Front. Comput. Sci.
https://doi.org/10.1007/s11704-018-7082-4

LETTER

A real-time hand-signs segmentation and classification system using fuzzy rule based RGB model and grid-pattern analysis

Muhammad Aminur RAHAMAN (✉)[1], Mahmood JASIM[1], Md. Haider ALI[1], Tao ZHANG[2], Md. HASANUZZAMAN[1]

1 Department of Computer Science and Engineering, University of Dhaka, Dhaka-1000, Bangladesh
2 Department of Automation, School of Information Science & Engineering, Tsinghua University, Beijing 100084, China

1 Introduction

Computer vision based researches of automatic hand-signs classification using simple camera based image are increasing rapidly in the present world. Hand-signs are the basic unit of the sign languages. So, successful sign language recognition depends on the recognition or classification of hand-signs successfully.

The contributions of the proposed system are in three areas: 1) Hand-signs detection and segmentation using proposed fuzzy rule based RGB (FRB-RGB) model combining with Haar-classifiers and motion queue in cluttered and dynamic background with illumination variation environment; 2) Rotation, translation and scale invariant window-grid vector (WGV) is proposed for feature vector; 3) We have proposed simple mathematical analysis based classifier that classifies the hand-signs accurately in real-time.

The technical details, proofs and evaluations can be found in the support information.

2 Proposed system description

2.1 Hand-sign segmentation

After capturing the image sequence $I^m_{rgb}(x, y)$, the proposed system starts to segment skin-like area from the ROI using proposed fuzzy rule based RGB (FRB-RGB) model. The proposed algorithm segments each skin-color pixel according to a set of fuzzy rules from the ROI containing hand-signs using RGB color spaces C_{rgb}, where each color pixel is represented by a color vector $P^k_C = \{P^R_C, P^G_C, P^B_C\} \in [0, 255]$. The system extracts probable binary hand-signs $I^m_{bin}(x, y)$ based on the skin-color pixels P^i_{skin} with specific motions. To remove noise, morphological closing and Gaussian smoothing operations are done on the binary image.

2.2 Feature vector generation

Before generating the grid-pattern represented by window-grid vector (WGV), the system rotates the hand-sign centered on center of gravity (COG) with respect to align the point of wrist (POW) to the vertical line which is drawn through the COG to make it rotation invariant. After clipping and normalization, the system generates WGV denoted by Ω represented by Eq.(1). In Eq.(1), each window-grid value (ω^η) is obtained by calculating the percentage of white pixels (“1”) presented in each window-grid (ω^η) of the window-grid mask (WGM) for each hand-sign.

$$\Omega = (\omega^0, \omega^1, \omega^2, \ldots, \omega^{M_\omega - 1}), \tag{1}$$

where ω^η represents each window-grid of Ω; $\eta = 0, 1, 2, \ldots, M_\omega - 1$; M_ω= (Size of input image)/(Size of each window)=$(150 \times 150)/(30 \times 30) = (5 \times 5) = 25$.

The area (Λ) of a hand-sign is calculated by counting the total number of black pixels in normalized image $I^m_{norm}(x, y)$.

Received March 7, 2017; accepted September 22, 2017

E-mail: aminur.wg@gmail.com

The mean height of the hand-sign is calculated from the area divided by total number of column ($\frac{\Lambda}{N}$) of the image. Finally, the features are collected as $\xi = [\Omega, \Lambda, \lambda]$ in the feature space for each hand-sign.

2.3 Training method

A training module is generated for each sign class i, the system calculates $j = 100$ features in each feature type of each sign X_i from the 10 signers where each signer performs 10 hand-signs. The system stores the three types of extracted features $\Omega_i^j, \Lambda_i^j, \lambda_i^j$ in to the three sub-classes of each sign class i and creates the training database.

2.4 Recognition method

In the recognition or classification module, the hand-signs are recognized by measuring the maximum similarity between the extracted features $\xi = [\Omega, \Lambda, \lambda]$ and the pre-trained features $\xi_i^j = [\Omega_i^j, \Lambda_i^j, \lambda_i^j]$ of the hand-signs using Eq. (2).

$$\begin{aligned} Sim_{\text{Max}}(\xi, \xi_i^j) = {} & \delta_1 \times Sim_{\text{Max}}(\Omega, \Omega_i^j) \\ & + \delta_2 \times Sim_{\text{Max}}(\Lambda, \Lambda_i^j) \\ & + \delta_3 \times Sim_{\text{Max}}, (\lambda, \lambda_i^j), \end{aligned} \tag{2}$$

where δ_1, δ_2, and δ_3 are the recognition coefficient with the condition, $\delta_1 + \delta_2 + \delta_3 = 1$; $Sim_{\text{Max}}(\Omega, \Omega_i^j)$ is the maximum similarity between test WGV and pre-trained WGV of the hand-signs; $Sim_{\text{Max}}(\Lambda, \Lambda_i^j)$ is the maximum similarity between test area and pre-trained area of the hand-signs; and $Sim_{Max}(\lambda, \lambda_i^j)$ is the maximum similarity between test mean height and pre-trained mean height of the hand-signs.

3 Experimental result and discussion

The proposed system uses a built-in webcam of ASUS A42F series laptop with Intel Corei3 processor of 2.40 GHz and 2GB RAM. The system uses Microsoft® Visual Studio® 2008 and EmguCV (C# and OpenCV wrapper) in 32-bit operating system of Microsoft Windwos7®, as system development platform.

For testing the system, six sets of images have been used for 46 hand-signs of Bengali sign language (BdSL) and 10 hand-signs of Chinese sign language (ChSL). The total number of testing images are 33,600 where $(4,600 \times 6) = 27,600$ for BdSL and $(1,000 \times 6) = 6,000$ for ChSL in six different background with illumination variation environments (E1, E2, E3, E4, E5, and E6) from 10 new signers. The system training and testing process in the six environments are described in details in supporting information.

3.1 Result of skin-color segmentation using FRB-RGB model

The proposed FRB-RGB model for skin-color segmentation is trained using 100 BdSL hand-signs from the 10 different skin-color people, where each people performs 10 hand-signs in different environment with illumination variation. To obtain the ground truth training dataset for FRB-RGB model, we manually segment and labeled 100 images which consists of 10 subsets for 10 signers. The result of the proposed 'FRB-RGB' model for skin-color segmentation and the comparison with some other existing methods using HSV [1]; Cascaded AdaBoost [2]; SPM (skin probability map) based RGB [3]; YCbCr-K-Means [4]; Fuzzy-YCbCr based Anfis [5] are shown in Table 1. The proposed skin-color segmentation method is also applied using HSV and YCbCr color model using the selected 33,600 off-line samples of hand-signs, in which the proposed model gives accurate results for skin-color segmentation with six challenging environments. For each method we compute the average computational time of three runs of 33,600 images.

Table 1 Hand-signs segmentation accuracy and average computation times for different methods

Skin-color segmentation methods	Accuracy /%		Average computation times/milliseconds	
	R1	R2	R1	R2
HSV [6]	86.03	77.12	148,100.25	348,200.87
Cascaded AdaBoost [2]	94.05	91.09	700,035.69	1,420,050.93
SPM based RGB [3]	89.97	86.57	69,027.35	145,070.75
YCbCr- K-Means [4]	94.43	91.47	819,040.55	1,628,035.75
Fuzzy-YCBCR based Anfis [5]	94.48	91.43	140,040.45	290,027.55
FRB-HSV (implemented by us)	95.97	93.93	81,000.95	167,050.88
FRB- YCbCr (implemented by us)	94.83	93.17	81,070.37	168,075.75
FRB-RGB (proposed)	96.47	94.53	81,090.15	168,035.95

Note: R1 represents the classification accuracies using different skin-color segmentation methods combining with Haar-classifiers; R2 represents the classification accuracies using different skin-color segmentation methods without combining Haar-classifiers

3.2 Result of hand-signs classification

The proposed system achieves mean accuracy of 95.67% for BdSL and 96.57% for ChSL with the computational cost of 8.01 milliseconds per frame in six challenging environments. From Table 2, we claim that the test results of hand-signs classification of the proposed system (denoted by 'HSSCS') show better performance than existing reputed BdSL recognition system [1, 6–10] in terms of accuracy and computational cost.

Table 2 Comparison of different methods for hand-signs classification in BdSL

Methods	Accuracy/%						Computational cost/(ms/f)
	E1	E2	E3	E4	E5	E6	
CM [1]	95.57	89.35	55.34	53.45	45.89	34.76	8.02
HSSCS	96.96	96.54	96.11	95.72	95.04	93.67	8.01
SHIFT-SVM [7]	88.67	87.08	79.35	77.89	67.75	63.55	90.09
LDA-ANN [8]	95.32	90.32	87.48	67.09	56.87	40.09	92.04
Skin-KNN [6]	95.56	93.88	87.07	79.38	73.05	63.85	96.46
Fuzzy-rule-based [10]	92.25	90.83	87.47	85.9	83.46	79.67	100.17
Haar-KNN [9]	93.85	93.9	90.45	89.49	87.48	86.95	94.88

4 Conclusion

This paper presents a computer vision-based hand-signs segmentation and classification system using proposed fuzzy rule based RGB (FRB-RGB) model and grid-pattern analysis. The system segments human skin-color robustly using the proposed FRB-RGB model in cluttered background with illumination variation. Finally, the system classifies the hand-signs with satisfactory results using the proposed feature vector [WGV, area, mean height] based classifier by calculating the maximum inter-correlation coefficient (ICC) between pre-trained feature vectors and test feature vector. The proposed system is applicable to assist as an interpreter for communication between sign and non sign people. It can also be used for human machine interaction (HMI) and virtual reality using hand-signs.

Acknowledgements This research is partially supported and funded by the Information and Communication Technology (ICT) Division, Ministry of Posts, Telecommunications and IT, Government of the People's Republic of Bangladesh.

Supporting information The supporting information is available online at journal.hep.com.cn and link.springer.com.

References

1. Rahaman M A, Jasim M, Zhang T, Ali M H, Hasanuzzaman M. Real-time bengali and chinese numeral signs recognition using contour matching. In: Proceedings of the IEEE International Conference on Robotics and Biomimetics. 2015, 1215–1220
2. Lü W, Huang J. Skin detection method based on cascaded adaboost classifier. Journal of Shanghai Jiaotong University (Science), 2012, 17(2): 197–202
3. Xu J, Zhang X. A real-time hand detection system during hand over face occlusion. International Journal of Multimedia and Ubiquitous Engineering, 2015, 10(8): 287–302
4. Qiu-yu Z, Jun-chi L, Mo-yi Z, Hong-xiang D, Lu L. Hand gesture segmentation method based on YCBCR color space and K-means clustering. International Journal of Signal Processing, Image Processing and Pattern Recognition, 2015, 8(5): 105–116
5. Iraji M S, Tosinia A. Skin color segmentation in YCBCR color space with adaptive fuzzy neural network (anfis). International Journal of Image, Graphics and Signal Processing, 2012, 4(4): 35–41
6. Rahaman M A, Jasim M, Ali M H, Hasanuzzaman M. Real-time computer vision based bengali sign language recognition. In: Proceedings of the 17th International Conference on Computer and Information Technology. 2014, 192–197
7. Yasir F, Prasad P W C, Alsadoon A, Elchouemi A. Sift based approach on bangla sign language recognition. In: Proceedings of the 8th IEEE International Workshop on Computational Intelligence and Applications. 2015, 35–39
8. Yasir R, Khan R A. Two-handed hand gesture recognition for bangla sign language using LDA and ANN. In: Proceedings of the 8th International Conference on Software, Knowledge, Information Management and Applications. 2014, 1–5
9. Jasim M, Zhang T, Hasanuzzaman M. A real-time computer vision-based static and dynamic hand gesture recognition system. International Journal on Image and Graphics, 2014, 14(01n02): 1–19
10. Ayshee T F, Raka S A, Hasib Q R, Hossain M, Rahman R M. Fuzzy rule-based hand gesture recognition for bengali characters. In: Proceedings of the IEEE International Advance Computing Conference. 2014, 484–489