



Received: 27 July 2016
Accepted: 09 October 2016
First Published: 24 October 2016

*Corresponding author: Rajesh Kaluri,
School of Information Technology and
Engineering, VIT University, Vellore,
India
E-mail: rajeshkaluri@yahoo.com

Reviewing editor:
Qingsong Ai, Wuhan University of
Technology, China

Additional information is available at
the end of the article

ELECTRICAL & ELECTRONIC ENGINEERING | RESEARCH ARTICLE

A framework for sign gesture recognition using improved genetic algorithm and adaptive filter

Rajesh Kaluri^{1*} and Ch. Pradeep Reddy¹

Abstract: Gesture based communication is the standard language utilized by the hard of hearing individuals for correspondence purpose. Despite the way that they precisely chat with each other by a method in sign language, they confront obscurity when they attempt to speak with individuals who can see sound, basically with the individuals who can't understand sign language. Consequently, an effective method to be produced to gain and recognize the sign motion language. In our proposed work we have planned a casing work for examining and distinguishing the sign motion language. The proposed technique is handled through various modules like Noise removal using adaptive filter, segmentation using region growing algorithm and feature extraction by using an improved genetic algorithm. Finally, the proposed technique will be assessed by contrasting with the support vector machine classifier.

Subjects: Cognitive Artificial Intelligence; Computational Linguistic & Language Recognition; Human Computer Intelligence; Machine Learning - Design

Keywords: adaptive filter; feature extraction; genetic algorithm; sign language recognition; speeded-up robust features (SURF)

1. Introduction

Communication via gestures as a sort of signal is a standout amongst the most characteristic methods for exchanging information for most hard of hearing individuals. The objective of gesture based



Rajesh Kaluri

ABOUT THE AUTHORS

Rajesh Kaluri obtained BTech in CSE from JNTU, Hyderabad, India and did MTech in CSE from ANU, Guntur, India and perusing PhD in majoring Gesture Recognition at VIT University, India. Currently, he is working as an assistant professor in School of Information Technology and Engineering, VIT University, India. He was a visiting professor in Guangdong University of Technology, China. His current research is in the area of computer vision.

Dr Ch. Pradeep Reddy is currently working with the School of Information Technology and Engineering, VIT University, Vellore, India. He received his MTech in Computer Science and Engineering from VIT University, India. He did his PhD in Computer Science and Engineering from VIT University, Vellore in 2014. His research interest includes mobile and wireless systems, cloud computing and pattern recognition.

PUBLIC INTEREST STATEMENT

Sign language is the most useful language in communicating with the people who can't speak and hear. Sign language recognition is the process where signs will be captured from the speaker and translates it into a text form. It helps to every individual to understand the feelings of the people who are lacking the power of hearing. In this paper, we compared various existing methods and have proposed a method to understand the signs easily from impaired hearing people.

communication, sign language recognition (SLR) is to give a productive and exact mechanism to translate gesture based communication into the content or discourse so that correspondence amongst hard of hearing and listening to society can be more helpful. SLR, as unitary of the important research areas of human-computer interaction (HCI), has spawned more and more interest in HCI society (Fang, Gao, & Zhao, 2004). Being as mind boggling as any talked language, SL has numerous signs, framed by outward appearances and important signals, including physical developments of the fingers, hands, wrists, arms, and head, each contrasting from another by minor changes close by movement, shape, position, and outward appearance. Appropriately, a communication via gestures can be considered as a gathering of important and easy to use hand motions, developments, and stances. Hand motion acknowledgment is the most ordinarily utilized methodology among other correspondence modalities in human-computer communication.

Dynamic hand motion correspondence is a more common and humanoid mode of correspondence with PCs, as contrasted and static hand motions (Abid, Petriu, & Amjadian, 2015). A phone based video calling framework for American sign language requires real time catch, encoding, and disentangling of advanced video on a cell gadget for transmission over the U.S. cell system. Cell gadgets have constrained handling force and battery life, forcing limitations on the multifaceted nature of the encoding and translating calculations (Ciaramello & Hemami, 2011). As per a standard Korean sign language (KSL) word reference, the 45-year-old Korean gesture based communication contains around 6,000 vocabulary words. Be that as it may, they are framed by consolidating a generally little number of essential motions. In addition, two sorts of motions of hands and fingers are utilized: one is static and the other is dynamic gestures. The previous comprises 31 particular stances communicating the dactylogy while the last is made up with evolving designs, constituting the primary body of the KSL and communicating distinctive implications of vocabulary words (Kim, Jang, & Bien, 1996).

The most regularly utilized measurements are hand shape/introduction, changes fit as a fiddle/introduction, hand area, developments of hand areas, hand-hand touching, hand-body touching (for the most part particular areas on the face), lip developments, outward appearance, and middle/shoulder posture and developments. Moreover, as a rule, connection is key to extraordinarily characterize the importance of a sign. In any case, that does not as a matter of course imply that the dynamical parts of gesture based communication have the same conduct and play the same semantic part as flow in talked languages. No less than three critical qualifications must be considered. Most importantly, the one-dimensionality of discourse makes it consecutive in nature (Lichtenauer, Hendriks, & Reinders, 2008).

The rest of the paper is organized as follows. Section 2 explains the researches related to our proposed method. Section 3 shows our proposed method being used for recognition of sign gesture using Region Growing, Genetic Algorithm and Adaptive Filter. Section 4 explains the result of the proposed methodology and finally the Section 5 concludes our proposed method with suggestions given for future works.

2. Literature review

Galka, Masior, Zaborski, and Barczewska (2016) have proposed a vigorous framework for Sign Language Gesture Acquisition and Recognition using an accelerometer glove and in addition its application in the acknowledgment of communication through signing signals. The fundamental information concerning movement sensors and the outline of the signal obtaining framework and also extend proposition are displayed. The assessment of the arrangement introduces the consequences of the motion acknowledgment endeavor by utilizing a chose set of gesture based communication motions with a depicted technique in light of hidden Markov model (HMM) and parallel HMM approaches.

Kosmidou and Hadjileontiadis (2009) has shown the work, information from five-channel surface electromyogram and 3D accelerometer from the endorser's prevailing hand were dissected utilizing intrinsic mode entropy (IMEn) for the robotized acknowledgment of Greek sign dialect (GSL) segregated signs. Discriminant examination was utilized to distinguish the successful sizes of the inherent

mode capacities and the window length for the computation of the IMEn that adds to the productive grouping of the GSL signs. Exploratory results from the IMEn examination connected to GSL signs relating to 60-word dictionary rehashed ten times by three local endorsers have indicated more than 93% mean grouping precision utilizing IMEn as the main wellspring of the arrangement highlight set.

Kelly, McDonald, and Markham (2011) have proposed a novel numerous case learning thickness grid calculation which naturally removes confined signs from full sentences utilizing the frail and boisterous supervision of content interpretations. The consequently separated secluded specimens are then used to prepare our spatiotemporal signal and hand stance classifiers. The investigations were done to assess the execution of the programmed sign extraction, hand stance characterization, and spatiotemporal signal spotting frameworks.

The temporal elements of a video-based motion are separated through forward, in reverse, and bidirectional expectations. The expectation mistakes are thresholded and collected into one picture that speaks to the movement of the grouping. Shanableh, Assaleh, and Al-Rousan (2007) have proposed highlight extraction plan was supplemented by straightforward grouping procedures, in particular, K closest neighbor (KNN) and Bayesian, i.e. probability proportion, classifiers. Exploratory comes about indicated grouping execution running from 97 to 100% acknowledgment rates.

Hikawa and Kaida (2015) proposed for hardware system for posture recognition using a hybrid network. The cross breed system comprises self-arranging map (SOM) and Hebbian system. Highlight vectors are removed from information stance pictures, which are mapped to a lower dimensional guide of neurons in the SOM. The Hebbian system is a solitary layer feedforward neural system prepared with a Hebbian learning calculation to distinguish classifications. Its power to revolution and scaling was enhanced by including irritation to the preparation information for the SOM-Hebb classifier. The entire framework is actualized on a field-programmable gate array utilizing novel video preparing design. The framework was intended to perceive 24 American gesture based communication hand signs, and its plausibility was checked through both recreations simulations and tests.

Zhou, Chen, Zhao, Yao, and Gao (2010) proposed a novel technique that adjusts the first model set to a particular endorser with his/her little measure of preparing information. Initially, partiality proliferation is utilized to separate the models of endorser free covered up Markov models; then the versatile preparing vocabulary can be consequently shaped. Taking into account the gathered sign signals of the new vocabulary, the blend of most extreme a posteriori and iterative vector field smoothing is used to produce underwriter adjusted models. The trial comes about on six underwriters exhibit that the proposed technique can decrease the measure of the adjustment information what's more, still can accomplish high recognition rate.

3. Proposed method

Our proposed method mainly consists of four steps. Firstly, in preprocessing process the noise removal can be done by using adaptive filtering and then the segmentation process will be carried out by region growing algorithm. After segmentation, feature extraction will be done by speeded-up robust features (SURF) algorithm with respect to the point feature. Classification will be done using improved genetic algorithm. Finally, the results will be compared with support vector machine and neural network classifiers. Figure 1 shows the flow of our proposed system.

3.1. Noise removal using adaptive filter

Adaptive filter is a versatile filter by having a linear filter which has a transfer function measured by variable parameters. According to an optimization algorithm, the parameters are adjusted with the transfer function (see Figure 2).

Figure 1. Framework for sign gesture recognition.

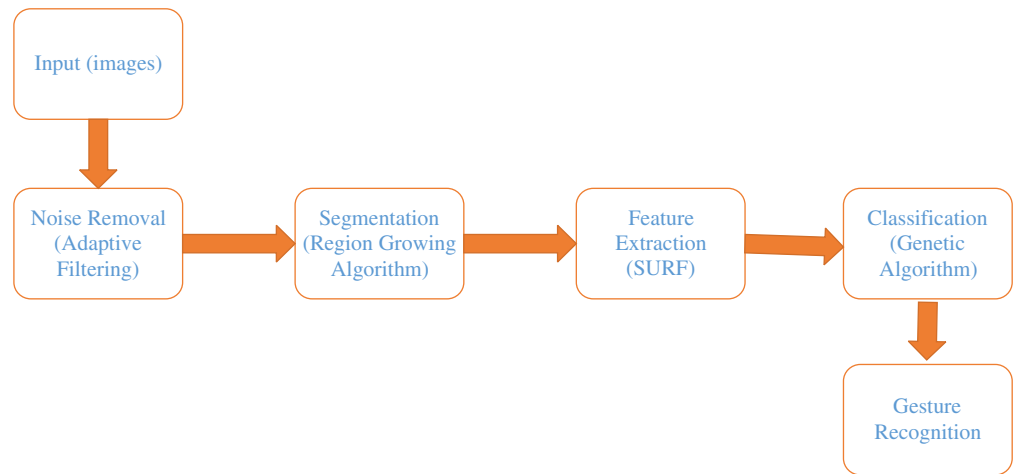
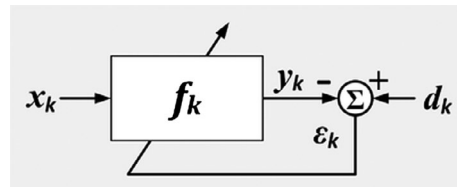


Figure 2. Noise removal in adaptive filter.



The working approach in adaptive filter is as follows.

- Step 1: Take the input image (true color image).
- Step 2: Convert the input color image to grayscale image.
- Step 3: Increase Gaussian noise to the image so that the image size is large.
- Step 4: Eliminate noise from the image by using wiener2 function $K = \text{wiener2}(J, [x, x])$.

3.2. Segmentation by region growing algorithm

Segmentation is a process of analyzing an image and isolating the parts of an image by its representation. A final segmentation can be improved with global labeling of a model by controlling its energy and region consistency (Liu, Seyedhosseini, & Tasdizen, 2015). Where an edge based method may endeavor to discover the item limits and after that find the article itself by filling them in, a locale based method takes the inverse methodology. Region growing algorithm is one of the best method to use for image segmentation process. Firstly, a seed point of an image will be identified and if the nearest pixel has a similar pixel ratio then seed point will move into the next pixel. It uses a well-known technique in order to find the blocks of an image. The basic formula for region growing algorithm is:

$$U_{i=1}^n R_i = R \tag{1}$$

where R is region and i is the pixel point.

Selecting seed point is the primary step in the region growing algorithm and which is based on user criteria. The regions are then grown from these seed points to adjacent points depending on a region membership criterion. Region growing algorithm makes the images rich with the conditions like minimum area of threshold value, better information of the image and similarity threshold value.

The region growing is a three phase method which comprises gridding, selection of seed point, applying region growing to the point. In gridding, a single image is distributed into numerous smaller images by sketching an illusory grid over it. That is, gridding grades in changing the image into numerous smaller grid images. The grids are generally square in shape and the grid number to which the novel image is divided into is a flexible. For our tentative estimation of the projected method, we have dripped the original image into 4, 18 and 24 grids. Gridding effects in smaller grids so that examination can be passed out simply. Assume the pixel is having the intensity value I_p , and the neighbouring pixel has the value I_N and the intensity threshold is set as T_p , then if $\|I_p - I_N\| \leq T_p$, then intensity constrain is encountered and contented.

3.3. Feature extraction by SURF

Distinctive individuals have diverse hand sizes, body sizes, marking propensities, marking rhythms, et cetera, which prompts assortments when they sign the same word. Good features will have the properties like accuracy, distinctiveness, efficiency, locality, quantity etc. (Tuytelaars & Mikolajczyk, 2008). The befuddle between the preparation information and the test information prompts poor acknowledgment execution. One contrasting option to tackle this issue is gathering enough information from various individuals to prepare SI models.

SURF uses basic Hessian matrix for interest point detection with respect to the integral images (P) at a location (x, y) (Bay, Ess, Tuytelaars, & Van Gool, 2008). Thus the sum of pixels of an input image (X) can be coined from the origin and P .

$$I_{\Sigma}(P) = \sum I \leq x_i = 0 \sum j \leq y_j = 0X(i, j) \quad (2)$$

The dimensionality of the descriptor has direct impact on both its computational complexity and point-matching robustness/accuracy. A short descriptor may be more robust against appearance variations, but may not offer sufficient discrimination and thus give too many false positives. The box filter of size 9×9 is an approximation of a Gaussian with $\sigma = 1.2$ and represents the lowest level (highest spatial resolution) for blob-response maps.

Interest points need to be found at different scales, not least because the search of correspondences often requires their comparison in images where they are seen at different scales. Scale spaces are usually implemented as an image pyramid. The images are repeatedly smoothed with a Gaussian and then sub-sampled in order to achieve a higher level of the pyramid.

3.4. Classification

Classification is the major section in recognizing the gestures. The classifier provides exact classification of sign gesture which aid in improved recognition rate. In our proposed system, an improved genetic algorithm has been used for recognizing the gestures from the images and which showed a better performance than the existing systems. Genetic algorithm is one which is mainly used in solving problems related to constrained and unconstrained approaches. It mainly consists of three basic steps like selection, crossover and mutation (see Figure 3).

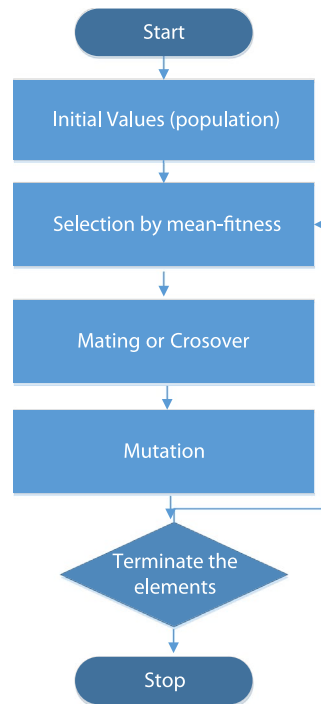
3.4.1. Improved genetic algorithm

Improved genetic algorithm firstly consider the segmented images from the previous process. It mainly concentrates on the quality images with high entropy values of the pixel image. The fitness value will be calculated for all the pixel and entropy values (Samra & Khalefah, 2014; Zhang, Wang, & Li, 2016).

$$X_i = [x_0^{(i)}, x_1^{(i)}, \dots, x_{N_L-1}^{(i)}] \quad 0 \leq i \leq N_p - 1 \quad 0 \leq j \leq N_L - 1 \quad (3)$$

From Equation (3), $x_j^{(i)}$ represents j th chromosome gene, N_p is population pool, N_L is chromosome length.

Figure 3. Step by step process of improved genetic algorithm.



A sort of target method is the fitness method, which is the top target parameter to the improved worth. By utilizing the accompanying equation, the fitness capacity is assessed.

$$F_i = \sum_{i=1}^n d_i/n \tag{4}$$

Cross over operation can be done with the help of two parent chromosomes. The level of the qualities will be picked by the recent child chromosomes with a hybrid rate (Yang, Ahuja, & Tabb, 2002). The fitness capacity is utilized to the recently delivered tyke chromosome subsequent to creating another chromosome. The chromosome for computing the hybrid rate is set as:

$$\text{Crossover rate } (C_r) = \frac{CO_g}{C_L} \tag{5}$$

where, CO_g is number of gene crossovered, C_L is length of chromosome.

Adaptive mutation operation is the union of the arrangement is stimulated. Based on the mutation rate (M_r) the mutation operation is transmitted out (Kaluri & Pradeep Reddy, 2016).

$$\text{Rate of mutation } (M_r) = \frac{M_p}{N_L} \tag{6}$$

where, M_p is the mutation point, N_L is the chromosome length.

Finally, the selection of chromosomes will be happened during the determination methodology (Madeo, Peres, & Lima, 2016; Mitra & Acharya, 2007), the N_p aimlessly produce chromosomes and the N_p new chromosomes are situated in a choice pool on the premise of their wellness values. The chromosomes that contain great wellness possess the top positions of the pool in the choice pool.

Figure 4. Sign gesture recognition for *meat*: (a) input frame, (b) noise removed, output, (c) segmented output.



Figure 5. Sign gesture recognition for *green*: (a) input frame, (b) noise removed, output, (c) segmented output.

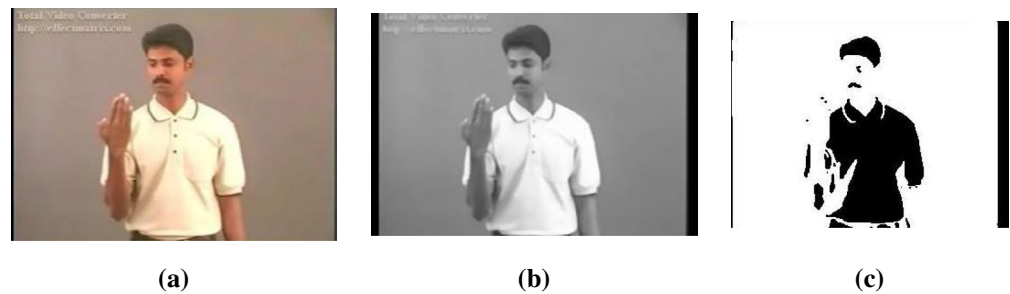
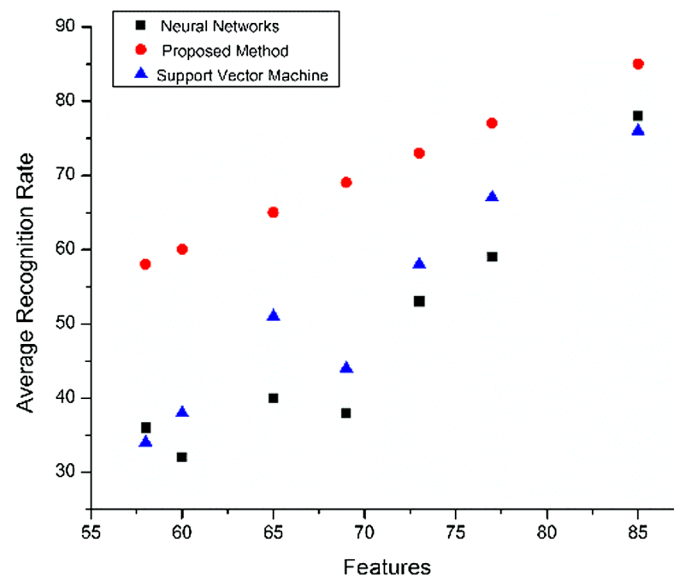


Figure 6. Recognition rate in proposed vs. existing methods.



Figures 4 and 5 represents the sign gesture recognition for *meat* and *green* color.

4. Experimental results

Adaptive Filtering works better in the noise removal process across the images with the function f_k and its parameters. Region growing algorithm worked with the seed points and hence its maintains volatility of the images. Recognition of sign gesture from the images is a very complex task and which was carried out successfully with the improved genetic algorithm and thus the recognition rate is more compared to the other classifiers. The detailed procedure is shown in Figure 6 by using three different frames namely input frame, noise removal frame and segmented output frame.

5. Conclusion

Sign gesture recognition have been successfully carried out in MATLAB. Based on our proposed technique the average recognition rate is high compared to the existing methods like neural networks and support vector machine. By using improved genetic algorithm, the segmentation rate is reduced gradually for the different images of the video. We employed feature extraction stage along with the segmentation which forms an efficient process in extracting the required measures for recognition. The results of our method gives a very good recognition rate compared to the exiting methods.

Funding

The authors received no direct funding for this research.

Author details

Rajesh Kaluri¹
E-mail: rajeshkaluri@yahoo.com
Ch. Pradeep Reddy¹
E-mail: pradeep@vit.ac.in

¹ School of Information Technology and Engineering, VIT University, Vellore, India.

Citation information

Cite this article as: A framework for sign gesture recognition using improved genetic algorithm and adaptive filter, Rajesh Kaluri & Ch. Pradeep Reddy, *Cogent Engineering* (2016), 3: 1251730.

References

- Abid, M. R., Petriu, E. M., & Amjadian, E. (2015). Dynamic sign language recognition for smart home interactive application using stochastic linear formal grammar. *IEEE Transactions on Instrumentation and Measurement*, 64, 596–605.
<http://dx.doi.org/10.1109/TIM.2014.2351331>
- Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, 110, 346–359.
<http://dx.doi.org/10.1016/j.cviu.2007.09.014>
- Ciaramello, F. M., & Hemami, S. S. (2011). A computational intelligibility model for assessment and compression of american sign language video. *IEEE Transactions on Image Processing*, 20, 3014–3027.
<http://dx.doi.org/10.1109/TIP.2011.2132730>
- Fang, G., Gao, W., & Zhao, D. (2004). Large vocabulary sign language recognition based on fuzzy decision trees. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 34, 305–314.
<http://dx.doi.org/10.1109/TSMCA.2004.824852>
- Galka, J., Masior, M., Zaborski, M., & Barczewska, K. (2016). Inertial motion sensing glove for sign language gesture acquisition and recognition. *IEEE Sensors Journal*, 16, 6310–6316.
- Hikawa, H., & Kaida, K. (2015). Novel FPGA implementation of hand sign recognition system with SOM-Hebb classifier. *IEEE Transactions on Circuits and Systems for Video Technology*, 25, 153–166.
<http://dx.doi.org/10.1109/TCSVT.2014.2335831>
- Kaluri, R., & Ch, P. R. (2016). An overview on human gesture recognition. *International Journal of Pharmacy and Technology*, 8, 12037–12045.
- Kelly, D., McDonald, J., & Markham, C. (2011). Weakly supervised training of a sign language recognition system using multiple instance learning density matrices. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 41, 526–541.
<http://dx.doi.org/10.1109/TSMCB.2010.2065802>
- Kim, J. S., Jang, W., & Bien, Z. (1996). A dynamic gesture recognition system for the Korean sign language (KSL). *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 26, 354–359.
- Kosmidou, V. E., & Hadjileontiadis, L. J. (2009). Sign language recognition using intrinsic-mode sample entropy on sEMG and accelerometer data. *IEEE Transactions on Biomedical Engineering*, 56, 2879–2890.
<http://dx.doi.org/10.1109/TBME.2009.2013200>
- Lichtenauer, J. F., Hendriks, E. A., & Reinders, M. J. (2008). Sign language recognition by combining statistical DTW and independent classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30, 2040–2046.
<http://dx.doi.org/10.1109/TPAMI.2008.123>
- Liu, T., Seyedhosseini, M., & Tasdizen, T. (2015). Image segmentation using hierarchical merge tree. *arXiv preprint arXiv:1505.06389*.
- Madeo, R. C. B., Peres, S. M., & Lima, C. A. (2016). Gesture phase segmentation using support vector machines. *Expert Systems with Applications*, 56, 100–115.
<http://dx.doi.org/10.1016/j.eswa.2016.02.021>
- Mitra, S., & Acharya, T. (2007). Gesture recognition: A survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 37, 311–324.
<http://dx.doi.org/10.1109/TSMCC.2007.893280>
- Samra, G. A., & Khalefah, F. (2014). Localization of license plate number using dynamic image processing techniques and genetic algorithms. *IEEE Transactions on Evolutionary Computation*, 18, 244–257.
- Shanableh, T., Assaleh, K., & Al-Rousan, M. (2007). Spatio-temporal feature-extraction techniques for isolated gesture recognition in arabic sign language. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 37, 641–650.
<http://dx.doi.org/10.1109/TSMCB.2006.889630>
- Tuytelaars, T., & Mikolajczyk, K. (2008). Local invariant feature detectors: A survey. *Foundations and Trends in Computer Graphics and Vision*, 3, 177–280.
- Yang, M. H., Ahuja, N., & Tabb, M. (2002). Extraction of 2D motion trajectories and its application to hand gesture recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24, 1061–1074.
<http://dx.doi.org/10.1109/TPAMI.2002.1023803>
- Zhang, Y., Wang, S., & Li, J. (2016). Improved particle filtering techniques based on generalized interactive genetic algorithm. *Journal of Systems Engineering and Electronics*, 27, 242–250.
- Zhou, Y., Chen, X., Zhao, D., Yao, H., & Gao, W. (2010). Adaptive sign language recognition with exemplar extraction and MAP/IVFS. *IEEE Signal Processing Letters*, 17, 297–300.
<http://dx.doi.org/10.1109/LSP.2009.2038251>



© 2016 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.

You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions

You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.



Cogent Engineering (ISSN: 2331-1916) is published by Cogent OA, part of Taylor & Francis Group.

Publishing with Cogent OA ensures:

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

Submit your manuscript to a Cogent OA journal at www.CogentOA.com

