ANN based Indian Sign Language Numerals Recognition using the Leap Motion Controller

Deepali Naglot
Research Scholar
Department of CSE (IT),
Vishwakarma Institute of Technology,
Pune, India
dnaglot1711@gmail.com

Milind Kulkarni
Assistant Professor
Department of CSE (IT),
Vishwakarma Institute of Technology,
Pune, India
milind.kulkarni@vit.edu

***Abstract*—Sign language is the basic medium of communication for the hearing and speech impaired people. It has evolved as one of the major areas of research and study in computer vision. Researchers in sign language recognition used different input devices such as data gloves, web camera, depth camera, color camera, Microsoft's Kinect sensor, etc. to capture hand signs. In this paper we display the importance of Leap Motion Controller and proposed technique for classification and their efficient results. The proposed system contains three stages: Data collection, Feature extraction and Classification. The proposed technique for classification is executed on a dataset of total 200 samples (consisting of 20 samples of each ISL numbers). Indian Sign Language has a total of 26 alphabets and 10 numerals using either one hand or both hands to show the sign. Recognition rate of proposed system is 100%.**

Keywords- Indian Sign Language (ISL), Leap Motion Controller(LMC), Multi-Layer Perceptron (MLP).

I. INTRODUCTION

To build an application for the deaf and dumb community has become of utmost importance today. With the changing world in terms of technology innovations and education opening new arenas of opportunities, it becomes necessary for everyone to be ahead of the others in this race. In order to enable the deaf and dumb community to create recognition and also to give them a standard platform to communicate and express their opinions with every other individual, this application is being created [12].

Sign Language is basically a language which uses manual communication and body language to convey meaning. This can involve combining hand shapes, orientation and movement of hands. It can be considered to be a replacement of speech for hearing and visually impaired people. Thus, because of which it has attracted many researchers in this field since long. Sign language recognition systems are classified into two categories: Glove based systems and Vision based systems. Glove based systems requires user to wear a sensor glove or a colored glove to extract features describing the hand sign. Cyber glove is a device which is used for extracting the features such as orientation, movements and color of the hands. It is widely used for sign language recognition. Vision based systems use digital image processing and machine learning techniques to extract features and recognize sign. In vision based system user does not have to wear any special device or hardware [2].

Many researchers have been working on different sign languages like Arabic Sign Language (ArSL), Australian Sign Language, British Sign Language (BSL), Danish Sign Language French Sign Language, Taiwan Sign Language and others but less work has been carried on in Indian Sign Language. In this paper, the fundamental experimentation is done on Indian Sign Language. Indian Sign Language was developed so that the deaf people in the society can interact with the normal people without any difficulties. In this paper, we have considered the numerals of ISL which involves the use of single hand. ISL comprises of around 3000 gestures and other common words. This application initiates to focus on the importance to be given to ISL neglecting the limitations widely known [12].

The rest of the paper is organized as follows: Section II discusses about the literature survey in the area of sign language recognition using different sensors.

Introduction of Leap Motion sensor is given in section III. The proposed method is presented in detail in section IV .Section V presents experimental results and discusses about performance evaluation parameter. Section VI gives the conclusions of the work.

II. LITERATURE SURVEY

In [2], Adithya and Vinod presented a method of recognizing the Indian sign language. Images are captured using web camera under lighting and environmental condition. Captured images are then converted to YCbCr colour space. Distance transform and Discrete Fourier Transform (DFT) is used for extract features from image and artificial neural network used for recognizing Indian sign language alphabets. Black background and static images were taken in the experiments, achieving a recognition rate of 91.11%.

In [8],Joyeeta Singha introduced a method for recognizing Indian Sign Language alphabets with continuous video sequences of the signs. The preprocessing stage includes skin filtering, histogram matching. Eigen values and Eigen Vectors were calculated for feature extraction and Eigen value weighted Euclidean distance classifier was used to recognize the sign. The achieved recognition accuracy was 96.25% for 24 characters.

Ching-Hua Chuan and Eric Regina proposed a system for ASL recognition using leap motion controller .This system uses leap motion sensor to collect data from the user. They used k-nearest neighbour and SVM to classify the 26 alphabets of American Sign Language. Average classification rate of k-nearest neighbour and support vector machine was 72.78% and 79.83% respectively [4].

In [5], M. Mohandes and S. Aliyu used Leap motion controller to recognize Arabic sign language. This system were collected 10 samples of each 28 letter using leap motion sensor. Out of 23 features returned by Leap Motion Controller for each frame of data they selected 12 most relevant features for further process. They used Nave Bayes classifier and Multilayer Perceptron (MLP) to classify 28 letters in Arabic sign language. A correct recognition rate of 99.1% achieved using Multilayer Perceptron and 98.3% using Nave Bayes classifier.

An approach for the recognition of Arabic sign language is addressed in [6]. A.S.Elons and Menna Ahmed used Leap motion sensor. This system works with 50 different dynamic signs from Arabic sign language. The signs were collected from 4 different persons, two signs set used for training and other two for testing. The fingers position and distance between the fingers in each frame were features sets provided to system. Proposed system used Artificial Neural Network (ANN) as classifier to recognize the gestures. Recognition accuracy of 88% was achieved.

III. LEAP MOTION CONTROLLER

The Leap Motion controller is a small USB peripheral device which is designed to allow users to control their computers with hand gestures alone. This sensor is 3D non-contact motion sensor which can detects and tracks hands, fingers, bones and finger-like objects reporting discrete position and motion. The heart of the device consists of two monochromatic IR cameras and three infrared LEDs as shown in Fig. 2. These track infrared light with a wavelength of 850 nanometers, which is outside the visible light spectrum [14].The device has a large interaction space of eight cubic feet, which takes the shape of an inverted pyramid – the intersection of the binocular cameras' fields of view. The Leap Motion Controller's viewing range was approximately 1 inch to 2 feet (60 cm) above the device. With the Orion beta software, this has been expanded to 2.6 feet (80 cm).The LEDs generate pattern-less IR light and the cameras generate almost 300 frames per second of reflected data. This is then sent to the host computer through a USB cable, where it is analyzed by the Leap Motion software using "complex maths" in a way that has not been disclosed by the Leap Motion Company, in some way synthesizing 3D position data by comparing the 2D frames generated by the two cameras [13]. Fig. 2 shows internal structure of Leap Motion Controller.

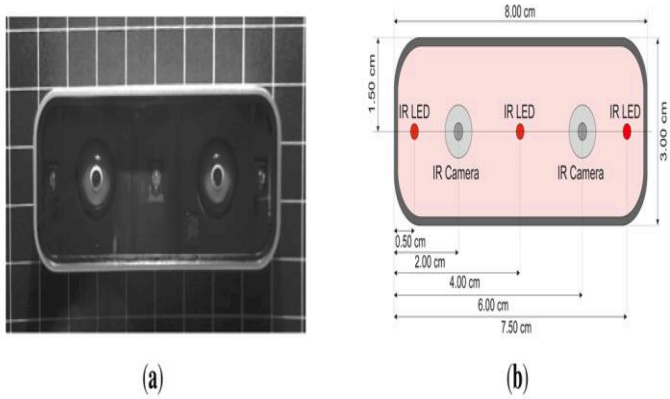

Fig. 1: Leap Motion Controller

The Leap Motion system employs a right-handed Cartesian coordinate system. The origin is centered at the top of the Leap Motion Controller. The x-axis lies in the horizontal plane and running parallel to the long edge of the device. The y-axis lies in vertical plane, with positive values increasing upwards. The z-axis has positive values increasing toward the user and it is lie in the horizontal plane [14].

IV. PROPOSED SYSTEM

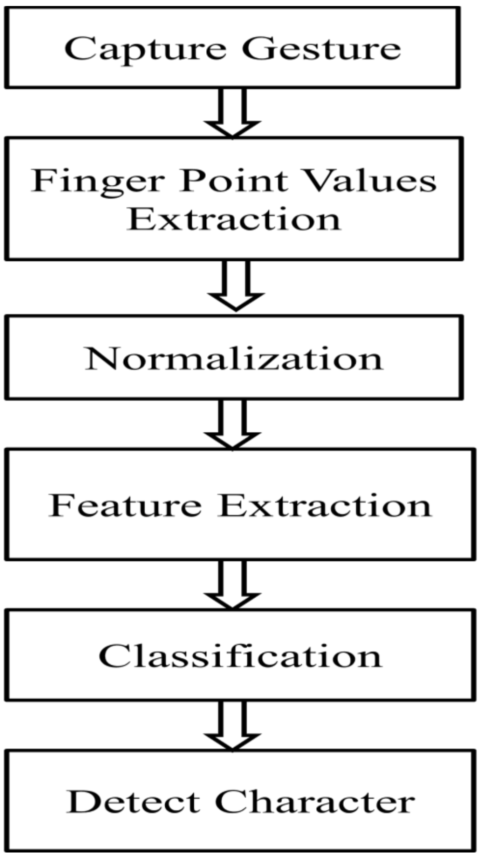

Fig. 2: Architecture of Proposed System

1) Data Collection

The proposed system will take 3D dynamic signs as input. In proposed system, data is collected using leap motion controller. Dataset of total 200 samples (consisting of 20 samples of each numerals) from which 100 used for training and remaining samples used for testing. Dataset is normalized before providing to ANN for training and testing.

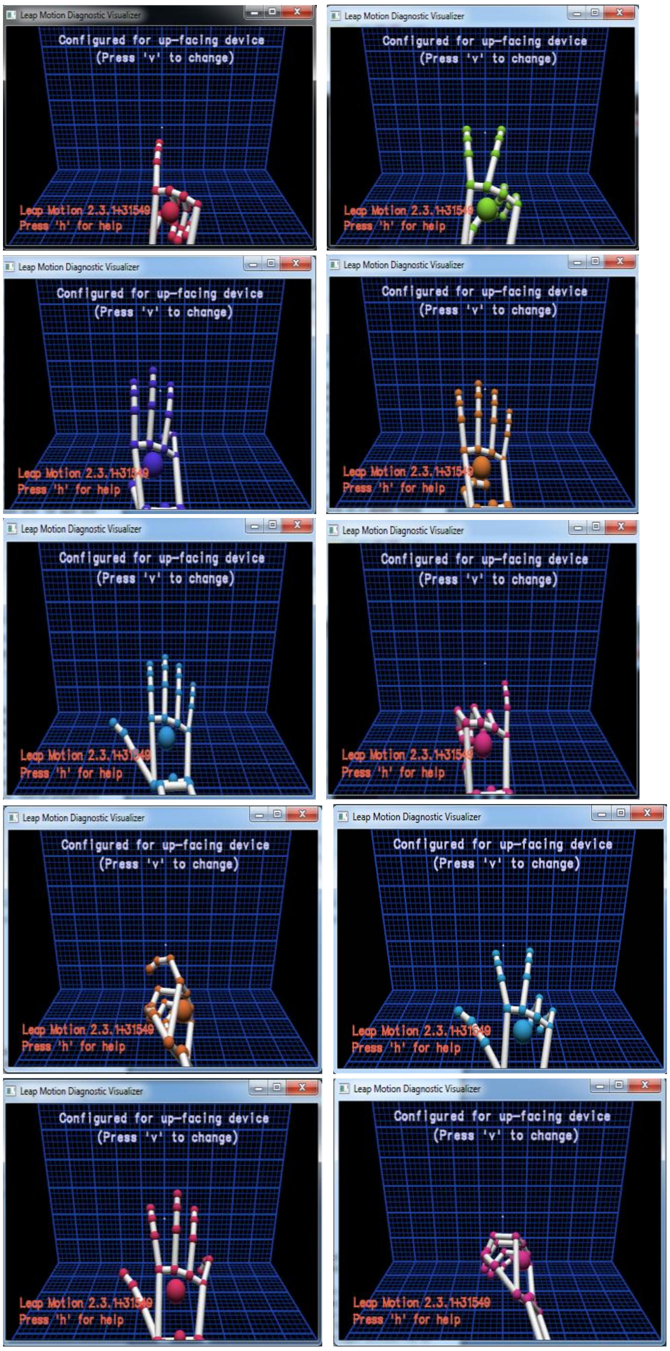

Fig.3: Some sample of 1, 2, 3, 4, 5, 6, 7, 8, 9, 0 digit signs of ISL on visualize tool of Leap Motion

2) Features Extraction

Leap motion API include different features for hand, fingers, bones and gesture. Some of them as follow:

Hand: The hand model provides information about the type (left, right or both hand), position (the center position of the palm in millimeters), velocity (in millimeter per second) and other characteristics of a detected hand, the arm to which the hand is attached, and lists of the fingers associated with the hand [15].

Fingers: Fingers related features include finger direction (a unit direction vector), finger length (in millimeter), width, tip position, tip velocity, dip position, pip position, mcp position.

From this detailed information, we select the Palm and finger dataset to be the features. Normalized features are provided to ANN. Features used for proposed system to train Artificial Neural Network (ANN) are:-

1) The Euclidian distances between the consecutive finger tip position to palm position.

2) The Euclidian distances between the finger tip position of each consecutive finger. In Euclidean three-space, the distance between points $x_1, y_1, z_1,$ and x_2, y_2, z_2 is calculated using following formula

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \quad (1)$$

3. Classification

1) *Multilayer Perceptron neural network (MLP)*

Multilayer Perceptron neural network (MLP) is self-learning, supervised algorithm used for pattern recognition and classification. In ANN neurons are interconnected which exchange messages between each other. The connections have numeric weights that can be tuned based on experience, making neural network adaptive to inputs and capable of learning. Backpropagation, an abbreviation for "backward propagation of errors", is a most commonly used algorithm for training Multilayer Perceptron neural network (MLP). The backpropagation algorithm trains a given feed-forward multilayer neural network for a given set of labelled training examples. The network examines its output response to the sample input pattern, when each entry of the sample set is presented to the network. The output response is then compared to the actual and expected results and the error value is calculated. The connection weights are adjusted based on the error. The MLP consists of three types of layers input, output and hidden layer with one or multi layer. The input layer neurons are connected to the hidden layer neurons and the hidden layer neurons are connected to output layer neurons by means of interconnection of weights as shown in fig.4. The size input layer is set by number of feature set we want the network to process. Similarly, the size of the output layer is set by the number of alphabets we want to recognize [2].

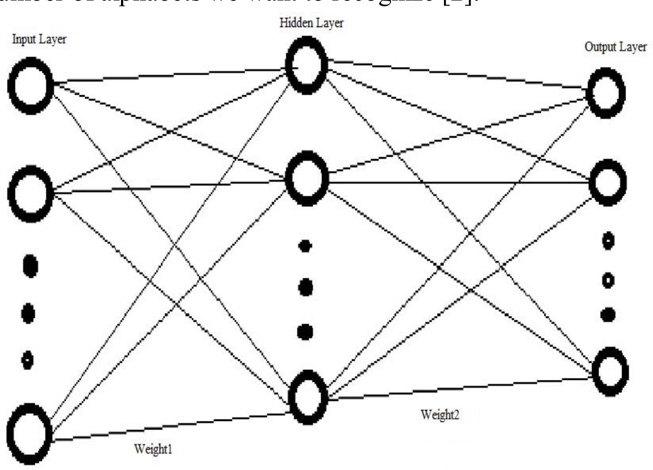

Fig. 4: Architecture of MLP neural network

Algorithm:

1. Initialization of weights between input layer to hidden layer units say vi , and hidden layer to output layer units say wj ,bias to some random values between -1 to 1.set value of learning rate .
2. Calculate net input to the each of the hidden layer unit

$$net_j = \sum_{i=0}^{d} Xi * Wji + b \qquad (2)$$

where xi is inputs given to the input layer, d is no of input unit.

3. Calculate output of each of hidden unit Yj by applying sigmoid activation function on net inputs from input layer units

$$Y_j = \frac{1}{1+e^{-net_j}} \qquad (3)$$

4. Apply output of hidden layer units to input of output layer units.
5. Calculate net input to each of the output layer unit

$$net_k = \sum_{i=0}^{h} Yi * Wkj + b \qquad (4)$$

where Yi is inputs given to the hidden layer, h is no of hidden unit.

6. Calculate output of each of the k output layer neurons Zk by applying sigmoid activation function to net input to the output layer unit.

$$Zk = \frac{1}{1+e^{-net_k}} \qquad (5)$$

7. Calculate error at output layer unit k

$$\delta_k = net_k(1 - net_k)(tk - Zk) \qquad (6)$$

where tk is target outputs

8. Back propagate error to each of hidden layer unit j using equation (7)

$$\delta_j = net_j(1 - net_j) \sum_k Wkj\ \delta_k \qquad (7)$$

9. Update weights of interconnection from hidden layer unit to output layer unit using equation 8,

$$\Delta\ Wkjnew_{new} = W_{old} + \Delta\ Wkj \qquad (8)$$

where, $\Delta\ Wkj$ is change in weight and is given by equation 9

$$\Delta\ Wkj = \alpha\ \delta_k Yj \qquad (9)$$

10. Update weights of interconnections from input layer neurons to hidden layer neurons using equation 10,

$$\Delta\ Wjinew = Wold + \Delta\ Wji \qquad (10)$$

$$\Delta\ Wji = \alpha\ \delta_j Xi \qquad (11)$$

V. RESULTS AND DISCUSSION

The proposed system was implemented using Intel(R) Core(TM) i5-4210U processor @ 2.40 GHZ speed and the code was written using Java programming language. Data is collected using 4 different persons through Leap motion sensor, total 200 (20 samples of each alphabet) input samples collected. Out of 200 input samples, 100 (10 samples of each numbers) samples used for training and 100 (10 samples of each numbers) samples used for testing. Output of the feature extraction step is feature set. It is used as input of the classifier to recognize the sign. Multi-Layer Perceptron (MLP) neural network with Back Propagation (BP) algorithm was used as the classifier and 10 ASL numerals were classified in the proposed system. Proposed system contains single input layer, hidden layer and output layer. The input layer has 9 neurons which is number of feature set (Euclidian distances between the consecutive finger tip position to palm center and Euclidian distances between the finger tip positions of each consecutive finger), 10 neurons in output layer which is number of numerals (classes). Number of neuron in hidden layer calculates using (input neuron + output neuron) /2. Sigmoid activation function is used in MLP neural network.

Table 1: Confusion Matrix for 2 class classification

	Predicted Class			
	Class	As Positive	As Negative	
Actual Class	Positive	TP	FN	P
	Negative	FP	TN	N
	Total	P’	N’	P+N

One of the approaches to evaluate performance of the supervised learning algorithms like MLP with BP is confusion matrix sometimes called as contingency table. Confusion matrix is analyzing how well our classifier can recognize samples of different class. A confusion matrix of size k x k associated with a classifier, demonstrates the predicted and actual/target classification, where k is the number of different classes. Table 1.presents confusion matrix, where TP contains true positive values (no of samples correctly classified), FN contains false negative (no of samples wrongly classified) values, FP contains false positive values and TN contains true negative values.

Accuracy is the percentage of test samples that are correctly recognized by the classifier. Precision is measure of exactness. Recall is measure of completeness.

$$\text{Accuracy}=\frac{\text{TP+TN}}{\text{P+N}} \qquad (12)$$

$$\text{Precision}=\frac{\text{TP}}{\text{TP+FP}} \qquad (13)$$

$$\text{Recall}=\frac{\text{TP}}{\text{TP+FN}} \qquad (14)$$

Table 2. shows confusion matrix for proposed system. Recognition rate of proposed system is 100% as there is no similarity between signs.

Table 2: Confusion matrix of proposed system

	Predicted Class															
	Sign	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Total**	**TP**	**FN**	**FP**	**Precision**	**Recall**
Actual Class	**0**	**10**	0	0	0	0	0	0	0	0	0	10	10	0	0	100%	100%
	1	0	**10**	0	0	0	0	0	0	0	0	10	10	0	0	100%	100%
	2	0	0	**10**	0	0	0	0	0	0	0	10	10	0	0	100%	100%
	3	0	0	0	**10**	0	0	0	0	0	0	10	10	0	0	100%	100%
	4	0	0	0	0	**10**	0	0	0	0	0	10	10	0	0	100%	100%
	5	0	0	0	0	0	**10**	0	0	0	0	10	10	0	0	100%	100%
	6	0	0	0	0	0	0	**10**	0	0	0	10	10	0	0	100%	100%
	7	0	0	0	0	0	0	0	**10**	0	0	10	10	0	0	100%	100%
	8	0	0	1	0	0	0	0	0	**10**	0	10	10	0	0	100%	100%
	9	0	0	0	0	0	0	0	0	0	**10**	10	10	0	0	100%	100%
	Total	10	10	10	10	10	10	10	10	10	10	**260**	**260**	**0**	**0**	**100%**	**100%**
Recognition Rate=100%																

VI. CONCLUSION

Sign Language Recognition is need for the hearing and speech impaired people to communicate with other people. This paper presents method for Indian Sign Language recognition using newly introduced Leap Motion Controller. Leap Motion Controller is 3D non-contact motion sensor which can detect hands and report discrete position and motion. Some of the advantages of Leap Motion Controller are: Robustness, Require less memory, Fast Processing. Proposed system does not require any specific background and environmental condition.

Features provided to ANN to train the system are distance between fingers tip position to palm center and distance between consecutive fingers tip position. Multi-Layer Perceptron (MLP) neural network with Back Propagation (BP) algorithm is used to recognize 0-9 digits of Indian Sign Language. Recognition rate of proposed system is 100%. The performance of the proposed system is improved in terms of accuracy and time to build the model using above approach. ISL alphabets use one or both hand to show signs. Recognition of both hand gestures and sentence or words will be implemented in the future.

VI. REFERENCES

[1] Cao Dong, M. C. Leu and Z. Yin, "American Sign Language alphabet recognition using Microsoft Kinect," *2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Boston, MA, 2015, pp. 44-52.

[2] Adithya, V.; Vinod, P.R.; Gopalakrishnan, U., "Artificial neural network based method for Indian sign language recognition," Information & Communication Technologies (ICT), 2013 IEEE Conference on , vol., no., pp.1080,1085,2011-12 April 2013.

[3] K. Y. Fok, N. Ganganath, C. T. Cheng and C. K. Tse, "A Real-Time ASL Recognition System Using Leap Motion Sensors" *Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), 2015 International Conference on*, Xi'an, 2015, pp. 411-414.

[4] C. H. Chuan, E. Regina and C. Guardino, "American Sign Language Recognition Using Leap Motion Sensor," *Machine Learning and Applications (ICMLA), 2014 13th International Conference on*, Detroit, MI, 2014, pp. 541-544.

[5] M. Mohandes, S. Aliyu and M. Deriche, "Arabic sign language recognition using the leap motion controller," *2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE)*, Istanbul, 2014, pp. 960-965.

[6] A. S. Elons, M. Ahmed, H. Shedid and M. F. Tolba, "Arabic sign language recognition using leap motion sensor," *Computer Engineering & Systems (ICCES), 2014 9th International Conference on*, Cairo, 2014, pp. 368-373.

[7] Gweth, Y.L.; Plahl, C.; Ney, H., "Enhanced continuous sign language recognition using PCA and neural network features," Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on , vol., no., pp.55,60, 16-21 June 2012.

[8] Joyeeta Singha, Karen Des "Recognition of Indian Sign Language in Live Video"International Journal of Computer Application,Volume 70-No 19,May 2013

[9] Samrudhi Naik "Depth Sensing Cameras- A Comparative Study, "International Engineering Research Journal (IERJ), Volume 1 Issue 11 Page 1585-1590 ISSN 2395-1621

[10] P. Mekala, Y. Gao, J. Fan and A. Davari, "Real-time sign language recognition based on neural network architecture," *2011 IEEE 43rd Southeastern Symposium on System Theory*, Auburn, AL, 2011, pp. 195-199.

[11] Dr.Sudeep Thepade, Arati Narkhede, Priti Kelvekar " Novel Technique for Background Removal from Sign Images for Sign Language Recognition System" International Journal of Computer Applications, (0975 – 8887) Volume 78 – No.10, September 2013.

[12] Dr. Sudeep.D.Thepade Ms.Arati Narkhede, Ms.Priti Kelvekar, Ms.Gandhali Kulkarni and Ms.Seema Tathe, " Sign Language Recognition using Color Means of Gradient Slope Magnitude Edge Images" published at International Conference on Intelligent Systems and Signal Processing (ISSP2013), 978-81-909376-6-5 2013IEEE.

[13] A. Colgan, "How Does the Leap Motion Controller Work?" (Accessed: 2014-09-05). [Online]. Available: http://blog.leapmotion.com/hardwareto-software-how-does-the-leap-motion-controller-workLeap Motion

[14] http://www.leapmotion.com (last access: August 2014)