Multimodal hand gesture recognition using single IMU and acoustic measurements at wrist

Nabeel Siddiqui, Rosa H. M. Chan *

Department of Electrical Engineering, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong, China

* rosachan@cityu.edu.hk

Abstract

To facilitate hand gesture recognition, we investigated the use of acoustic signals with an accelerometer and gyroscope at the human wrist. As a proof-of-concept, the prototype consisted of 10 microphone units in contact with the skin placed around the wrist along with an inertial measurement unit (IMU). The gesture recognition performance was evaluated through the identification of 13 gestures used in daily life. The optimal area for acoustic sensor placement at the wrist was examined using the minimum redundancy and maximum relevance feature selection algorithm. We recruited 10 subjects to perform over 10 trials for each set of hand gestures. The accuracy was 75% for a general model with the top 25 features selected, and the intra-subject average classification accuracy was over 80% with the same features using one microphone unit at the mid-anterior wrist and an IMU. These results indicate that acoustic signatures from the human wrist can aid IMU sensing for hand gesture recognition, and the selection of a few common features for all subjects could help with building a general model. The proposed multimodal framework helps address the single IMU sensing bottleneck for hand gestures during arm movement and/or locomotion.

Introduction

Human fingers are one of the main means of interaction with the world and are an essential body part in the study of gesture recognition technologies in the field of human–computer interaction (HCI). Gesture recognition technology allows humans to interact with a remote system without physical contact. Many types of sensors have been used in wireless systems for gesture recognition including cameras, sensor gloves, and muscle-based gadgets [1]. The conventional data glove records finger orientation by measuring flexion, vision-based devices commonly use a camera with a depth-based sensor, and muscle activity-based apparatuses record muscular contractions, known as surface electromyography (sEMG). The recordings of these time-series measurements of hand gestures are linked to some instruction in a computer. Despite the high accuracy (>90%) of each of these modalities, there are some limitations in their long-term usage [2,3]. For instance, wearing the data glove for longer durations is

inconvenient, and having a camera positioned throughout daily life to track fingers is impractical. Nonetheless, we anticipate a practical gesture recognition system could be integrated with a smartwatch, the most widely used wearable wrist gadget. Finger movements are linked with the physical operation of tendons, bones, and ligaments at the wrist [4–7]. Some studies have shown that an optical sensor [4], accelerometer and gyroscope [5,6], an array of barometric sensors [8], and a combination of sEMG, accelerometer, and gyroscope [9] around the wrist could be used for hand gesture recognition at the wrist. An inertial measurement unit (IMU), containing an accelerometer and a gyroscope, is a lightweight tiny chip that can be easily placed over the wrist [6]. However, many studies have shown that an IMU sensor placed over the wrist in skin contact could be insufficient as a sole modality for hand gesture recognition, especially during arm movements [9]. If the IMU is not placed in skin contact, it may not detect information that could be recorded from the vibration over the skin while making a gesture. Many researchers have used IMU with other modalities to increase the accuracy of gesture recognition, e.g., IMU with sEMG [9].

In this study, we aimed to use pressure-based sensing with IMU at the wrist using pressure-based sensing units. Pressure-based transducers are more suitable where body movement is unavoidable [10]. Using microphones as pressure-based sensors, we proposed that relevant data could be collected at the wrist for hand gestures using microphones in our previous study [11]. This setup is also supported by another recent study [12]. The microphone signal, or acoustic measurement, is related to the vibrations generated over the skin due to the lateral movement of tissues inside the skin during any mechanical action. Microphones are used as pressure-based sensors in monitoring muscular data, where skin displacement causes a pressure change in the air chamber. Since the chamber is attached to the sound port of the microphone, the pressure change displaces the diaphragm of the microphone, which generates an electrical signal. This skin displacement is due to the low-frequency oscillations produced from muscle contractions [10,13,14]. Acoustic measurements taken from the muscle belly are commonly known as mechano-myography, sound myography, and acoustic myography [15–17].

Our approach of using multimodal sensors (microphones and IMU) at the wrist does not require specific muscle(s) for sensor placement, unlike sEMG [18]. However, specific area over the wrist which would provide significant information to support the use of acoustic sensors were yet to be explored. Therefore, a band of 10 acoustic sensors was created and spread around the wrist with an IMU attached to the band to record limb movement, as shown in Fig 1. Unlike some of the previous studies, the forearm is not fixed; therefore, the IMU is allowed to record every natural movement of the limb during gesture recording. Our main objective was to investigate if a subset of microphone units out of 10 microphone units could be used with an IMU to increase classification accuracy without confining natural limb movements.

First, we recorded 10 acoustic channels and 6 channels of IMU data (10 microphones, three-axis accelerometer, and three-axis gyroscope) from the wrists of 10 subjects. These subjects performed 10 trials for each of the 13 daily life gestures: hand lift, hands up, thumbs up/down, single/double tap, hand/finger swipe, okay sign, victory sign, and fist. Over 7000 features were then extracted from each channel, and the features most relevant to gesture recognition were selected using a mutual information-based algorithm. The results support a multimodal sensor approach using a single microphone unit placed at mid-anterior wrist along with an IMU. Using only two sensors (an IMU and a microphone unit at the mid-anterior wrist) with 25 features, the average accuracy for the general model was 75% for the given gesture set, whereas the average intra-subject classification accuracy exceeded 80% using the same features. These results can be achieved using either a support vector machine (SVM) or linear discriminant analysis (LDA) classifier. Another analysis was also performed using the
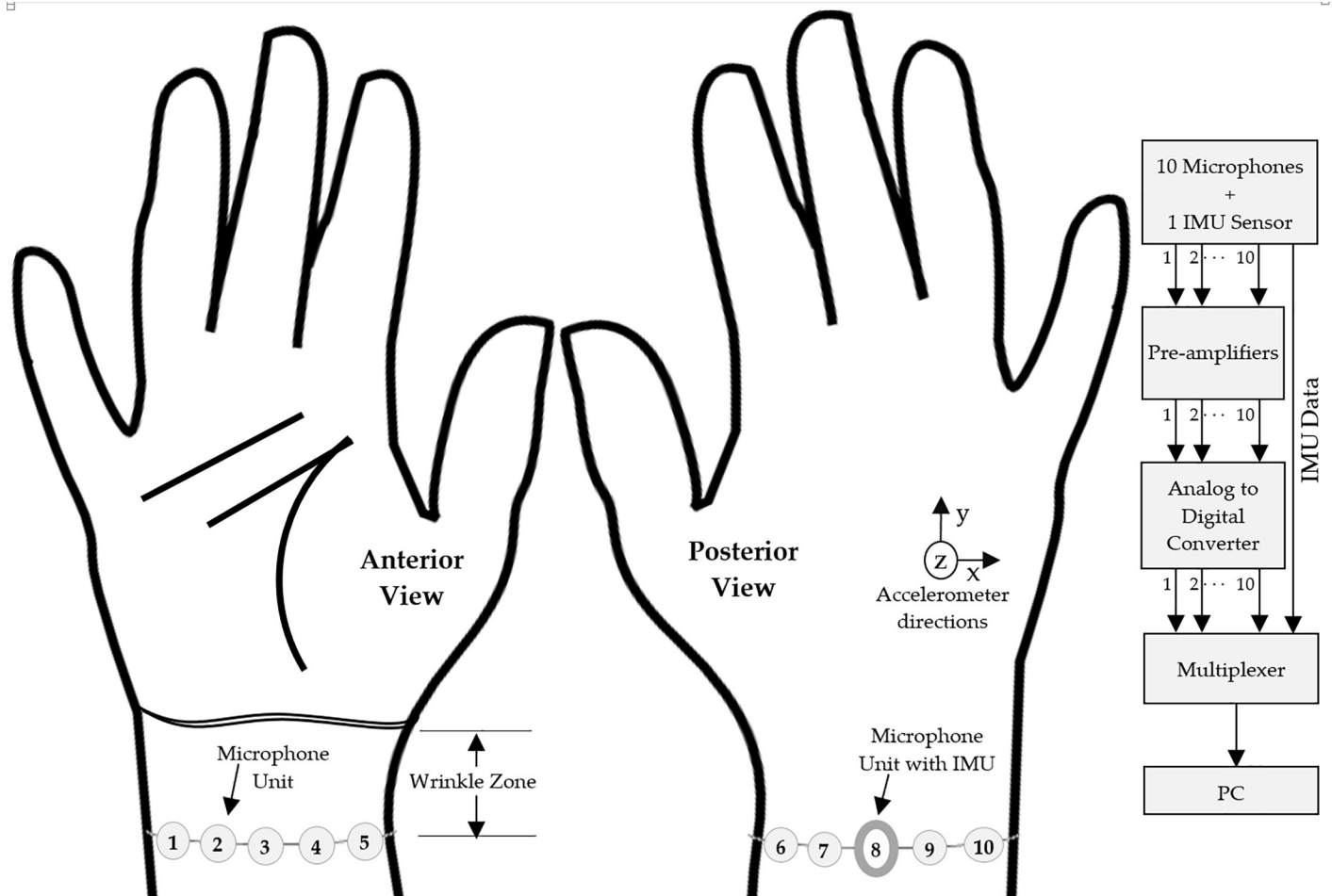
Fig 1. Illustration of hardware prototype. Sensor arrangement on the right hand's anterior and posterior sides is shown at the left and the signal acquisition process is shown at the right.

data from IMU only, which showed that using a microphone with an IMU increased accuracy by around 5% compared with using IMU alone in the general model.

Data acquisition system setup

Fig 1 illustrates the design of our hardware prototype, including the placement of all 10 microphone units. Each of these units had a bottom-ported microphone (f-3dB = 6 Hz; Model: ICS40300, InvenSense, San Jose, CA, USA) that was soldered to a custom designed circuit board (FR-4, 1.6 mm thickness). The cylindrical hole through the PCB acted as a hollow chamber between the microphone diaphragm and the skin surface. This setup is shown in Fig 2A and is similar to the one shown in Kim et al. [19]. An IMU (Model: LSM9DS0, STMicroelectronics, Geneva, Switzerland) was placed over microphone unit #8, which was positioned over the middle of exterior wrist, or, in other words, at the top of the entire sensor arrangement over the wrist, as illustrated in Fig 1. Our prototype utilized only one IMU to record limb movement and it was not in skin contact with the wrist. The procedure was adopted from Repnik et al. [20]. In a preliminary study, we used a single band of five microphones to validate the performance of hand gesture recognition from acoustic recordings [10]. In this study, we examined the multimodal functioning of our custom designed microphone units with an

IMU. The distance between adjacent microphone units in a sensor band varies due to variation in the circumference of wrist amongst subjects, whereas the placement distance of sensor band from the hand over the wrist is discussed in detail in the next section.

As shown on the right side of Fig 1, the microphone units were connected to 10 pre-amplifier (preamp) units in parallel. Each preamp had an op-amp (Model: OPA344, Dallas, TX, USA) with a 67 dB gain arrangement and a lower cutoff frequency of 4 Hz. The pre-amplified acoustic signals and the signals from the IMU were fed to the recording computer via a data acquisition device (Model: Arduino DUE, Somerville, MA, USA). The sampling frequency of each acoustic channel was set to 200 Hz, because previous studies reported that the most significant bio-acoustic information occurs below 100 Hz [21]. The sampling rates for accelerometer and gyroscope were 100 Hz and 50 Hz respectively. Arduino was programmed at 200 Hz per channel sampling for recording every analog/digital signal. Since the sampling rates of accelerometer and gyroscope were different, therefore, these IMU signals were down-sampled in MATLAB to match their operating frequencies.

Experimental setup

Microphone unit placement

Before the start of the experiment in the first session, each subject was asked to form the smallest possible angle between the wrist and the palm of their right arm by the hand as much as possible. This practice enables a distance measurement from the distal bracelet line under which we observed maximum skin displacement (wrinkles) at the wrist with naked eyes. We call this area wrinkle zone as indicated in Fig 1, where the placement of microphone units is avoided to reduce the sensor dislocation during hand motion. Table 1 provides wrinkle zone measurements of the subjects.

In the start of each experimental session, we marked the wrinkle zone length on the wrist. We used this marking for placement of the sensor band at a consistent distance from the hand during all the experiment sessions with each subject. Next, we placed the sensors as indicated in Fig 1. Here, each microphone was placed equidistant from adjacent microphones in the band.

Data collection

Ten right-handed subjects, 5 men and 5 women, participated in the study. Participants provided their informed consent to the experimental procedure, which was approved by the ethics
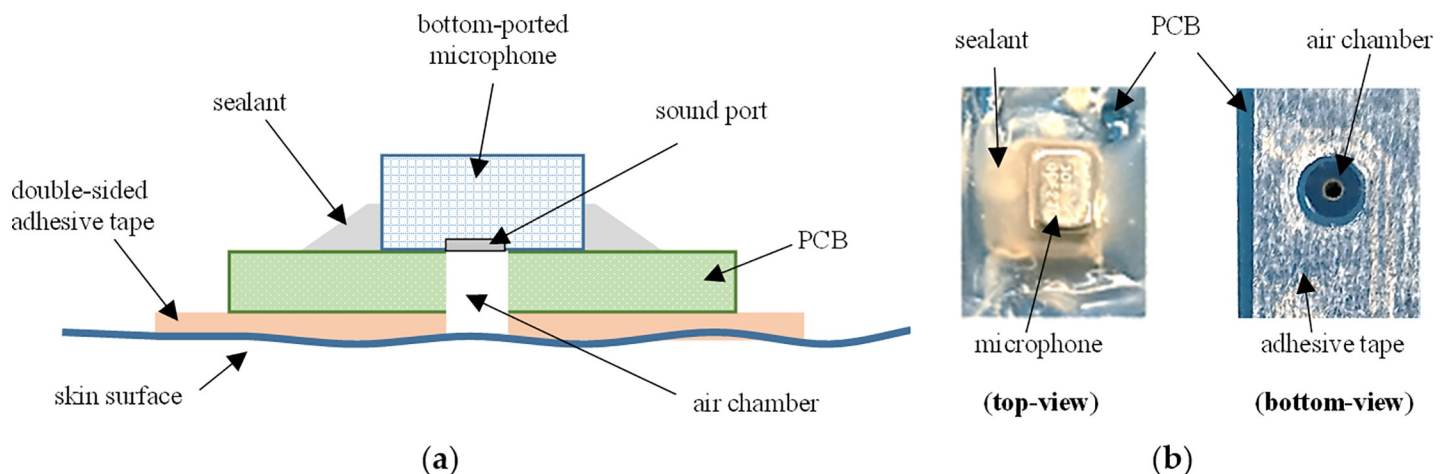
Fig 2. Microphone sensor unit placed on skin surface with double-sided adhesive tape. (a) Schematic cross-sectional side-view; (b) real-life shots.

Table 1. Details of subjects.

Subject	Sex	Age	Wrinkle Zone Length (cm)	Forearm Length (cm)
1	Male	26	3.2	29
2	Female	22	1.8	25
3	Male	29	3.5	26
4	Male	26	3.0	25
5	Female	24	2.0	26
6	Female	23	2.1	27
7	Female	22	2.5	30
8	Male	24	3.3	29
9	Female	27	2.5	23
10	Male	30	2.3	28

https://doi.org/10.1371/journal.pone.0227039.t001

committee at the City University of Hong Kong. The hand gestures under investigation included the 13 common gestures shown in Fig 3, and 1 additional gesture corresponding to the relaxed stage, where the hand is relaxed during the recording, as shown in Fig 3N.

Each subject practiced the gesture before the start of gesture recordings. During the recording session, the subjects were asked to rest their right forearm and elbow on a chair armrest. The position of hand was maintained parallel to the armrest with the palm facing down. They were allowed to make any natural movement of the forearm when they felt it was necessary for the gesture, e.g., in case of thumbs down when there could be a natural twist at the elbow for some of the subjects. Each trial started from the hand in a relaxed position with the palm down and ended at the same resting position. Hence, every channel recording shows the waveform in the action sequence of rest-activity-rest-activity-rest, where the first activity refers to the transformation of hand from the relaxed state to the gesture state and the later activity is back to the relaxed state at the natural pace of the subject.

Each subject performed 10 trials for each gesture. The experiments were conducted in two to three sessions to avoid muscle fatigue, and the duration for each session was no longer than 30 minutes. After sensors placement, an average of 60 trials were recorded in each session. The sensors band was dislodged from the limb at the end of every session and reattached in the
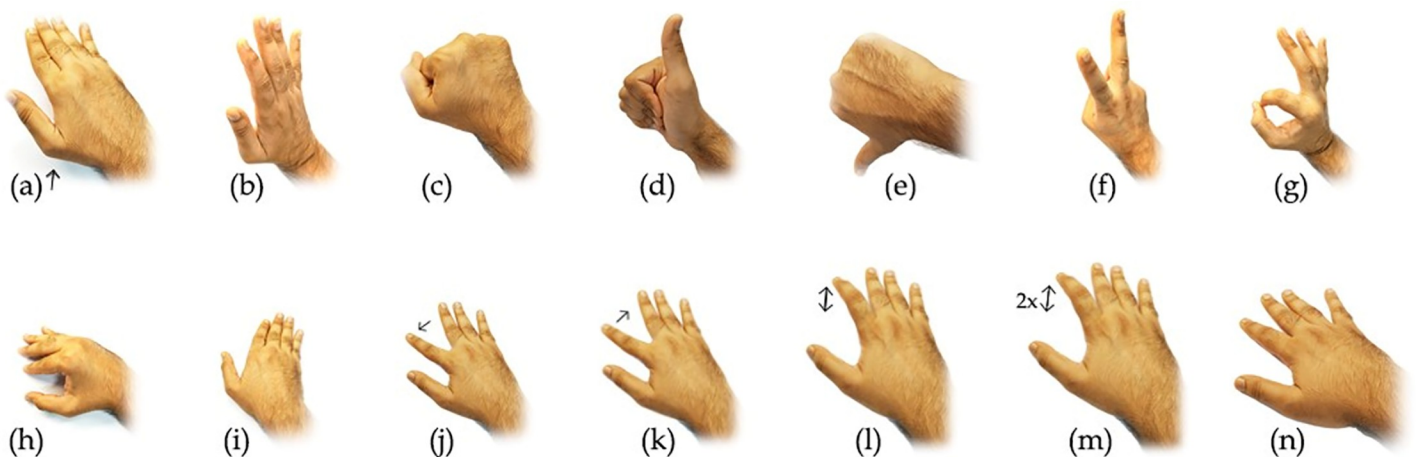
Fig 3. 14 hand gestures used in the experiment. (a) Hand lift, (b) hand up, (c) fist, (d) thumbs up, (e) thumbs down, (f) victory, (g) okay sign, (h) hand swipe left, (i) hand swipe right, (j) index finger swipe left, (k) index finger swipe right, (l) index finger single tap, (m) index finger double tap, and (n) relaxed.

https://doi.org/10.1371/journal.pone.0227039.g003

subsequent sessions. Therefore, all the 140 trials were recorded from each subject over 2–3 experiment sessions. The data collected allowed us to investigate the reliability of hand gesture recognition across sessions using an acoustic modality.

Data analysis

Dataset

The dataset for each subject was a collection of 140 trials (14 gestures × 10 trials), where each trial contained parallel acoustic recordings from 10 microphones, along with motion recordings from the IMU. Each trial spans over three seconds.

Feature extraction

M_i represents the microphone $i = 1, 2,. . ., 10$; and A and G represent the accelerometer and gyroscope with independent values in each specific direction x, y, and z, respectively, for subject n. Each trial (tr) has 16 independent sensor readings at each time point. A total of 7873 features were extracted using the highly comparative time-series analysis code toolbox in MATLAB [22] from each of these sensor readings, as shown in Eq (1)

$$\boldsymbol{F} = [\boldsymbol{f}_{tr,1}, \boldsymbol{f}_{tr,2}, \cdots, \boldsymbol{f}_{tr,7873}]_{tr=1,2,\cdots,140} \tag{1}$$

where $\boldsymbol{F}$ is the feature matrix with 140 trials and 7873 calculations per trial of any time-series signal. Using $\boldsymbol{F}$ for each of the sensors readings, we formed the following cascaded feature matrix:

$$\boldsymbol{S}_n = [\boldsymbol{F}_{M_1}, \boldsymbol{F}_{M_2}, \cdots, \boldsymbol{F}_{M_{10}}, \ \boldsymbol{F}_{A_x}, \boldsymbol{F}_{A_y}, \boldsymbol{F}_{A_z}, \ \boldsymbol{F}_{G_x}, \boldsymbol{F}_{G_y}, \boldsymbol{F}_{G_z}] \tag{2}$$

where $\boldsymbol{S}_n$ represents the feature matrix of a dimension of 140 × 125,968 (16 signals × 7873 features) for subject n. After obtaining $\boldsymbol{S}_n$, feature scaling was performed using standardization (per feature, per channel) for each subject, so that every feature column of $\boldsymbol{S}_n$ had a zero mean and unit variance. This can be demonstrated mathematically:

$$\boldsymbol{S}_n{'} = \frac{\boldsymbol{S}_n - \bar{\boldsymbol{S}}_n}{\boldsymbol{\sigma}_n} \tag{3}$$

where $\bar{\boldsymbol{S}}_n$ and $\boldsymbol{\sigma}_n$ are the mean and standard deviation, respectively, of individual features across all sessions for the subject n.

Feature selection and classification

For the proof-of-concept of the use of multimodal sensors, we employed a generic approach using widely-used feature selection and classification algorithms [23]. Minimum redundancy maximum relevance (mRMR), was applied to the standardized feature matrix obtained from Eq (3). This method ranks the relevance of each feature for classification tasks in large multi-channel, multi-featured datasets. The procedure was adopted from Estevez et al. [24].

After extraction and selection of relevant features, two classifiers were used for classification; a multiclass SVM classifier (using Gaussian kernel) with 10-fold cross-validation, and an LDA classifier with Monte-Carlo cross-validation using 50 runs. In the latter, each run consisted of 112 random training trials (8 trials/gesture) with the remaining 28 trials (2 trials/gesture) acting as the testing set.
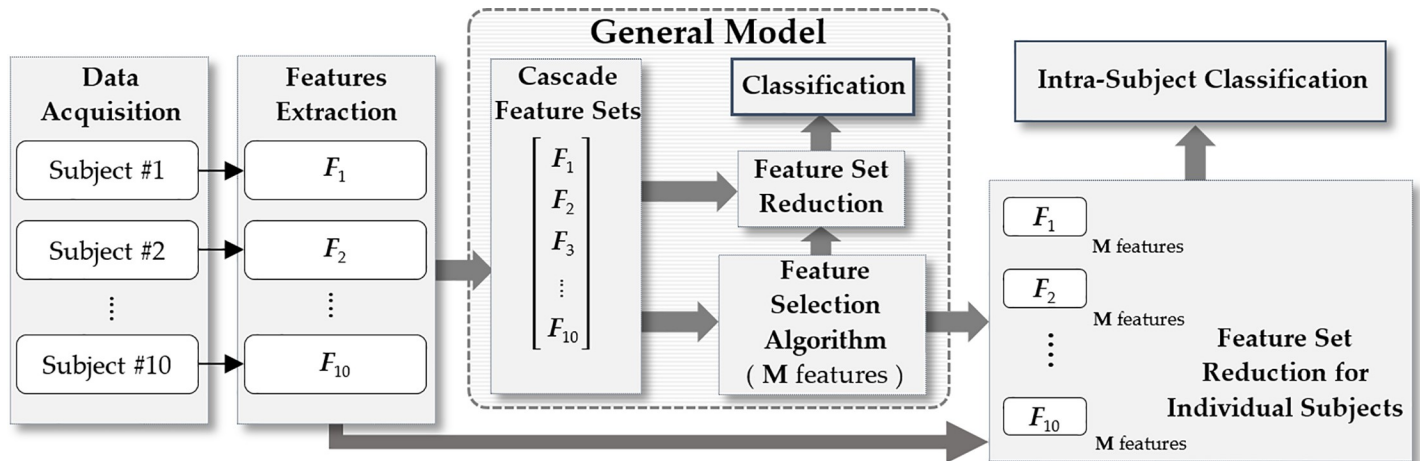
Fig 4. Procedure for data analysis. The first step is feature extraction from individual subjects, then consolidating extracted feature sets from all subjects and applying mRMR to obtain a general model with M features. Then, we applied the same features from the general model on individual subjects to obtain intra-subject classification accuracy.

Results

The overall analysis was conducted in two parts after feature extraction, which is illustrated in Fig 4. First, we created a general model that includes cascaded standardized feature matrices S_n' from all subjects, $n = 1, 2, \ldots, 10$. Next, a mutual information-based feature selection algorithm mRMR was applied for feature selection over the cascaded feature set. The top 25 features that resulted from mRMR were noted, which are defined in S1 Appendix, whereas the rankings and locations of the independent signal sources along with feature names are listed in Table 2. The x, y, and z directions of the accelerometer and placement of microphone sensor unit #3 are depicted in in Fig 1. The x, y, and z directions associated with gyroscope represent the angular velocity around the respective directions. Each feature rank corresponds to a feature computation associated with 1 of the 16 independent sources (10 microphones, 3 accelerometers, and 3 gyroscopes). The two features in Table 2, ranked #9 and #23, were selected from the microphone sensor unit at position #3, whereas the rest of the features were selected from IMU data as the result of mRMR. The microphone sensor unit in position #3 indicates the mid-anterior wrist. Later, SVM and LDA classification methods were applied using the 25 selected features in the cascaded feature matrix, and the results are plotted in Fig 5. This resulted in a 75% accuracy of the general model, notably using acoustic information from the mid-anterior wrist and the data from the IMU chip. The same mRMR feature selection analysis with LDA and SVM was also conducted using the data from the IMU only, and the results are also plotted in Fig 5. The comparison shows an improvement of almost 7% using IMU and microphone #3 compared with IMU-only sensing in the general model with basic feature modeling.

The second part of the analysis was based on using the selected 25 features from the general model for the individual subjects employing the same classification techniques. The results obtained from this analysis are presented in Fig 6, which demonstrates that the features obtained from the general model worked satisfactorily by yielding average accuracy results higher than 80% using the two sensors: the IMU and the microphone placed at position #3. Next, the confusion matrix was formed using the selected features with SVM classification from the individual subjects and averaged across all the subjects, which is shown in Fig 7C. A similar approach was used for the microphones-only and IMU-only cases, whose confusion

Table 2. Feature rankings with the independent signal source.

Feature Rank	Feature	Signal Source		
		Accelerometer	Gyroscope	Microphone
1	Shannon Entropy	.	y	.
2	The separation between the maxima of normal distribution and empirical distribution	.	x	.
3	Minimum Bayesian information criterion	.	y	.
4	Mean of autocorrelation of signal segments	.	y	.
5	Standard deviation based stationarity	y	.	.
6	The maximum mean to mean variance by splitting the signal into x segments	.	y	.
7	Kurtosis	z	.	.
8	Shannon Entropy	x	.	.
9	The separation between the maxima of normal distribution and empirical distribution	.	.	3
10	Pearson Skewness	.	x	.
11	Approximate entropy	.	y	.
12	Minimum Bayesian information criterion	x	.	.
13	Mean of autocorrelation of signal segments	z	.	.
14	Permutation Entropy	.	y	.
15	The highlowmu statistic	.	x	.
16	Burstiness	.	y	.
17	The second-order moment	.	z	.
18	The highlowmu statistic	x	.	.
19	The maximum mean to mean variance by splitting the signal into x segments	z	.	.
20	Mean and Standard deviation based stationarity	.	y	.
21	Proportion of data points within ρ standard deviations of the mean	.	x	.
22	Interquartile range	.	y	.
23	Permutation Entropy	.	.	3
24	Mean of 25% (highest and lowest) trimmed signal	x	.	.
25	Kurtosis	y	.	.

matrices for comparison are shown in Fig 7A and Fig 7B, respectively. These confusion matrices indicate prominent confusion in some of the gestures that are apparently similar. For example, gestures #2, #6, and #7, which are hand stop, victory, and okay sign, respectively, involve a similar major twist at the wrist. Likewise, no major wrist movement occurs when forming the click/tap and double click/tap gestures i.e., gestures #11 and #12, respectively. Fig 7 indicates that microphones could also enhance hand gesture classification accuracy.

Discussion

In this study, we explored the feasibility of using a device with an acoustic sensor with an IMU for multimodal hand gesture recognition. Ten microphone sensor units were placed in skin contact around the wrist in the shape of a band. One IMU was attached to the band for recording limb movement during hand gesture. Ten subjects were recruited, and each performed 13 daily life hand gestures (10 trials/gesture) and a relaxed gesture. A total of 7873 features were used to build a feature vector for each recorded trial from 16 independent sources. To create a general model, the top 25 features were selected using a mutual information algorithm (mRMR) by cascading all individual feature matrices from every subject. Next, these selected feature sets from the general model were used for classification in individual subjects. The intra-subject average accuracy was almost 80%. The confusion matrices indicate difficult in distinguishing similar gestures.
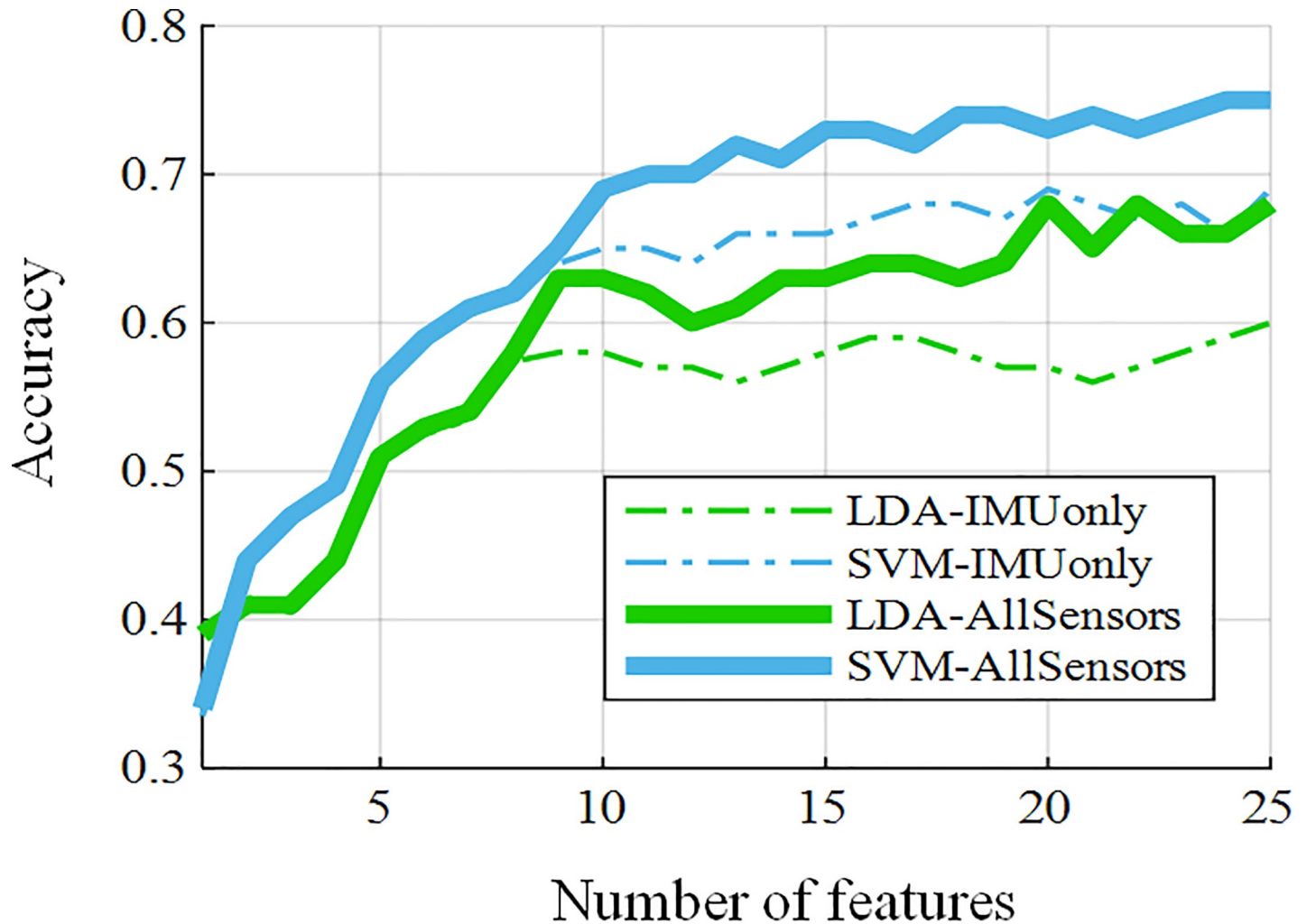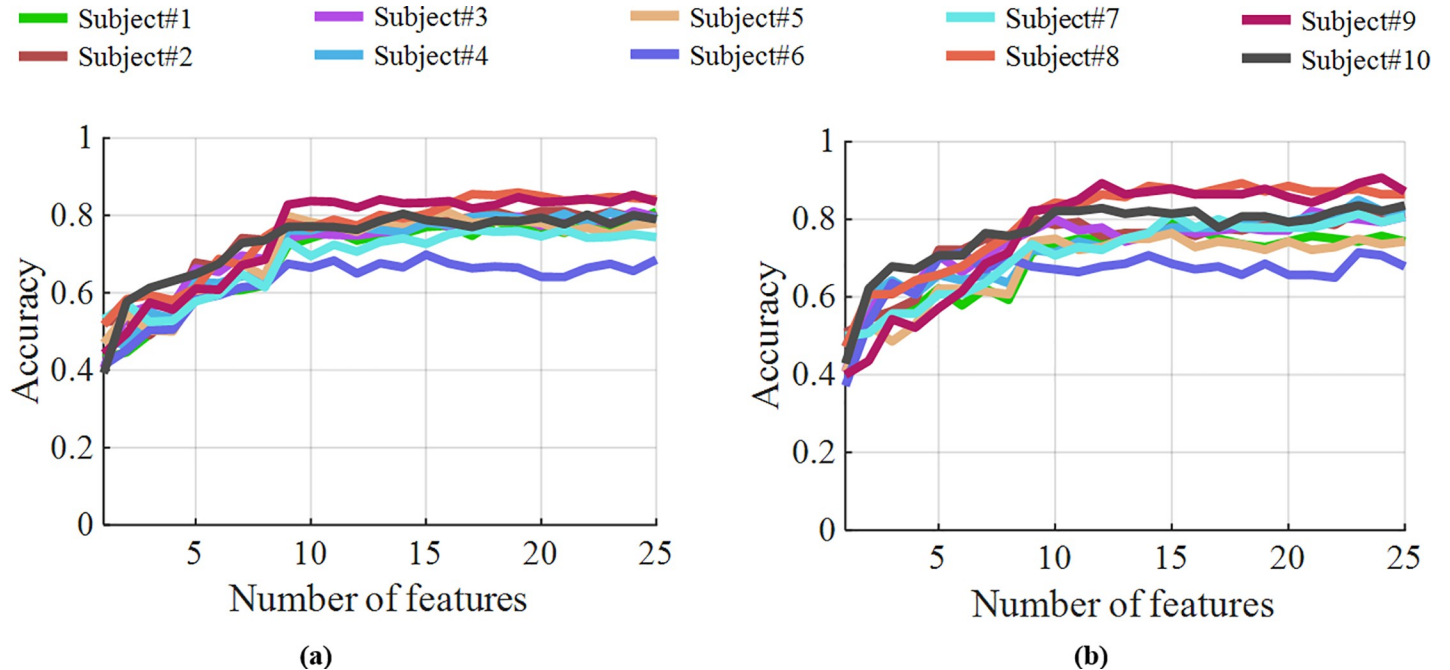
Fig 5. Feature selection with an increasing feature set using concatenated data from all subjects.

https://doi.org/10.1371/journal.pone.0227039.g005

From these 25 features, two features were chosen from the recordings of the microphone sensor unit at the mid-anterior wrist, notably at ranks #9 and #23, whereas the rest of the features were selected from the IMU that was attached to the band at the mid-posterior wrist. These results showed that an extra acoustic sensor(s) positioned at mid-anterior wrist could improve hand gesture recognition accuracy of common wrist wearables with an IMU. The information from microphone units could resolve the signal corruption issue faced by IMU during limb movements because pressure-based sensing only records the data when it records the vibrations over the skin due to hand movement at the wrist; therefore, the microphones are virtually free from the major impacts of limb movement. Future improvements in the design of the microphone sensing units could improve sensing and increase the capability of pressure-based sensors in multimodal systems to recognize gestures.

Subjects were tested on the current gesture set for benchmarking purposes, and more hand gestures will be included in future studies. Our current analysis builds a framework for using IMU with acoustic sensing. Currently, the main constraint with using microphones is the

Fig 6. Feature selection with an increasing feature set for each subject separately using the feature set obtained from the global model. (a) classification using LDA and (b) classification using SVM.

inconvenient coupling mechanism with the wrist that uses double-sided adhesive tape to demonstrate how vibrations measured at wrist during finger movements could facilitate hand gesture recognition. Future work will involve the implementation of a neural network to maximize the accuracy obtained with the dual sensors and better coupling mechanism of acoustic sensors at wrist. We also envision these new findings could facilitate the development of electronic tattoo for human-computer interface [25,26]. This will enable the development of a low-cost wrist band with an IMU and an acoustic sensor(s) that can accurately recognize hand gestures in daily life.
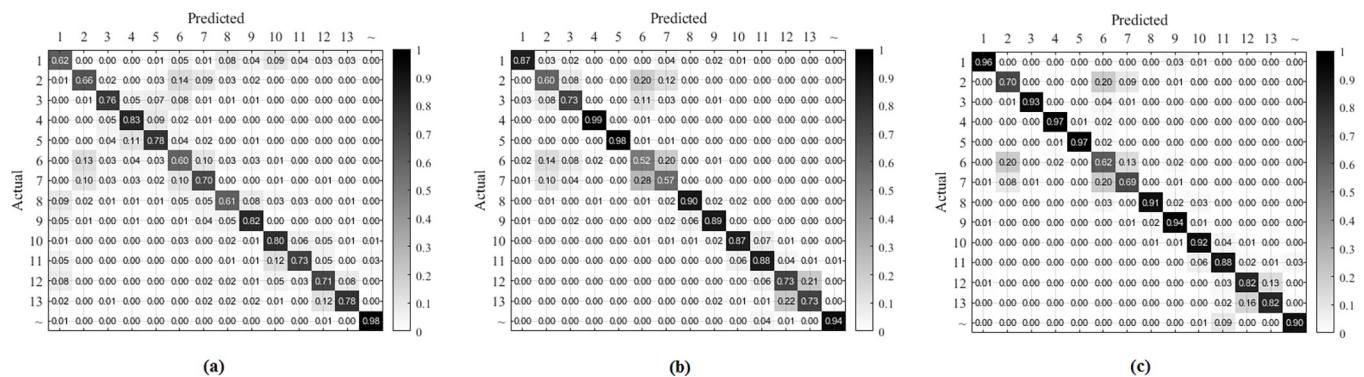
Fig 7. Average confusion matrices using the top 25 feature set obtained from the general model and used for each subject separately with SVM classifier. (a) microphones only; (b) IMU only; (c) all sensors or, more specifically, IMU and Microphone at position #3.

Supporting information

S1 Appendix. Selected features and their parameters.
(DOCX)

Author Contributions

Conceptualization: Nabeel Siddiqui.

Data curation: Nabeel Siddiqui.

Formal analysis: Nabeel Siddiqui.

Funding acquisition: Rosa H. M. Chan.

Investigation: Nabeel Siddiqui.

Methodology: Nabeel Siddiqui.

Project administration: Rosa H. M. Chan.

Software: Nabeel Siddiqui.

Supervision: Rosa H. M. Chan.

Validation: Nabeel Siddiqui.

Visualization: Nabeel Siddiqui.

Writing – original draft: Nabeel Siddiqui.

Writing – review & editing: Rosa H. M. Chan.

References

1. Jiang X, Merhi L-K, Menon C. Force Exertion Affects Grasp Classification Using Force Myography. IEEE Trans Human-Mach Syst. 2018 Apr; 48(2):219–26.

2. Savur C, Sahin F. Real-Time American Sign Language Recognition System Using Surface EMG Signal. In: 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA). Miami, FL, USA: IEEE; 2015. p. 497–502.

3. Galka J, Masior M, Zaborski M, Barczewska K. Inertial Motion Sensing Glove for Sign Language Gesture Acquisition and Recognition. IEEE Sensors J. 2016 Aug; 16(16):6310–6.

4. Jeong-Mook Lim, Dong-Woo Lee, Bae-Sun Kim, Il-Yeon Cho, Jae-Cheol Ryou. Recognizing hand gestures using wrist shapes. In: 2010 Digest of Technical Papers International Conference on Consumer Electronics (ICCE). Las Vegas, NV: IEEE; 2010. p. 197–8.

5. Xu C, Pathak PH, Mohapatra P. Finger-writing with Smartwatch: A Case for Finger and Hand Gesture Recognition using Smartwatch. In: Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications—HotMobile '15. Santa Fe, New Mexico, USA: ACM Press; 2015. p. 9–14.

6. CoolSo. Available from: http://www.coolso.com.tw/ (accessed on 10 August 2019)

7. Zhu Y, Jiang S, Shull PB. Wrist-worn hand gesture recognition based on barometric pressure sensing. In: 2018 IEEE 15th International Conference on Wearable and Implantable Body Sensor Networks (BSN). Las Vegas, NV, USA: IEEE; 2018. p. 181–4.

8. Shull PB, Jiang S, Zhu Y, Zhu X. Hand Gesture Recognition and Finger Angle Estimation via Wrist-Worn Modified Barometric Pressure Sensing. IEEE Trans Neural Syst Rehabil Eng. 2019 Apr; 27 (4):724–32. https://doi.org/10.1109/TNSRE.2019.2905658 PMID: 30892217

9. Jiang S, Lv B, Guo W, Zhang C, Wang H, Sheng X, et al. Feasibility of Wrist-Worn, Real-Time Hand, and Surface Gesture Recognition via sEMG and IMU Sensing. IEEE Trans Ind Inf. 2018 Aug; 14 (8):3376–85.

10. Watakabe M, Mita K, Akataki K, Itoh Y. Mechanical behaviour of condenser microphone in mechanomyography. Med Biol Eng Comput. 2001 Mar; 39(2):195–201. https://doi.org/10.1007/bf02344804 PMID: 11361247

11. Siddiqui N, Chan RHM. A wearable hand gesture recognition device based on acoustic measurements at wrist. In: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) [Internet]. Seogwipo: IEEE; 2017. p. 4443–6.

12. Zhang C, Starner T, Inan O, Abowd GD, Xue Q, Waghmare A, et al. FingerPing: Recognizing Fine-grained Hand Poses using Active Acoustic On-body Sensing. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems—CHI '18. Montreal QC, Canada: ACM Press; 2018. p. 1–10.

13. Beck TW, Housh TJ, Cramer JT, Weir JP, Johnson GO, Coburn JW, et al. Mechanomyographic amplitude and frequency responses during dynamic muscle actions: a comprehensive review. Biomed Eng Online. 2005 Dec 19; 4:67. https://doi.org/10.1186/1475-925X-4-67 PMID: 16364182

14. Orizio C, Liberati D, Locatelli C, De Grandis D, Veicsteinas A. Surface mechanomyogram reflects muscle fibres twitches summation. J Biomech. 1996 Apr; 29(4):475–81. https://doi.org/10.1016/0021-9290(95)00063-1 PMID: 8964777

15. Guo W, Sheng X, Liu H, Zhu X. Mechanomyography Assisted Myoeletric Sensing for Upper-Extremity Prostheses: A Hybrid Approach. IEEE Sensors J. 2017 May 15; 17(10):3100–8.

16. Harrison AP. A more precise, repeatable and diagnostic alternative to surface electromyography—an appraisal of the clinical utility of acoustic myography. Clin Physiol Funct Imaging. 2018 Mar; 38(2):312–25. https://doi.org/10.1111/cpf.12417 PMID: 28251802

17. Orizio C. Soundmyogram and EMG cross-spectrum during exhausting isometric contractions in humans. J Electromyogr Kinesiol. 1992; 2(3):141–9. https://doi.org/10.1016/1050-6411(92)90011-7 PMID: 20719607

18. Hermens HJ, Freriks B, Disselhorst-Klug C, Rau G. Development of recommendations for SEMG sensors and sensor placement procedures. J Electromyogr Kinesiol. 2000 Oct; 10(5):361–74. https://doi.org/10.1016/s1050-6411(00)00027-4 PMID: 11018445

19. Kim T-K, Shimomura Y, Iwanaga K, Katsuura T. Influence of force tremor on mechanomyographic signals recorded with an accelerometer and a condenser microphone during measurement of agonist and antagonist muscles in voluntary submaximal isometric contractions. J Physiol Anthropol. 2008 Jan; 27(1):33–42. https://doi.org/10.2114/jpa2.27.33 PMID: 18239348

20. Repnik E, Puh U, Goljar N, Munih M, Mihelj M. Using Inertial Measurement Units and Electromyography to Quantify Movement during Action Research Arm Test Execution. Sensors. 2018 Aug 22; 18(9):2767

21. Orizio C, Perini R, Veicsteinas A. Muscular sound and force relationship during isometric contraction in man. Eur J Appl Physiol Occup Physiol. 1989; 58(5):528–33. https://doi.org/10.1007/bf02330708 PMID: 2759079

22. Fulcher BD, Jones NS. hctsa: A Computational Framework for Automated Time-Series Phenotyping Using Massive Feature Extraction. Cell Syst. 2017 22; 5(5):527–531.e3. https://doi.org/10.1016/j.cels.2017.10.001 PMID: 29102608

23. Liu L, Wang S, Hu B, Qiong Q, Wen J, Rosenblum DS. Learning structures of interval-based Bayesian networks in probabilistic generative model for human complex activity recognition. Pattern Recognition. 2018 Sep; 81:545–61.

24. Estevez PA, Tesmer M, Perez CA, Zurada JM. Normalized Mutual Information Feature Selection. IEEE Trans Neural Netw. 2009 Feb; 20(2):189–201. https://doi.org/10.1109/TNN.2008.2005601 PMID: 19150792

25. Liu X, Vega K, Maes P, Paradiso JA. Wearability Factors for Skin Interfaces. In: Proceedings of the 7th Augmented Human International Conference 2016 on—AH '16. Geneva, Switzerland: ACM Press; 2016 p. 1–8.

26. Liu Y, Norton JJS, Qazi R, Zou Z, Ammann KR, Liu H, et al. Epidermal mechano-acoustic sensing electronics for cardiovascular diagnostics and human-machine interfaces. Sci Adv. 2016 Nov; 2(11): e1601185. https://doi.org/10.1126/sciadv.1601185 PMID: 28138529