

Sign Language Gesture Recognition with Bispectrum Features using SVM

Hasmath Farhana Thariq Ahmed¹, Hafisoh Ahmad^{1, a)}, Swee King Phang¹,
Chockalingam Aravind Vaithilingam¹, Houda Harkat^{2, 3} and Kulasekharan
Narasingamurthi⁴

¹ School of Engineering, Taylor's University, 1, Jalan Taylor's, 47500, Subang Jaya, Selangor, Malaysia.

² Faculty of Sciences and Technologies, University of Sidi Mohamed Ben Abdellah, Route Imouzzar Fez, BP 2626, FES 30000, Morocco.

³ Faculty of Science and Technology, University of Algarve, Reitoria, Campus de Gambelas, 8005-139, FARO, Portugal.

⁴ Simulation Metier, Valeo India Pvt Ltd., 1/396, Old Mahabalipuram Road, Navallur, Chennai, TamilNadu 600130, India.

a) Corresponding Author: Hafisoh.Ahmad@taylors.edu.my

Abstract. Wi-Fi based sensing system captures the signal reflections due to human gestures as Channel State Information (CSI) values in subcarrier level for accurately predicting the fine-grained gestures. The proposed work explores the Higher Order Statistical (HOS) method by deriving bispectrum features (BF) from raw signal by adopting a Conditional Informative Feature Extraction (CIFE) technique from information theory to form a subset of informative and best features. Support Vector Machine (SVM) classifier is adopted in the present work for classifying the gesture and to measure the prediction accuracy. The present work is validated on a secondary dataset, SignFi, having data collected from two different environments with varying number of users and sign gestures. SVM reports an overall accuracy of 83.8%, 94.1%, 74.9% and 75.6% in different environments/scenarios.

INTRODUCTION

Device-free sensing adopt optical or depth cameras [1-3], commercial devices leveraging the radio frequency signals [4-6] for predicting human gestures with higher accuracy. With recent advancements in wireless technology, sensing using Wi-Fi devices are preferred over other technologies as it is easy to deploy and it performs sensing in a privacy-preserving manner. Wi-Fi based sensing has a wide spread scope in the field of computer vision [7-11]. Human Gesture Recognition (HGR) system recognizes gestures by capturing Channel State Information (CSI) values of Wi-Fi signal reflection pattern in subcarrier level. Device free sensing system adopts learning approach for performing the classification task and measures the recognition accuracy. State of the art reports extensive research work carried out for extracting and selecting informative feature information, as it highly influence the accuracy or recognition performance of learning algorithms. For achieving remarkable recognition accuracy, Deep Learning approaches demands sizable data for auto feature extraction and classification, however, have poor data interpretability. Whereas, Machine Learning (ML) approaches perform well with constrained data size and attribute towards the handcrafted features.

The Wi-Fi signal reflection from commercial devices exhibits non-Gaussianity and adds high frequency noise in the raw CSI value extracted [12]. The conventional sensing systems filter out the noise by processing the signal value and perform linearization to obtain the phase change of the signal from the CSI. Recent research advancements in the field of gesture recognition leveraging CSI metric uses the second order statistical method and deals with Gaussian signal distribution for extracting the feature information. There are many second order statistical tool like Principal Component Analysis (PCA) to pre-process the large set of signal information and reduce its dimensionality by picking

only the principal components. However, second order statistical method has limitation to address the non-Gaussian signal distribution. The aforementioned is a limitation in the reported research which motivated the present work to explore the Higher Order Statistical (HOS) method using bispectrum features (BF). The proposed idea uniquely uses Conditional Informative Feature Extraction (CIFE)[13], a feature selection (FS) method derived from information theory for estimating the recognition accuracy. Proposed study adopts multilevel ML for measuring the gesture recognition performance. It benchmarks the application of proper feature extraction technique for machine learning classifier.

The contribution of present work is summarized as follows:

1. Higher order bispectrum features are extracted from the CSI values (without applying any pre-processing technique), and applies a feature selection method, considering the non-Gaussianity nature of the Wi-Fi signals.
2. Classifying and estimating the recognition accuracy with the feature subset derived implementing CIFE selection algorithm, adopting a machine learning approach.

To the extent of work in the device free sensing using Wi-Fi CSI, this is one of its kinds of work to propose the combination of BF + CIFE method adopting ML approach for sign gesture recognition using CSI values and reports the classification accuracy. The organization of the paper is as follows: Related work section reports the literature that are related to the present work; Preliminaries section provides the basic concepts and understanding of the work as preliminaries; System overview section introduces the system under study and the methodology; Implementation and evaluation section briefs about the implementation and evaluates the methodology of the proposed work; Conclusion section concludes the discussion on device free gesture recognition.

RELATED WORK

State of the art device free sensing leveraging CSI applies PCA, a feature extraction method on the pre-processed CSI traces. For example, PCA was adopted in detection application such as localizing human, and able to achieve 97% accuracy using SVM with less localization error in across different environment[14]. The other works such as WFID[15], R-TTWD[16], FallDeFi[17], R-DEHM[18] and WiFind[19] uses PCA and achieves remarkable recognition accuracy. BodyScan[20] computes power spectral density (PSD) from PCA features for recognizing activities with 72.5% and detecting breathing rate with 97.4% accuracy. The CSI dynamics is also studied using model-based methods that interprets the gesture and perform recognition using mathematical formulations, without signal pre-processing[20-23]. Some of the works adopts feature selection paradigm for maximizing the recognition accuracy. SVM reports considerable performance using forward and backward feature selection method and selects optimal set of 14 features from the set of 24 features[24]. Other recognition system such as WiHear[25] and Wi-Fi-ID[26] also incorporates feature selection step using Multiclass/Multicluster selection and Relief methods respectively for extracting best features for achieving recognition performance.

To the best of our knowledge, no existing work reports the implementation of Bispectrum Features (BF) in gesture recognition domain using Wi-Fi signals. In this paper, we explore a new gesture recognition paradigm using the raw CSI values, without pre-processing the signal information, as summarized in the Introduction section. The information on how the BF will perform in the gesture recognition task is unknown in the technical community. This work will fill the void in this domain.

Channel State Information (CSI)

The channel condition of a wireless medium will be unstable and encounter constant change in signal reflection depending on the environmental factors. Recent advancements in wireless technologies with introduction of Orthogonal Frequency Division Multiplexing (OFDM) enables higher transmission rate of packets with minimal Bit Error Rate (BER). Commercial Wi-Fi devices that comply IEEE 802.11n with OFDM also implement Multiple Input Multiple Output (MIMO) concept, permitting data transmission over multiple antennas. Optimizing channel over multiple antennas becomes possible in IEEE 802.11n devices by accessing the physical layer information in the form of CSI. OFDM complying devices can read the Channel Frequency Response (CFR) as CSI mainly with amplitude and phase change of the signal for every subcarrier as,

$$Rx = HTx + N \tag{1}$$

Where, Rx and Tx is the receiver and transmitter signal strength; H is the CSI channel matrix and N is the noise.

Bispectrum Feature (BF)

HOS has third, fourth and higher order spectra, which is referred as ‘Polyspectra’. Third order polyspectrum commonly referred as Bispectrum is preferred than the other higher order extensions, as it is computationally easy. The proposed study leverages the bispectrum features by performing Fourier transform on the third order moments of cumulant value. The bispectrum is represented as follows,

$$B(f_i, f_j) = E[(X(f_i) \times X(f_j) \times X^*(f_i + f_j))] \quad (2)$$

Where, $X(f)$ represents Fourier transform of random signal $x(nT)$, n is the integer and T is the sampling interval. $E[.]$ and X^* represents the expectation operator and the complex conjugate. $B(f_i, f_j)$ represents a symmetric and complex valued function of frequencies (f_i, f_j) , known as bispectrum. The symmetric function of bispectrum can extract different type of features and computes the bispectrum in a form of triangular region[27]. The proposed study considers 3 features (F_1 , F_2 and F_3): denotes sum of amplitudes (logarithmic), sum of amplitudes (logarithmic) of the diagonal elements and first order moments of amplitudes of the diagonal elements of the bispectrum[28].

SYSTEM OVERVIEW– HGR USING WI-FI SIGNALS

Present work highlights the evidence of nonlinearity present in the CSI values by applying a higher order statistical method and extract bispectrum features. Fig. 1 show the process flow carried out in the present work. The signal reflections in terms of CSI dynamics for different set of gestures is stored with respective labelling as a dataset. The proposed work derives third order features, without applying any pre-processing technique on the acquired CSI values. Subsequently, selection algorithms are applied on the extracted set of features for minimizing the computational cost. Lastly, using a machine learning approach the selected set of features are classified and the performance in terms of accuracy will be measured.

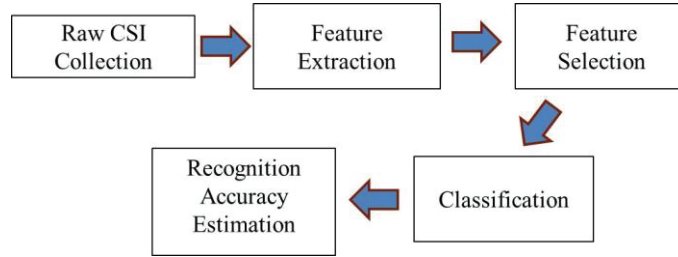


FIGURE 1. Process adopted for gesture recognition using Wi-Fi signals

Fig. 2 shows the detailed process of gesture recognition from raw CSI signal for estimating accuracy. The process starts with extraction of Bispectrum features from the CSI data containing raw signal information as explained earlier. Fig. 2(a) shows that the total number of extracted features, F_n , from the raw CSI traces. Information theory feature selection method chose an optimal set of features F_a , with mutual information as a selection metric from the total F_n features as presented in Fig. 2(b). The value F_a , will be assigned experimentally for selection algorithm to emphasize the impact of number of selected features and its corresponding accuracy. Lastly, selected set of bispectrum features will be handled using Machine Learning approach and estimate the classification accuracy as shown in Fig. 2(c).

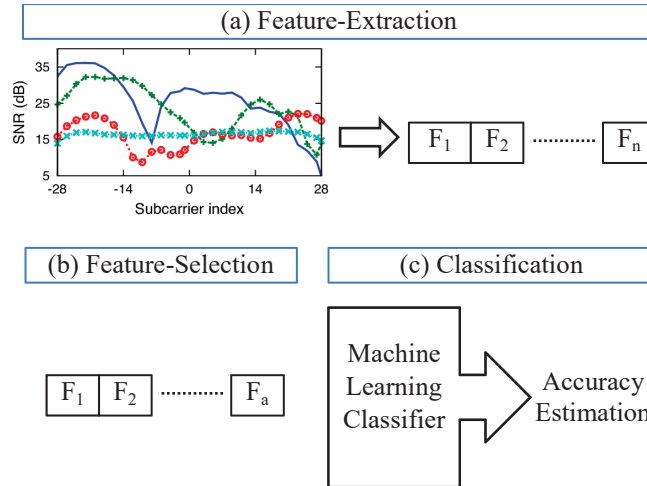


FIGURE 2. HGR estimation process

Feature Extraction

Higher order statistical method applied on CSI data will extract the bispectrum features[27], considering the number of transmitting channels (N_t) and receiving channels (N_r), by equipping the receiver with Intel 5300 chip. The receiving channel encapsulates the CSI measurements from every sub carriers (N_s), with a predefined set instance for a CSI gesture samples. The proposed study, extracts three bispectrum features (F_1, F_2, F_3) from N_t channels having N_s subcarriers per channel as shown in Fig. 3. Thus, a total of F_n bispectrum features ($F_n = 3 \times N_t \times N_s$) were extracted as explained earlier.

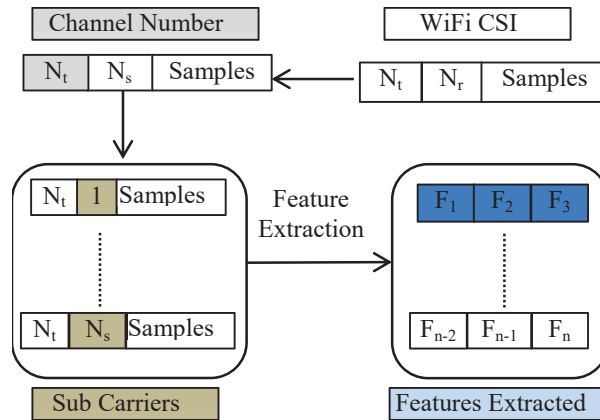


FIGURE 3. Extraction of bispectrum features from the receivers

Feature Selection

The proposed study performs features selection using CIFE, which could extract informative features for reducing the computational time. CIFE reduces the F_n bispectrum features to F_a selected features[29]. The selection process assigns the number of features to be selected through experimental results, in this case the value assigned as $F_a = 90$. Selection step incorporated here plays a major role in eliminating the redundant feature information; in turn achieve results with less execution or training time. The machine learning algorithm measures the performance of CIFE algorithm with accuracy as a metric.

Gesture Recognition

SVM Gaussian kernel evaluates the classification accuracy of sign language gesture recognition task by mapping the nonlinear samples to a higher dimensional space. SVM performs the classification task using regularization parameter C and γ classification parameter (non-linear) of the Gaussian kernel. The present work adopt CIFE, that takes one values amongst the given C values : 0.003, 0.01, 0.03, 0.1, 0.3, 1, 3, 10, 30, 100 and 300 in a sequence, one at time to the classification parameter γ performing $11 \times 11 = 121$ iterations. The results are validated using 10×5 -cross-validation in 80/20 basis for portioning/testing. The present work uses the LIBSVM[30] package, for implementing the classification task involving multiclass gestures.

IMPLEMENTATION AND EVALUATION

The implementations of the gesture recognition methodology and evaluation scheme of the results are briefly explained in this section. The accuracy estimation mainly considers selected feature of CIFE and reports the results using SVM. The proposed idea will be validate using a popular sign gesture database, SignFi[31].

Implementation on SignFi Dataset

The sign gesture database used in this study[31] acquire CSI values for predefined gesture in lab and home environment for 276 signs with 20/10 instances for each sign (single user). Hence, this data set looks promising for multi-class gesture recognition task. Through extensive experimentation, 8,200 gestures instances acquired from the lab and home environment from one user for 276 gestures; 5,520 from the lab with dimension of $13\text{m} \times 13\text{m}$ and 2,760 from the home environment with dimension of $4.11\text{m} \times 3.86\text{m}$. Also, SignFi[31] acquired data in lab environment for 150 sign gestures with 7500 instances of data for each gesture (5 users/multiple). Data acquisition done with 3 transmitting antennas ($N_t = 3$), and 1 receiving antenna ($N_r = 1$) operating at 5 GHz. CSI values can be recorded with the specialized tool (802.11n CSI tool)[32] and stores the signal information for first 30 subcarriers ($N_s = 30$).

Bispectrum Feature Extraction

The process of extracting the bispectrum as feature vector is shown in Fig. 4. A contour plot of the bispectrum feature will be made between two frequencies f_1 and f_2 . To obtain accurate estimates of the bispectrum, large quantity of data is needed, which makes it difficult to estimate higher order polyspectra. Hence, to quantify the extent of signal phase coupling a normalized third order bispectrum, called bicoherence is used by varying values from 0 to 1.

Bicoherence is plotted between independent frequency axes k and l and the frequency plane bring out the bicoherence magnitude. Considering the symmetric nature of the bicoherence plot, a segment of it will be normally considered for estimation of the features. Bicoherence measures the skewness of the stochastic wideband signal. Also, it captures the non-linearity that present in deterministic signals with the presence of Quadratic Phase Coupling (QPC) as evidence. The bicoherence plot shows peak at the areas of significant bicoherence magnitude indicating QPC between frequency components at the frequencies k , l and $k+l$, which indicate the QPC between frequency components at the triplets.

Proposed study extracts 3 bispectrum features two different environment with single and multiple users data acquisition, and extracts feature vector of size $F_n = 270$ ($3 \times N_t \times N_s = 3 \times 3 \times 30$). Present study extracts the same number of feature vector from different dataset to show the impact of interference in data acquisition environment on recognition accuracy. However, the total extracted features were processed by a CIFE selection algorithm, to filter the best feature set, reducing the computational complexity. Feature selection section explains in detail about the selection criteria.

Fig. 5 shows the contour plot of the bispectrum features between two frequencies f_1 and f_2 , that reveals the basic symmetry of the gesture samples. The contour plot of the Bispectrum Distribution of the CSI signal of third channel and the tenth subcarrier from the lab dataset shows different gestures: (a) the first sample labelled as gesture 'FINISH', (b) the second sample labelled as gesture 'GO', (c) the fifth sample labelled as gesture 'LOVE', (d) the ninth sample labelled as gesture 'NOT'. The bispectrum feature plots are symmetrical along the diagonal of the plot. This can be clearly observed from Fig. 5(a), where the adjacent triangular regions (I – IV) exhibit clear symmetric behaviour along the diagonal lines. Due to this symmetry, one region (say I) will be used to plot the bicoherence and the bispectrum features are estimated from the areas of significant bicoherence magnitude and reported.

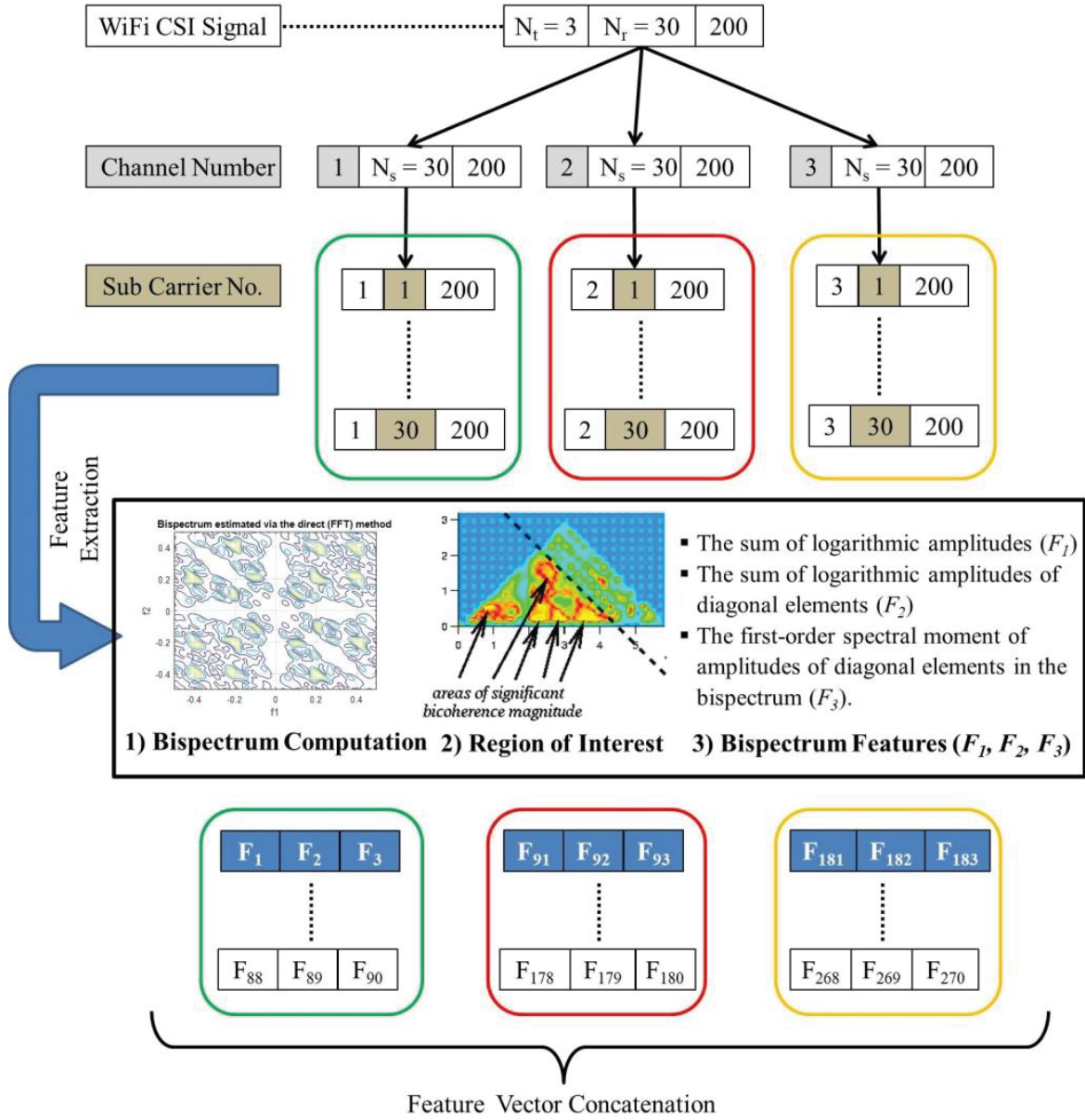


FIGURE 4. Process of extracting bispectrum features

CIFE Feature Selection

CIFE[13] encapsulate the conditionally informative mutual features to maximize the relevancy and reduce the redundancy among the feature pair. The uncertainty present in the extracted set of features will be measured computing the entropy, $E(B)$ of the bispectrum features. Therefore, the mutual information among the feature set is represented as

$$I(B; C) = E(B) - E(B|C) \quad (3)$$

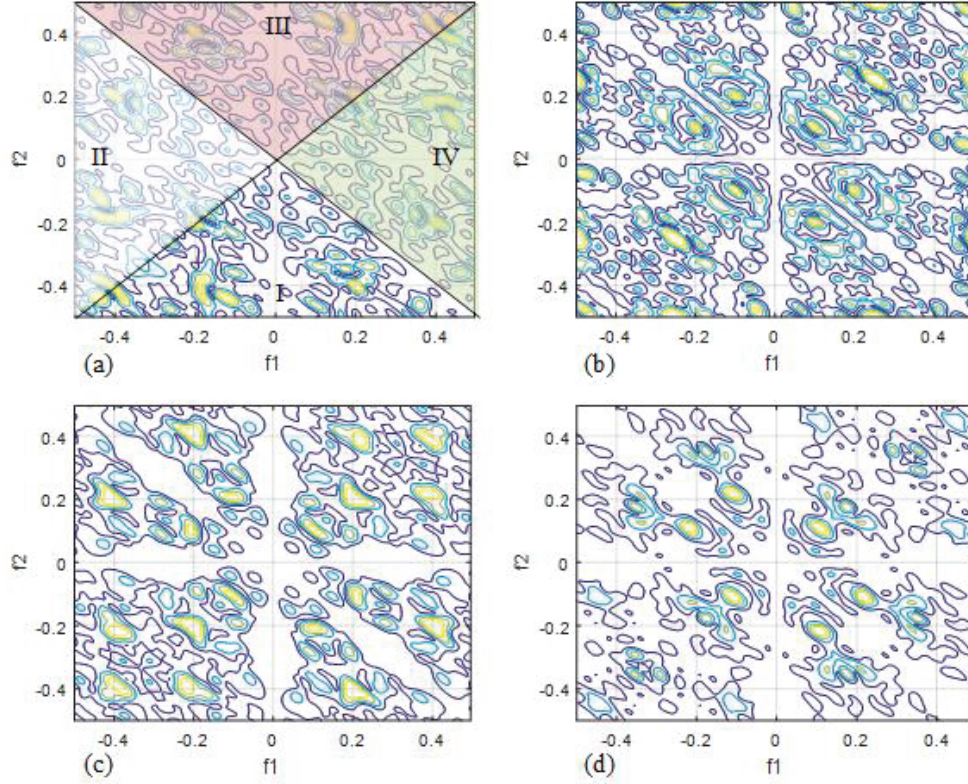


FIGURE 5. The contour plot of the bispectrum distribution of the CSI signal

Eq. 3 indicates that the uncertainty of the gesture class will be reducing with the evidence of the extracted set of bispectrum features. In the proposed study, the algorithm computes the joint class information by assigning $k_s=90$ as explained in Equation (2).

$$I(B; C) = \sum_{t=1}^n I(B(t); C) - \sum_{u=1}^{t-1} R_C(B(u); C(t)) \quad (4)$$

Eq. 4[13] extracts the BF sequentially given that $(t-1)$ BF are extracted, the t^{th} feature can be extracted by optimizing the conditional informative objective as,

$$\theta_t = \underset{\theta_t}{\operatorname{argmax}} \{I(B^{(t)}; C)\} - \sum_{u=1}^{t-1} R_C(B(u); B(t)) \quad (5)$$

where θ , the t^{th} BF parameter and Eq. 5 derives the conditional informative set of BF[13].

HGR Performance Evaluation

The present study adopts SVM Gaussian Kernel for classification and evaluates the recognition performance. The classification performance was evaluated using CIFE by varying the learning parameters (C and γ) with values: 0.003, 0.01, 0.03, 0.1, 0.3, 1, 3, 10, 30, 100 and 300. The values mentioned are supplied in sequence for each iteration ($11 \times 11 = 121$), using 10×5 -cross-validation [33]. The recognition accuracy reported for different environments for 276 and 150 signs with CIFE ($k_s = 90$ features) shown in Table 1. The value for k_s chosen to be 90 features based on experimental trials and found the accuracy declines if less than 90 features are selected. In case, more than 90 features are selected then the accuracy seems more stable but leads to increase in the computational time. The accuracies reported by the SignFi adopting CNN with and without signal processing for datasets from different environments shows Third order bispectrum features show highest gesture recognition accuracy of 94.1% for dataset from *home* environment for 276 gestures. *Lab* environment and *lab + home* environment dataset for the same 276 gestures showed recognition accuracy of 83.8% and 74.9% respectively. Higher recognition accuracy was observed for *home* environment compared to other environments, as it contains comparatively lesser gesture samples.

A recognition accuracy of 75.6% was observed for *lab* environment with 150 gestures and 7500 gesture samples, obtained from 5 users. It is observed that the recognition accuracy declines with increase in number of users or with higher gesture samples. The recognition accuracy can be improved by performing various experimental trails on different datasets to find the optimized parameter that return best result. Also, the volume of data attribute towards the performance. Therefore, performing more experimental trial for optimizing the parameter with other feature selection algorithms using bispectrum features will be of interest in future work.

TABLE 1. Comparison of HGR accuracy with literature and present work

| Dataset (Gestures) | Number of Gesture Samples | Overall Accuracy | | |
|--------------------|---------------------------|------------------|------|------|
| | | (1) | (2) | (3) |
| Lab (276) | 5520 | 95.7 | 98.0 | 83.8 |
| Home (276) | 2760 | 93.9 | 98.9 | 94.1 |
| Lab + Home (276) | 8280 | 92.2 | 94.8 | 74.9 |
| Lab (150) | 7500 | | 86.6 | 75.6 |

(1) SignFi + CNN without signal processing[31]; (2) SignFi + CNN with signal processing[31];
(3) **Present work:** SignFi + 3rd HOS Bispectrum Features + SVM

The training and testing time consumed per instance of gesture enacted in different environment (lab, home, lab+home) shown in Fig. 6. The time take for testing is higher than the training in all the environments. Lab environment takes more time for training and testing compared to home environment.

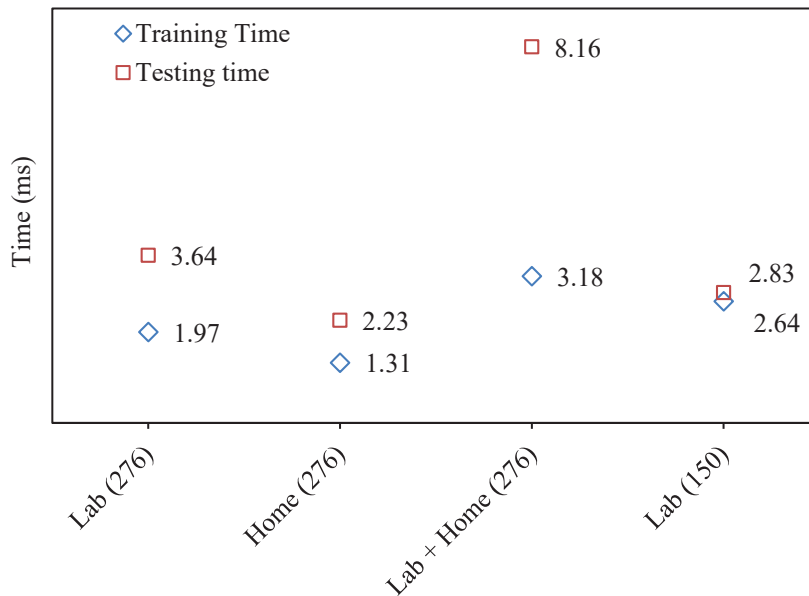


FIGURE 6. Time consumed per instance of gesture in different environments

CONCLUSION

The present work adopts HOS measure for estimating human gesture recognition accuracy using SVM classifier. Third order statistical measure is chosen for feature extraction as it is easy to compute. CIFE, feature selection algorithm is introduced following the feature extraction as it reduces the computational complexity of the SVM. SVM with bispectrum features and CIFE reports 83.8%, 94.1%, and 74.9% in *lab*, *home*, and *lab + home* environment with one user. Accuracy of 75.6% is reported for *lab* environment for 150 gestures with 5 different users. This work can be extended by studying the influence of number of gesture samples in the dataset and number of feature input to the classifier. Furthermore, using cumulant features and other feature selection algorithms could be tested for selecting optimal parameters to obtain highest gesture recognition accuracy for different datasets.

ACKNOWLEDGMENTS

This work was supported by Taylor's University through its TAYLOR'S PhD SCHOLARSHIP Programme. The authors of this article would like to acknowledge the authors of SignFi dataset for providing Wi-Fi CSI traces, labels, and videos of the 276 sign words in GitHub. <https://github.com/yongsen/SignFi>

REFERENCES

1. R. Saini, P. Kumar, P. P. Roy, and D. P. Dogra, "A novel framework of continuous human-activity recognition using Kinect," *Neurocomputing*, Article vol. 311, pp. 99-111, 2018.
2. V. Ranga, N. Yadav, and P. Garg, "American sign language fingerspelling using hybrid discrete wavelet transform-gabor filter and convolutional neural network," *Journal of Engineering Science and Technology*, vol. 13, no. 9, pp. 2655-2669, 2018.
3. G. A. Rao and P. Kishore, "Selfie sign language recognition with multiple features on adaboost multilabel multiclass classifier," *Journal of Engineering Science and Technology*, vol. 13, no. 8, pp. 2352-2368, 2018.
4. B. Kellogg, V. Talla, and S. Gollakota, "Bringing gesture recognition to all devices," in *Proceedings of the 11th USENIX Conference on Networked Systems Design and Implementation*, Seattle, WA, 2014, pp. 303-316: USENIX Association.
5. M. Zhao, F. Adib, and D. Katabi, "Emotion recognition using wireless signals," in *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*, 2016, pp. 95-108: ACM.
6. S. P. Rana, M. Dey, M. Ghavami, and S. Dudley, "Signature Inspired Home Environments Monitoring System Using IR-UWB Technology," *Sensors*, vol. 19, no. 2, p. 385, 2019.
7. W. Li, B. Tan, and R. J. Piechocki, "WiFi-based passive sensing system for human presence and activity event classification," *IET Wireless Sensor Systems*, Article vol. 8, no. 6, pp. 276-283, 2018.
8. Y. Zeng, P. H. Pathak, and P. Mohapatra, "Analyzing shopper's behavior through wifi signals," in *Proceedings of the 2nd workshop on Workshop on Physical Analytics*, Florence, Italy, 2015, pp. 13-18: ACM.
9. R. Chereshevnev and A. Kertész-Farkas, "GaIn: Human gait inference for lower limbic prostheses for patients suffering from double trans-femoral amputation," *Sensors (Switzerland)*, Article vol. 18, no. 12, 2018, Art. no. 4146.
10. J. Yang, H. Zou, H. Jiang, and L. Xie, "Device-Free Occupant Activity Sensing Using WiFi-Enabled IoT Devices for Smart Homes," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3991-4002, 2018.
11. H. Lee, C. R. Ahn, N. Choi, T. Kim, and H. Lee, "The Effects of Housing Environments on the Performance of Activity-Recognition Systems Using Wi-Fi Channel State Information: An Exploratory Study," *Sensors*, vol. 19, no. 5, 2019.
12. P. Mirowski *et al.*, "Probability kernel regression for WiFi localisation," *Journal of Location Based Services*, vol. 6, no. 2, pp. 81-100, 2012.
13. D. Lin and X. Tang, "Conditional infomax learning: an integrated framework for feature extraction and fusion," in *European Conference on Computer Vision*, 2006, pp. 68-82: Springer.
14. R. Zhou, X. Lu, P. Zhao, and J. Chen, "Device-free presence detection and localization with SVM and CSI fingerprinting," *IEEE Sensors Journal*, vol. 17, no. 23, pp. 7990-7999, 2017.
15. F. Hong, X. Wang, Y. Yang, Y. Zong, Y. Zhang, and Z. Guo, "WFID: Passive device-free human identification using WiFi signal," in *Proceedings of the 13th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, Hiroshima, Japan, 2016, pp. 47-56, New York, NY, USA: ACM, 2016.
16. H. Zhu, F. Xiao, L. Sun, R. Wang, and P. Yang, "R-TTWD: Robust device-free through-the-wall detection of moving human with WiFi," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 5, pp. 1090-1103, 2017.
17. S. Palipana, D. Rojas, P. Agrawal, and D. Pesch, "FallDeFi: Ubiquitous fall detection using commodity Wi-Fi devices," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 4, 2018.
18. J. Zhao, L. Liu, Z. Wei, C. Zhang, W. Wang, and Y. Fan, "R-DEHM: CSI-Based Robust Duration Estimation of Human Motion with WiFi," *Sensors*, vol. 19, no. 6, 2019.
19. W. Jia, H. Peng, N. Ruan, Z. Tang, and W. Zhao, "WiFind: Driver fatigue detection with fine-grained Wi-Fi signal features," *IEEE Transactions on Big Data*, 2018.
20. L. Gong, W. Yang, D. Man, G. Dong, M. Yu, and J. Lv, "WiFi-based real-time calibration-free passive human motion detection," *Sensors*, vol. 15, no. 12, pp. 32213-32229, 2015.

21. L. Gong *et al.*, "An adaptive wireless passive human detection via fine-grained physical layer information," *Ad Hoc Networks*, vol. 38, pp. 38-50, 2016.
22. Q. Gao, J. Wang, X. Ma, X. Feng, and H. Wang, "CSI-based device-free wireless localization and activity recognition using radio image features," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 10346-10356, 2017.
23. W. Wang, Y. Chen, and Q. Zhang, "Privacy-preserving location authentication in Wi-Fi networks using fine-grained physical layer signatures," *IEEE Transactions on Wireless Communications*, vol. 15, no. 2, pp. 1218-1225, 2016.
24. Y. Wang, X. Jiang, R. Cao, and X. Wang, "Robust indoor human activity recognition using wireless signals," *Sensors*, vol. 15, no. 7, pp. 17195-17208, 2015.
25. G. Wang, Y. Zou, Z. Zhou, K. Wu, and L. M. Ni, "We can hear you with wi-fi!," *IEEE Transactions on Mobile Computing*, vol. 15, no. 11, pp. 2907-2920, 2016.
26. J. Zhang, B. Wei, W. Hu, and S. S. Kanhere, "Wifi-id: Human identification using wifi signal," in *2016 International Conference on Distributed Computing in Sensor Systems (DCOSS)*, 2016, pp. 75-82: IEEE.
27. A. Swami, J. M. Mendel, and C. L. Nikias, "Higher order spectral analysis toolbox, for use with MATLAB, The MathWorks," 1998.
28. M. E. Hossain, W. A. Jassim, and M. S. Zilany, "Reference-free assessment of speech intelligibility using bispectrum of an auditory neurogram," *PloS one*, vol. 11, no. 3, p. e0150415, 2016.
29. G. Brown, A. Pocock, M.-J. Zhao, and M. Luján, "Conditional likelihood maximisation: a unifying framework for information theoretic feature selection," *Journal of machine learning research*, vol. 13, no. Jan, pp. 27-66, 2012.
30. C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1-27, 2011.
31. Y. Ma, G. Zhou, S. Wang, H. Zhao, and W. Jung, "Signfi: Sign language recognition using wifi," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 1, pp. 23:1--23:21, 2018.
32. D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Linux 802.11 n CSI tool," *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 1, 2010.
33. H. Farhana Thariq Ahmed, H. Ahmad, S. K. Phang, C. A. Vaithilingam, H. Harkat, and K. Narasingamurthi, "Higher Order Feature Extraction and Selection for Robust Human Gesture Recognition using CSI of COTS Wi-Fi Devices," *Sensors*, vol. 19, no. 13, 2019.