. Junker. "Using multiple sensors for mobile sign language recognition."Seventh IEEE International Symposium on Wearable Computers, 2003Proceedings., pages 45–52, 2003.

[19] Hossein Hajimirsadeghi, Majid Nili hmadabadi, and Babak Nadjar Araabi , "Conceptual mitation Learning Based on Perceptual and Functional Characteristics of Action." IEEE Transactions On Autonomous Mental Development .2013

[20] Annamária R. Várkonyi-Kóczy, and Balázs Tusor, "Human–Computer Interaction for Smart Environment Applications Using Fuzzy Hand Posture and Gesture Models". IEEE Transactions On Instrumentation And Measurement, Vol. 60, No. 5, May 2011 ,1505

[21] Philippe Dreuw, Jens Forster, and Hermann Ney. Tracking benchmark databases for video~based sign language recognition. In ECCV International Workshop on Sign Gesture and Activity SGA Crete Greece September 2010.

[22] Jonathan Alon, Quan Yuan, and Stan Sclaroff, "A Unified Framework for Gesture Recognition and Spatiotemporal Gesture Segmentation" IEEE Transactions On Pattern Analysis And Machine Intelligence, VOL. 31, NO. 9, SEPTEMBER 2009

[23] P. Buehler, M. Everingham, D. P. Huttenlocher, and A. Zisserman. "Long term ann and hand tracking for continuous sign language TV broadcasts." In British Machine Vision Conference, 2008.

[24] T. Shanableh, K. Assaleh, and M. Al-Rousan. "Spatio-temporal feature-extraction techniques for isolated gesture recognition in Arabic Sign language." Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 37(3):641–650, June 2007

[25] Qi Wang, Xilin Chen, Liang-Guo Zhang, Chunli Wang, and Wen Gao. "Viewpoint invariant sign language recognition." Computer Vision and Image Understanding, 108(1-2):87–97, 2007.

[26] Ignazio Infantino, Riccardo Rizzo,and Salvatore Gaglio "A Framework for Sign Language Sentence Recognition by Commonsense Context" IEEE Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews, VOL. 37, NO. 5, SEPTEMBER 2007

[27] Helen Cooper and Richard Bowden. "Large lexicon detection of sign language." In ICCV, Workshop Human Comp. Inter, 2007.

[28] T Deselaers ,P Dreuw, , D Rybach, D Keysers, and H Ney. "Tracking using dynamic programming for appearance based sign language recognition." 7th International Conference on Automatic Face and Gesture Recognition FGR06, 62(1):293-298,2006.

[29] Ulrich von Agris, Daniel Schneider, Jorg Zieren, and Karl-Friedrich Kraiss. "Rapid signer adaptation for isolated sign language recognition." In CVPRW '06: Proceedings of the2006 Conference on Computer Vision and Pattern Recognition Workshop, page 159, IEEE ComputerSociety Washington, DC, USA, 2006.

[30] C.-F. Juang and Ksuan-Chun Ku. "A recurrent fuzzy networkfor fuzzy temporal sequence processing and gesture recognition.Systems" Man, and Cybernetics, Part B: Cybernetics, IEEETransactions on, 35(4):646–658, Aug. 2005.

[31] Eng-Jon Ong and Richard Bowden. "A boosted classifier tree for hand shape deteetion." In Proceedings of the Sixth IEEE international conference on Automatic face and gesture recognition, . IEEE Computer Society FGR' 04, pages 889-894, Washington, DC, USA, 2004.

[32] Gaolin Fang, Wen Gao, and Debin Zhao. "Large vocabulary sign language recognition based on hierarchical decision trees."Proceedings of the 5th international conference on Multimodal interfaces - ICMI '03, page 125, 2003.

[33] Ming Hsuan Yang, Narendra Ahuja, and Mark Tabb. "Extraction of 2d motion trajectories and its application to hand gesture recognition." IEEE PAMI, 24(8):1061–1074, 2002.

[34] T Starner, J Weaver, and A Pentland. "Real-time American sign language recognition using desk and wearable computer based video." IEEE Transactions onPattern Analysis and Machine Intelligence, 20(12):1371-1375, 1998.

[35] Marco Maisto, Massimo Panella, Member, IEEE, Luca Liparulo, and Andrea Proietti "An Accurate Algorithm for the Identification of Fingertips Using an RGB-D Camera" IEEE Journal On Emerging And Selected Topics In Circuits And Systems, VOL. 3, NO. 2, JUNE 2013

[36] M. Van den Bergh and L. Van Gool. "Comhining RGB and ToF cameras for realtime 3D hand gesture interaction." In Applications of Computer Vision (WACV), 2011 IEEE Workshop on, pages 66-72, 2011.

[37] Antonis A Argyros and Manolis I A Lourakis. "Binocular hand tracking and reconstruction based on 2d shape matching", ICPR, pages 207-21O, 2006.

[38] H Nanda and K Fujimura , "Visual tracking llsing depth data". Conference on Computer Vision and Pattern Recognition Workshop, OO(C):37-37, 2004.

[39] Zhou Ren, Junsong Yuan, Jingjing Meng, and Zhengyou Zhang," Robust Part-Based Hand Gesture Recognition Using Kinect Sensor", IEEE Transactions On Multimedia, VOL. 15, NO. 5, AUGUST 2013

[40] Chao Sun, Tianzhu Zhang, Bing-Kun Bao, Changsheng Xu, and Tao Mei, "Discriminative Exemplar Coding for

Sign Language Recognition with Kinect" IEEE Transactions On Cybernetics 2013

[41] Cheoljong Yang, Yujeong Jang, Jounghoon Beh, David Han, and Hanseok Ko. "Gesture recognition using depth-based hand tracking for contactless controller application." In Consumer Electronics (ICCE), 2012 IEEE International Conference pages 297 -298, 2012.

[42] Jagdish L Raheja, Ankit Chaudhary, and Kunal Singal. "Tracking of fingertips and centers of palm using

KINECT". 2011 Third International Conference on Computational Intelligence Modelling Simulation, pages 248--252, 2011.

[43] Jos L. Hernndez-Rebollar, Robert V. Lindeman, and Nicholas Kyriakopoulos. "A multi-class pattern recognition system for practical finger spelling translation" In Proceedings of the 4th IEEE International Conference on Multimodal Interfaces,pages 185-190, 2002.