

STUDY OF HAND GESTURE RECOGNITION AND CLASSIFICATION

SRINIVAS K, MANOJ KUMAR RAJAGOPAL*

School of Electronics Engineering, VIT University, Chennai – 600 127, Tamil Nadu, India. Email: manojkumar.r@vit.ac.in

Received: 23 January 2017, Revised and Accepted: 03 March 2017

ABSTRACT

To recognize different hand gestures and achieve efficient classification to understand static and dynamic hand movements used for communications. Static and dynamic hand movements are first captured using gesture recognition devices including Kinect device, hand movement sensors, connecting electrodes, and accelerometers. These gestures are processed using hand gesture recognition algorithms such as multivariate fuzzy decision tree, hidden Markov models (HMM), dynamic time warping framework, latent regression forest, support vector machine, and surface electromyogram. Hand movements made by both single and double hands are captured by gesture capture devices with proper illumination conditions. These captured gestures are processed for occlusions and fingers close interactions for identification of right gesture and to classify the gesture and ignore the intermittent gestures. Real-time hand gestures recognition needs robust algorithms like HMM to detect only the intended gesture. Classified gestures are then compared for the effectiveness with training and tested standard datasets like sign language alphabets and KTH datasets. Hand gesture recognition plays a very important role in some of the applications such as sign language recognition, robotics, television control, rehabilitation, and music orchestration.

Keywords: Human-computer interaction, Hand gestures, Gesture recognition system, Static and dynamic hand gesture.

© 2017 The Authors. Published by Innovare Academic Sciences Pvt Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>) DOI: <http://dx.doi.org/10.22159/ajpcr.2017.v10s1.19540>

INTRODUCTION

Body language is one the non-verbal forms of communication and hand gestures and arm movement, posturing, face movement and eye gaze forms the non-verbal form of body language communication [1].

Gestures are considered as one of the methods used for conveying information through motion of the body. In hand gesture recognition system (GRS), hand gestures can be a continuous gesture referred as dynamic gestures or can be fixed or single gesture referred as static gestures [2].

Hand gesture plays a vital role in the day-to-day interaction devices such as smart homes, smartphones, and other gadgets; these devices use hands for communication, interfacing, and networking with the environment. Gesture recognition is considered as a multidisciplinary approach combining computer vision, pattern recognition, motion analysis, and machine learning [3].

Commonly used hand recognition techniques are (a) vision-based recognition and (b) non-vision-based recognition method. In vision-based approach, algorithms are used to derive hand pose and trajectory information and environmental factors such as background illumination, hands occlusion, and skin color plays a significant role during the recognition process. In non-vision-based scheme, users' needs to wear data gloves and handle the cabling part of recognizing system [2].

GRS is divided into: (a) extraction, (b) estimation and classification, (c) recognition [4].

Segmentation is the primary process in extraction; in this process, input data are partitioned into demarcated sections. Skin color is used as one of the common inputs for segmentation process. Various parametric tools exist for segmentation process like Gaussian model, Gaussian mixture model [4].

Hand shape and features are estimated using either hand contour or fingertip positions or palm center; also algorithms like self-growing and self-organized neural gas can be used for features estimation [4].

With segmented and feature estimated inputs, classification algorithms are used for gesture recognition, and algorithms such as hidden Markov model (HMM), finite state machine, principal component analysis are the widely used to recognize the gestures [4].

In this paper, Section II provides an introduction to gestures, and Section III has details of commonly used gesture capture devices for capturing hand gesture movement. Section IV highlights the commonly used gesture recognition algorithms and Section V elaborates on standard datasets to be used for training and experimentation. Usage of hand GRS in different applications is addressed in Section VI.

GESTURES

Hand gestures used for communication and conveying emotions can be acquired from images, video sequences, and datasets. Gestures may be 2D gestures (plane) or 3D gestures (free-form) or static gestures or dynamic gestures. Static gestures are configuration based on hand position, and dynamic gestures are motion-based hands movement [3].

Badi *et al.* [5] state static hand gesture or hand posture as "posture is a combination of hand position, orientation and flexion observed at some time instance." Static hand gestures do not vary with respect to time signals. Static gestures can be evaluated using minimal set of images (single image or group of images captured at a specific time). Facial information is a good example for static gesture and also hand gestures such as "OK" or "STOP" signs provide complete understanding of the sign used for communication. Dynamic hand gestures are defined as continuous sequences of hand postures connected by motions/movements. In a continuous video sequence each frame defines a posture, and entire video sequence defines gesture (single action gesture and multiple gestures). Examples for dynamic gestures includes "goodbye," "come here" which is a combination of multiple gestures.

GESTURE CAPTURE DEVICES

Kinect

Tashev [6] details about Kinect, device specifically designed for human-computer interaction (HCI) which includes features

supporting Gestures and speech in addition to skeletal and face tracking. Furthermore, Kinect provides voice identification, and by fusion techniques, 3D data can be constructed. Kinect device made its first appearance in 2010 and the public version for Windows kernel development kit (KDK) was available in 2011 and 2013 KDK 1.7 version was released. Kinect sensor device has mainly four components: Element microphone array (4), color camera, depth camera, and tilting mechanism. Depth camera works on unstructured scattered light principle in which image is converted to depth image, and pixel value is the gap to object. Color camera is the traditional web camera. Microphone array consists of four supercardioid microphones operating at 16 kHz sampling rate, and camera position adjustments are done by tilting mechanisms.

Bhattacharya *et al.* [7] present approaches for gesture segmentation and classification by applying machine learning on Kinect's data stream. Kinect device is capable of identifying human skeleton from the skeletal joints position of up to two people. Kinect processes skeletal data and delivers at 30 frames/second. With increasing distance from sensors, depth measurements increases and depth resolution decreases, and the recommended range is between 1 and 3 m.

Ibañez *et al.* [8] propose uses of easy Gesture recognition tool along with Kinect for gesture recognition. This tool helps developers to overcome the need for machine learning algorithms, helping amateur users to archive gestures by building newer training datasets. These training datasets can be used for further machine learning techniques.

Chen *et al.* [9] use the concept of fusion of Kinect along with hand-held sensors for gesture identification. Input data are acquired from Kinect sensor and wireless sensors are secured to the user's wrist, and the subject operates within Kinect specified operating range, <3 m. By this blending method, relatively better results are achieved than operated individually.

Sensors

Berman and Stern [10] describes sensors as equipment or group of gadgets which can be used as detectors for grasping gesture signals. Two of the most important platform characteristics of sensors are to check if the user needs to be a large number of user attached sensors or to be unencumbered and also does the system needs single sensor or numerous sensors. These multiple sensor platforms can be alike or dissimilar. Sensors that offer robustness and are flexible exhibit challenges in complexity and computational overhead. Some of the predominant features of sensors include resolution, pace of acknowledgment, range and dimensions, cost of sensors and its operating conditions and environment. Optical or light-based sensors and acoustics or sound-based sensor systems are the prime examples for unencumbered recognition setups. In encumbered sensor setups, devices like data gloves, markers, and infrared detectors are attached to the subjects making gesture capture a simpler process. Users must handheld or wear relevant devices while using magnetic sensors, accelerometer (ACC) and Gyro-type recognition systems and in electric-based configurations, users must wear gloves or should physically touch the screen. Data glove devices are heterogeneous sensors devised mainly for hand motion recognition, and these sensors provide details of hand shape, fingers position and movement and different angles of palm. Hand glove-based sensors are extensively used in robotics and sign language communications. Accuracy and response speed of these glove systems are high with minimal interventions from hand surrounding environment but has challenges like higher computing processing systems, user's hand size and shape and user-specific calibration processes.

Cetin and Erden [11] uses a camera along with sensors to achieve higher recognition rates. In this configuration setup, differential pyroelectric infrared (PIR) sensor array is used along with an ordinary camera. As PIR sensors recognize all higher temperature bodies in the view range, it does not differentiate hand motions from other face, and body

movements and hence camera is mainly used for hand identification. Any type of mobility will be sensed by any one of the sensors and camera will check if the hand move is causing the locomotion then the data is processed in real-time.

Kurita [12] presents a non-contact scheme for measuring human hand kinetics using portable wireless motion-detection sensors. In this method, the current produced due to the capacitance difference between subject's hand and electrodes is calculated. Using electrostatic induction current, the sensors will be locating the direction of hand movement.

Electrodes

Ouyang *et al.* [13] evaluate the method of non-linear features extraction and classification to categorize different hand manipulations using (surface electromyogram [sEMG]) surface electromyography signals. The non-linear hand features are measured referring from the recurrence plot which is used for representing dynamical attributes of sEMG signals during hand locomotion. Positions of electrodes to measure sEMG are defined by the musculoskeletal systems of the forearms and substantiated by muscle specific contractions like hand expansion and seizure.

ACC

Li *et al.* [14] used glove-based sensing technique for acquiring the rotary movement of hand and fingers. These sensing devices were fitted with bending sensors and ACC. ACC is capable of calculating dynamic accelerations like vibrations and static accelerations like gravity. Control applications-based consumer devices use extensively ACC for day-to-day operations.

3D-ACC in hand maneuver recognition system calculates velocity change rate along three axes (x, y, and z). In these systems, effective segments are scoped and hypothesized as array of feature vector. Using linear min-max scaling method, amplitude of operative segments is scaled. Using normalization techniques, range, and momentum are improved to identify the gestures.

GESTURE RECOGNITION ALGORITHMS

Multivariate fuzzy decision tree (MFDT)

Bien *et al.* [15] propose MFDT to learn and classify hand gestures based on decision tree learning method. In fuzzy decision tree method of cataloguing hand gestures, count of nodes is very large as the data is split using fuzzy membership; having large number of nodes causes lower performance rates. MFDT has smaller number of nodes as compared to fuzzy decision trees. In decision tree learning method, discrete-valued target functions are approximated and learned functions are symbolized as a decision tree. ID3 and C4.5 are the well-known decision tree algorithms. As compared to fuzzy decision tree, in MFDT, decision tree is formed using multivariate concept, but fuzzy decision tree uses univariate split concept.

HMM

Richarz and Fink [16] state that HMM is one of the most widely accepted tool for time-series analysis as this method is capable of exhibiting temporal relationships among different models and samples in addition to segmentation and classification. This capability is extensively used for training and inference by a different variation of HMM.

Yang *et al.* [2] define that, HMM is considered as one of the most popular approaches for dynamic hand gesture recognition. In HMM, hidden factors ensure different states and variants from these state transitions are visible. The output of HMM also consists of hidden information of state sequence, and though each state is not noticeable, potential output has possibility distribution.

Dynamic time warping (DTW) framework

DTW framework is one of the temporal classifier method, addresses the issue of motion velocity variation, and this method uses matching

schemes technique. This algorithm searches for best warping among sequences of gestures movement using distance matrices. The sequences can be of different dimensions or referenced misalignment. A robust cost function is used by DTW to carry out signal positions. DTW can also be used with other algorithms including HMM for improving the overall performance [17].

Latent regression forest (LRF)

Chang *et al.* [18] present LRF algorithm for estimating 3D hand posturing from a single depth image. This algorithm is targeted mainly for ordered coarse-to-fine search specific to depth images. In this framework, binary decision trees are learned with the process of repeatedly partitioning the input image into subregions, these subregions will correlate to specific skeletal joint. LRF can be considered as an aggregation of binary decision trees, trained on bootstrap sample. Three types of nodes form the tree; types of nodes are split nodes, division nodes, and leaf nodes. Split nodes determine the route, division nodes splits the search object into disjoint objects for input data proliferation and leaf nodes are the final terminating points that enact as skeletal joints.

Most discriminating segment longest common subsequence (MDSLCS)

Stern *et al.* [19] address the LCS method, a predecessor of DTW. Due to stable dimensional data strings, LCS has an advantage over Euclidean and Manhattan; in addition, LCS similarity measure is sturdy for noises. In LCS, noisy components of path movement will not be compared. Authors propose to use the classification algorithm most discriminating subsegments (MDSs) based on LCS algorithm; hence, named as MDSLCS. The key idea of MDSLCS is the automatic identification and derivation of MDSs which makes it a better classifier than extracting full gestures. Representing each gesture as MDSs is analogous to phonemes in speech or strokes in handwriting. A recognition result of 92.6% was achieved as compared to 89.5% recognition using HMM.

Support vector machine (SVM)

Glette *et al.* [20] use SVM gesture recognition technique; SVM is based on the principle of structural risk minimization which is an advantage leading to good generalization performance. SVM finds a solution using a nonlinear transformation referred as feature space in which an optimal hyperplane is determined. Hyperplanes with maximal margins are referred as optimal; margins are the minimal distance from separating hyperplanes to closest data points which is called as Support vectors.

sEMG

Naik *et al.* [21] define sEMG as an electrical recording of muscle actions and movements. sEMG is linked to muscle contraction strength and can be used only for control of prostheses as these sensors are not suitable for people with physical amputations.

Micera *et al.* [22], define sEMG as a mode for measurement of Electrical signals using electrodes positioned over the skin. Capturing sEMG signal is impacted by physiological, anatomical, and biochemical fiber type structure of muscles, depth, and location of active fibers. sEMG signals are also influenced by electrodes dimensions, materials made-up of, gaps between surfaces, placing of electrodes on muscle surfaces. sEMG signals are learned through different conjurations and methods like monopolar in which only one electrode is placed on muscle surface for recording and capturing sEMG signals.

DATASETS USED FOR EXPERIMENTATION AND VERIFICATION

Sign language

Kalsh *et al.* [23] specify that one of the popular standard datasets for hand movement recognitions are hand sign languages.

Badi *et al.* [5] express human sign language as one the convenient and mandate method for communicating to hearing challenged person and deaf people. Sign language is one of the visual languages communicated using lip signs, facial utterances, hands and arm actions. Sign language understanding and identification process involve in precisely translating the sign actions and movements to speech/voice or textual messages. Sign languages are not global but instead are region specific giving rise to different forms such as American sign language, Indian sign language, Chinese sign language, German sign language and so on.

Sun *et al.* [24] specify that all the 26 alphabets in English vocabulary can be unique hand gestures. Recognition processes involve extracting the meaningful features from these gestures and interpret the specific meanings and finally generate tone or text messages.

Fig. 1 shows the mapping of alphabets and numbers to unique hand gestures in American sign language [25].

Color tip dataset

ColorTip dataset is one the public datasets used for hand gesture recognition and fingertip identification. These datasets consist of set of

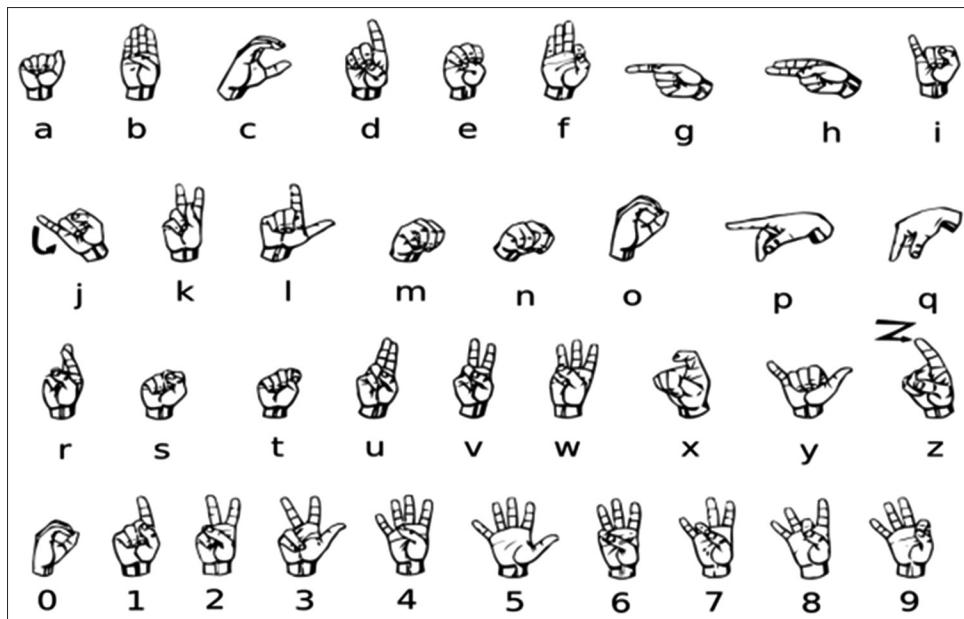


Fig. 1: American sign language alphabets

recordings and annotations captured from Kinect sensor camera. In this dataset, gestures are recorded from subjects wearing colored gloves as shown in Fig. 2. Each gesture varies in orientation and translation with complexity and challenging factor depending on the amount of intra-gesture variability [26].

EMG

sEMG signals are captured from human forearms by placing electrodes on the subjects as shown in Fig. 3. These signals can be used to manage real prosthesis or pragmatic devices.

EMG-controlled devices rely on the identification and analysis of forearm muscle contractions. Patterns derived from EMG activities specify each hand movement. Gesture recognition involves obtaining EMG signals using sensors, features extraction followed by features identification [27].

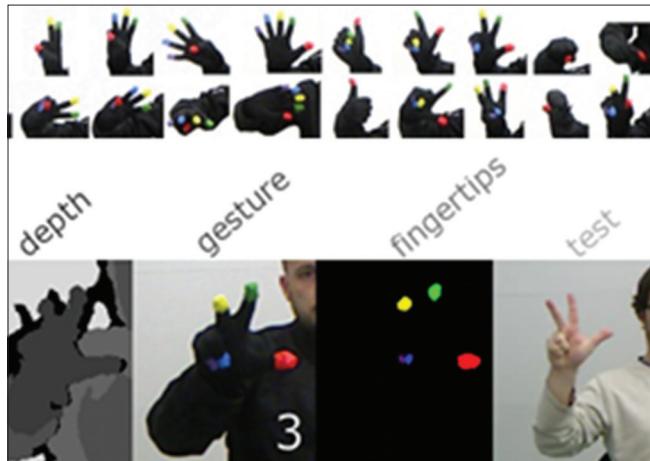


Fig. 2: Color tip dataset

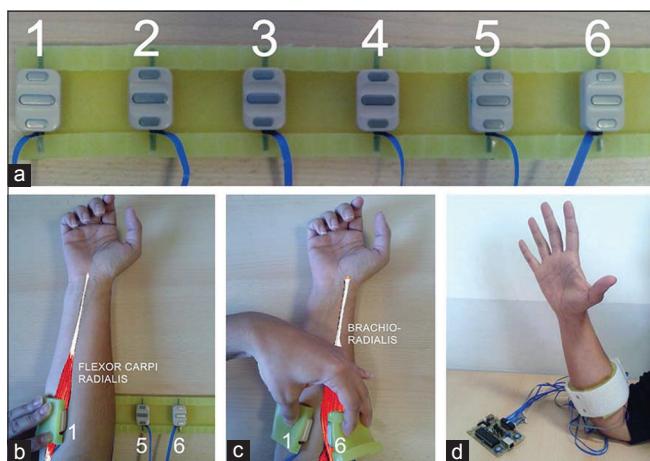


Fig. 3: Electromyogram sensors and bracelet (a) sensors with bracelet (b) sensor one position (c) sensor six position (d) bracelet position on forearm

KTH dataset

KTH dataset is an action collection dataset with six actions: Boxing, hand clapping, hands waving, jogging, running, and walking. These six different actions as in Fig. 4 are executed by 25 subjects in four different environments such as indoors, outdoors, outdoors with scale variation and outdoors with different clothing [28].

Weizmann action dataset

Weizmann action dataset is a collection of video tracks at a resolution of 180×144 pixels. These action dataset consists of nine actors performing ten different actions including walk, wave, bend, jump, and run as shown in Fig. 5 [29,30].

APPLICATIONS OF GRS

Robotics

Yan *et al.* [31] specify the usage of gesture recognition in managing mobile robots. Some of the social robots are being administered by unskilled users and hence there is a need for simple and powerful interfaces between human and robots for effective communication. Interfacing through LAN to instruct the robots is one of the methods of interacting. Using sensors hand gesture movements are extracted as real-time data and are streamed as feature vectors. Predefined commands such as robot movement in different directions, change of momentum, stop and move were trained from users. Gesture recognition methods provided a simple, flexible and accurate system to control the robots. Users should exhibit all the patterns few times beforehand to train and then robots memorize these patterns and recognize specified gestures.

Manganelli *et al.* [32] explore growing neural gas (GNG) algorithm as recognition scheme for robot movements and responses. In this technique, initial training is not needed instead motion captions are fetched from sequential skeletal depth user's data. This data are clustered by GNG and profiled to robotic response through reinforcement learning. Kinect sensor camera is used for user data capture.

Television control

Hu *et al.* [33] focus on automatic TV user's posture recognition system to identify and trigger the camera-based GRS to control the TV. By this method, energy usage is reduced as more power is utilized to keep the camera device and gesture identification modules always turned-on so that maintenance expenditure is high. Sensors perceive user's actions and face-based assessment to identify and interpret user's behavior into finite state method. Some of the user's actions studied include absent, other-action, watching and controlling.

Rehabilitation

Srinivasan *et al.* [34] focus on human movement's classification and identification in health-care applications with emphasis on rehabilitation. In telerehabilitation, gesture recognition plays a vital role due to a shortage of therapists and workspaces. Using the ultrasonic clustered Doppler sensing methods, movements are identified from the trained datasets. This method determines the momentum and angular information from doppler frequencies and direction of arrival (DoA) from the signals at the receiver sensors. Doppler frequencies collect DoA using Root-Music algorithm, and using Bayesian classifier method, movements are determined.

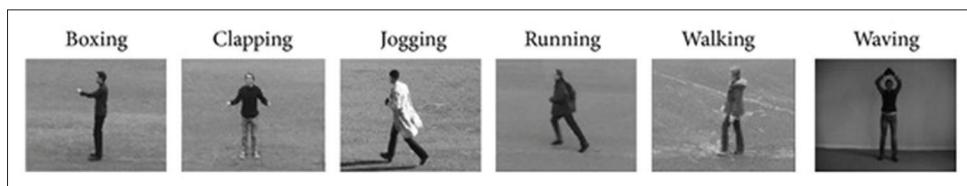


Fig. 4: KTH dataset

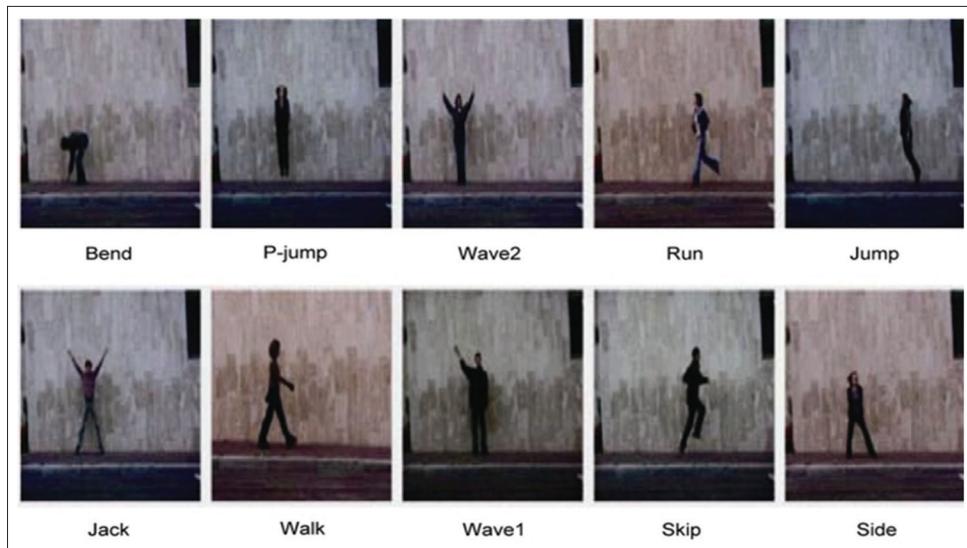


Fig. 5: Weizmann action dataset

Gesture for conducting music

Jung *et al.* [17] focus on the tool to assist the learning of music gestures for amateurs known as meter-mimicking. In musical concerts, each maestro uses unique different sets of gestures and also should ensure that these gestures are to be inferred by both amateurs and also expert musicians. Learning tool is based on the programmed detection of musical metrics by study of hand gestures. The musical patterns are recognized by DTW technique.

CONCLUSION

HCI is one of the advanced techniques for direct interfacing with computers as compared to keyboard and mouse. Hand gesture recognition along with other recognition techniques including face recognition, torso movement and eye gaze are the predominant HCI techniques. Hand gestures are communicated through dynamic movement like hand waving or through static poses like victory sign. GRS constitutes of different stages made up of gestures, gesture capture devices, tracking algorithm, feature extraction, classification algorithm. Kinect device and sensors are the well-known gesture capture device; DTW, HMM are the popular gesture identification algorithms. Standard datasets like Sign language, sEMG datasets are available for comparative studies.

REFERENCES

1. Yao Y, Fu Y. Contour model-based hand-gesture recognition using the Kinect sensor. *IEEE Trans Circuits Syst Video Technol* 2014;24(11):1935-44.
2. Yang S, Premaratne P, Vial P. Hand Gesture Recognition: An Overview. *Proceedings of IEEE IC-BNMT*; 2013.
3. Çağrılıboz N, Güdükbay U. A hand gesture recognition technique for human-computer interaction. *J Vis Commun Image Represent* 2015;28:97-104.
4. Khan RZ, Ibraheem NA. Hand gesture recognition: A literature review. *Int J Artif Intell Appl IJAIA* 2012;3(4):161-74.
5. Badi H, Hussein SH, Kareem SA. Feature extraction and ML techniques for static gesture recognition. *Neural Comput Appl* 2014;25(3-4):733-41.
6. Tashev I. Kinect Development Kit: A Toolkit for Gesture and Speech-Based Human-Machine Interaction. *IEEE Signal Processing Magazine*. University of Washington, September; 2013.
7. Bhattacharya S, Czejdo B, Perez N. Gesture Classification with Machine Learning using Kinect Sensor Data. *Third International Conference on Emerging Applications of Information Technology (EAIT)*. Fayetteville State University. 2012.
8. Ibáñez R, Soria A, Teyseyre A, Campo M. Easy gesture recognition for Kinect. *Adv Eng Softw* 2014;76:171-80.
9. Chen C, Jafari R, Kehtarnavaz N, Liu K. Fusion of Inertial and Depth Sensor Data for Robust Hand Gesture Recognition. Vol. 14. No. 6. *IEEE Sensors Journal*. University of Texas at Dallas, June. 2014.
10. Berman S, Stern H. Sensors for gesture recognition systems. *IEEE transactions on systems, man, and cybernetics-part C: Appl Rev* 2012;42(3):277-90.
11. Cetin A, Erden F. Hand gesture based remote control system using infrared sensors and a camera. *IEEE Transactions on Consumer Electronics*. Vol. 60. No. 4. Ankara, Turkey: Hacettepe University (Beytepe campus), November; 2014.
12. Kurita K. Noncontact hand motion classification technique for application to human-machine interfaces. *IEEE Transactions on Industry Applications*. Vol. 50. No. 3. Japan: Kinki University, May/June; 2014.
13. Ouyang G, Wilamowska-Korsak M, Liu H, Ju Z. Surface EMG based hand manipulation identification via nonlinear feature extraction and classification. *IEEE Sensors Journal*. Vol. 13. No. 9. Portsmouth: University of Portsmouth, September; 2013.
14. Li Y, Lantz V, Wang K, Yang J, Zhang X, Chen X. 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.
15. Bien Z, Jeon MJ, Lee SW. User adaptive hand gesture recognition using multivariate fuzzy decision tree and fuzzy garbage model. *Int J Fuzzy Syst Appl* 2011;1(3):15-31.
16. Richarz J, Fink GA. Visual recognition of 3D emblematic gestures in an HMM framework. *J Ambient Intell Smart Environ* 2011;3(3):193-211.
17. Jung CR, Miranda ER, Schramm R. Dynamic Time Warping for Music Conducting Gestures Evaluation. *IEEE Transactions on Multimedia*. Vol. 17. No. 2. Porto Alegre, Brazil: University of Federal do Rio Grande do Sul, February; 2015.
18. Chang HJ, Tejani A, Kim TK, Tang D. Latent regression forest: Structured estimation of 3D articulated hand posture. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Columbus, OH, London: Imperial College; 2014.
19. Stern H, Shmueli M, Berman S. Most discriminating segment - Longest common subsequence (MDSLCS) algorithm for dynamic hand gesture classification. *Pattern Recognit Lett* 2013;34(15):1980-89.
20. Glette K, Gruber T, Platzner M, Kaufmann P. Classification of electromyographic signals: Comparing evolvable hardware to conventional classifiers. *IEEE Transactions on Evolutionary Computation*. Vol. 17. No. 1. Paderborn: University of Paderborn, February; 2013.
21. Naik GR, Kumar DK, Jayadeva. Twin SVM for gesture classification using the surface electromyogram. *Inf Technol Biomed IEEE Trans* 2010;14(2):301-8.
22. Micera S, Carpaneto J, Raspopovic S. Control of hand prostheses using peripheral information. *Biomed Eng IEEE Rev* 2010;3:48-68.
23. Kalsh EA, Garewal NS. Sign language recognition system. *Int J Comput Eng Res* 2013;3(6):15-21.

24. Sun C, Zhang T, Bao BK, Xu C, Mei T. Discriminative exemplar coding for sign language recognition with Kinect. *IEEE Trans Cybern* 2013;43(5):1418-28.
25. Available from: <http://www.sinkind.com/hand-signs-letters.html>.
26. Suaua X, Alcoverroa M, López-Méndezb A, Hidalgoa JR, Josep R. Real-time fingertip localization conditioned on hand gesture classification. *Image Vis Comput* 2014;32(8):522-32.
27. Riillo F, Gruppioni E, Saggio G. Evaluating the influence of subject-related variables on EMG-based hand gesture classification. *Medical Measurements and Applications (MeMeA)*, 2014. IEEE International Symposium on IEEE; 2014.
28. Qiao Y, Tang X, Wang L. Latent hierarchical model of temporal structure for complex activity classification. *IEEE Tran Image Proc* 2014;23(2):810-822.
29. Athitsos V, Jangyodsuk P, Escalante HJ, Wan J. CSMMI: Class-specific maximization of mutual information for action and gesture recognition. *IEEE Trans Image Proc* 2014;23(7):3152-65.
30. Available from: <http://www.hindawi.com/journals/isrn/2012/872131/> fig8.
31. Yan R, Tee KP, Chua Y, Li H, Tang H. Gesture recognition based on localist attractor networks with application to robot control. *IEEE Computational Intelligence Magazine*. Singapore: Institute for Infocomm Research, February; 2012.
32. Manganelli J, Merino J, Threatt AL, Yanik PM. A gesture learning interface for simulated robot path shaping with a human teacher. *IEEE Trans Hum Mach Syst* 2014;44(1):41-54.
33. Hu W, Wang K, Lian S. Automatic user state recognition for hand gesture based low-cost television control system. *IEEE Trans Consum Electron* 2014;60(1):107-15.
34. Srinivasan S, Pandharipande A, Sommen PC, van Sloun RJ. Ultrasonic array doppler sensing for human movement classification. *IEEE Sens J* 2014;14(8):2782-91.