Journal of Mobile Multimedia, Vol. 12, No. 3&4 (2017) 225–242

EVALUATING THE PLACEMENT OF ARM-WORN DEVICES FOR RECOGNIZING VARIATIONS OF DYNAMIC HAND GESTURES

KATHRIN KEFER CLEMENS HOLZMANN RAINHARD DIETER FINDLING
Department of Mobile Computing, University of Applied Sciences Upper Austria
Softwarepark 11, Hagenberg, 4232, Austria
{*kathrin.kefer*},{*clemens.holzmann*},{*rainhard.findling*}*@fh-hagenberg.at*

Dynamic hand gestures have become increasingly popular as touch-free input modality for interactive systems. There exists a variety of arm-worn devices for the recognition of hand gestures, which differ not only in their capabilities, but also in their positioning on users' arms. These differences in positioning might influence how well gestures are recognized, leading to different gesture recognition accuracies. In this paper, we investigate the effect of device placement on dynamic hand gesture recognition accuracy. We consider devices being strapped to the forearm on two positions: the wrist and below the elbow. These positions represent smart watches being worn on the wrist and devices with EMG sensors for the additional detection of static hand gestures (e.g spreading the fingers) being worn right below the elbow. Our hypothesis is that wrist-worn devices will have better recognition accuracy, caused by higher acceleration values of a bigger action radius of dynamic hand gestures. We conducted a comparative study using an LG G Watch and Thalmic Labs' Myo armband, for which we recorded a total of 12960 gesture samples of eight simple dynamic gestures in three different variants with eight participants. We evaluated a potential difference in gesture recognition accuracies using different feature sets and classifiers. Although recognition accuracies for wrist-worn devices seem higher, the difference is not statistically significant due to substantial variations in accuracy across participants. We thus cannot conclude that different positions of gesture recording devices on the lower arm have significant influence on correctly recognizing arm gestures.

Keywords: Gesture recognition; Hand gestures; Accelerometer; Sensor placement; Arm-worn devices

1 Introduction

Hand gestures are becoming increasingly popular as touch-free input modality, e.g. in cars or for home entertainment systems. However, most gesture recognition systems rely on cameras and can be negatively affected by poor lighting conditions or obstructions, which limit the range in which gestures can be used (cf. [1, 2]). These problems can be avoided by deriving gesture information directly from arm-worn devices, such as smartwatches or armbands. One example of the latter is Thalmic Labs' Myo armband.[a] In contrast to wrist-worn devices like smartwatches or activity trackers, it not only allows detection of dynamic gestures using inertial measurement units, but also static gestures/poses with the hand [3] using electromyography sensors that read electrical signals from the muscles in the forearm. One constraint, however, comes with the placement of such devices. In order to properly read muscle signals from the forearm, they have to be strapped to the forearm just below the elbow. While this

[a]The Thalmic Labs' Myo armband: `https://www.myo.com/`

placement is a necessity for detecting hand poses, we expect it to have a negative effect on the detection of dynamic hand gestures. Because of the lower action radius, accelerations get lower the closer sensing devices are worn to the elbow, which acts as rotary joint.

We therefore investigate a possible negative effect on gesture recognition accuracy when placing an arm-worn device right below the elbow, as this is relevant for interactive systems that use both static and dynamic gestures.

Summarizing, the contributions of our work are:

- We investigate the influence of devices being strapped to either the wrist or the lower arm on acceleration-based gesture recognition. Our hypothesis is that wrist-worn devices will allow higher gesture recognition accuracies than devices worn below the elbow. To evaluate this hypothesis, we built an acceleration based gesture recognition system that combines features and models from previous work in gesture and activity recognition.

- We recorded a total of 12960 gesture samples of eight different dynamic arm gestures that have previously been used for gesture recognition [4, 5, 6, 7] (Fig. 2), from eight participants in three different variants of execution. In a prior paper [8], we have already tested the effect of device placement for the first variant of gestures. As recording devices we use an LG smartwatch[b] and Thamic Labs' Myo armband, which both represent off-the-shelf, state-of-the-art hardware available to consumers, thereby facilitating a realistic estimate of the influence of such devices on gesture recognition accuracy.

- We test our hypothesis on our dataset with different feature sets and classifiers. We thereby find that gesture recognition accuracies seem to be partially higher with wrist-worn devices. But due to large variations in gesture recognition accuracies across individuals, we find no statistically significant difference, hence have to reject our hypothesis.

2 Related Work

In this section we summarize previous work in both acceleration based arm gesture recognition and using different sensor positions on the human body for e.g. activity recognition.

2.1 Gesture Recognition Using Arm-Worn Devices

Much previous research on gesture recognition relies on built-in accelerometers. Devices are often held in hands, such as with Wu et al. [9]. They used a Nintendo Wii controller with built-in accelerometers to distinguish 12 gestures: tilting the hand to the left, right, forward and backward, drawing a circle, square, right-angle, and the characters C, S, V, W, and Z into the air. They recorded a total of 3360 gesture samples and used FDSVM (Frame-based Descriptor and multi-class SVM), a self developed gesture recognition system, to distinguish gestures. FDSVM first represents the collected acceleration data by a frame-based descriptor and extracts discriminative information per frame. Then, a SVM-based multi-class classifier is built for gesture recognition in nonlinear feature space. They compared their approach with four other models (Naive Bayes, J48, DTW, WiiGee) and achieved a recognition rate of 95.21% for the user-dependent, respectively 89.29% for the user-independent evaluation, which outperformed all other models. A similar work is presented by Marasović and Papić [10].

[b] The LG smartwatch: `http://www.lg.com/us/smart-watches/lg-W100-lg-watch`

They used an Android smartphone held in the hand to capture accelerometer readings and distinguish seven gestures: right, down, square, circle, triangle, drawing an L and N into the air. They recorded each gesture from a single participant six times and extract 37 features per sample. By applying Principal Component Analysis (PCA), they reduced their feature set to contain only the 80% most discriminative features, then used these to employ a k-nearest neighbour (kNN) and Linear Discriminant Analysis (LDA) classifier. Their evaluation yielded high recognition accuracies for both models (LDA: 87.6%, kNN: 86.7) – however, due to the limited size of the underlying evaluation data, those results might not be generalizable.

In other previous work, wrist-worn devices were used for gesture recognition, instead of devices held in hands. For example, Rekimoto [11] used a gesture detection unit mounted on users' wrists. They recognized hand gestures by capacitively measuring wrist-shape changes and forearm movements with built-in accelerometers. By using three pairs of capacitance transmitters and receivers placed around the users wrist, their system could clearly distinguish between two hand shapes (making a fist and pointing at something). When using these wrist-shape changes in combination with accelerometer readings, they were able to create a set of gesture commands that use the two hand shapes to separate the gesture commands and then detect the transformation between two consecutive arm positions, like palm up, palm down, palm right, palm left, forearm up, and forearm down, as one command. Porzi et al. [7] proposed a gesture recognition system that used a smartwatch with built-in accelerometers to assist people with visual impairments. They distinguished eight dynamic gestures: moving the hand left, right, up, down, and drawing a circle clockwise, counter clockwise, a square, and an arrow to the right side into the air. To achieve effective user independent gesture recognition they developed their own version of a Global Alignment Kernel (GAK). For evaluation, they used three different Support Vector Machines (SVM), including their own Global Alignment Kernel and the standard implementation of it, and a DTW (Dynamic Time Warping) model as reference. Overall, they achieved up to 92.33% recognition accuracy with their own recognition system.

2.2 Placement of Accelerometers for Gesture Recognition

Different placements of accelerometers on the human body have been investigated mainly in the field of activity recognition. Cleland et al. [13] conducted a study about the optimal placement of accelerometers for detection of everyday activities like walking, sitting, lying, standing and walking up and down stairs. They placed accelerometers on the chest, left wrist, left hip, left foot, left thigh and lower back and tested multiple combinations of them for their recognition accuracy. Therefore, they calculated a total of 26 features, such as mean, standard deviation (SD), energy, kurtosis, and skewness for each axis from a sliding window running over the raw acceleration data and containing 512 of these values. With those features they trained four different classifiers (SVM, J48 decision tree, Naive Bayes, and Neural Network). Using only one accelerometer, they achieved more than 91% accuracy for evaluated accelerometer positions, with the SVM being the best classifier, and the hip being the best position. The simultaneous use of three sensors yielded best results, while using e.g. five or six sensors simultaneously even decreases the activity recognition accuracy, compared to using two or three sensors.

The activity recognition accuracy for different accelerometer positions has also been investigated by Olguin and Pentland [14]. They used a Hidden Markov Model (HMM) to test eight different common activities like walking, performing hand movements, or sitting. For sensor positioning, they used combinations of up to three wireless accelerometers placed on the wrist, the left hip, and the chest. Using only one accelerometer on the chest resulted in an average activity recognition accuracy of 62.45%, which they found to improve by 20% when adding a second accelerometer on the hip or wrist. When using accelerometers at all three positions they achieved the best results (92.13% accuracy).

Another study on the accelerometer placement by Gjoreski et al. [15] distinguished fall detection cases (falling slowly, falling from chair slowly, falling from chair quickly) as well as postures (lying, sitting, standing, on all fours, sit on ground, going down, standing up) which might be recognized subsequently to a previous fall. They used up to four accelerometers on different sensor positions, such as chest, waist, right tight, and right ankle. For recognition, they combined raw acceleration data with a set of 18 extracted attributes. The resulting feature set has been used to train four models (SVM, Naive Bayes, J48 and Random Forest). They found posture recognition with a single sensor on the chest (75%) or waist (77%) causing difficulties in distinguishing sitting and standing or on all fours and lying on the stomach. Recognition accuracies improved remarkably when using two (91.5%), three (97%) or four acceleration sensors (99%). For fall detection, they got similar results as for detecting postures.

2.3 Comparison and Discussion

In previous work in acceleration based gesture recognition devices are either held in hands (cf. [9, 10]) or are limited to sensor position on the wrist (cf. [14, 15]). While those works provide a good overview of different approaches towards dynamic (hand) gesture recognition, none investigates the differences of positioning sensors on either wrist or lower arm, and its subsequent influence on gesture recognition accuracy. However, in terms of gesture similarity, we want to point out the work of Porzi et al. [7], which uses a gesture set very similar to the one we investigate in our approach. Similarly, while previous work in acceleration based activity recognition investigates different sensor positions, few of them consider the wrist as sensor position (cf. [13, 14]). Further, those works are limited to distinguishing activities and do not focus on distinguishing arm or hand gestures. However, those approaches use feature extraction and models that can be used with acceleration based gesture recognition alike (cf. [13, 15]), which thereby can be beneficial to our approach.

Summarizing, to the best of our knowledge, no previous research combines acceleration based gesture recognition with sensors placed on the wrist and the forearm of users (Tab. 1). Consequently, we investigate the influence of sensor placements on the wrist and the forearm to acceleration based gesture recognition accuracy. We further do not rely on potential cumbersome prototypical hardware, but instead use available off-the-shelf hardware (LG G Watch on the wrist and Myo armband strapped to the lower arm just below the elbow joint). This facilitates obtaining a realistic estimate of the impact sensor positioning has with products readily available to consumers.

Table 1. Overview of related work for acceleration based arm gesture and activity recognition.

Work	Device	Placement	Measurement(s)	Recordings	Features	Models
Wu et al. [9]	Nintendo Wii controller	Held in the hand	Acceleration	Gestures (left, right, go, come, circle, square, right-angle, C, S, V, W, Z)	Mean, Energy, Entropy, SD, Correlation between the axes	Frame-based Descriptor and multi-class SVM Naive Bayes, J48, DTW, WiiGee
Marasović and Papić [10]	Smartphone	Held in the hand	Acceleration	Gestures(right, down, square, circle, triangle, L, N)	For each axis: Average, Average Absolute Difference, Variance, Standard Deviation, Root Mean Square, Zero-Crossings, Signal-to-Noise Ratio. Duration, Correlation Coefficient, Acceleration, Average Resultant, Binned Distribution	kNN with PCA LDA
Rekimoto [11]	Gesture recording unit	Wrist	Capacitive wrist-shape changes and acceleration	Gestures (making a fist and pointing in combination with six arm positions: palm up, palm down, palm right, palm left, forearm up, forearm down)	-	-
Porzi et al. [7]	Smartwatch together with a Smartphone	Wrist (Smartwatch)	Acceleration	Gestures (left, right, up, down, circle clockwise, circle counter clockwise, square, arrow to the right)	Global Alignment Kernel	SVM, DTW
Cleland et al. [13]	Accelerometers	Left wrist, chest, left hip, left foot, left thigh, lower back	Acceleration	Activities (walking, jogging, sitting, lying, standing, walking up and down stairs)	For each axis: Mean, Energy, Standard Deviation, Kurtosis, Skewness Over all axes: Average Standard Deviation, Average Mean, Average Skewness, Average Kurtosis, Average Energy Correlations between/within the axes	J48, NB, SVM, NN (Multilayer Perceptron)
Olguin and Pentland [14]	Accelerometers	Right wrist, left hip, chest	Acceleration	Activities (sit down, run, squat, walk, stand, crawl, lay down on the chest, hand movements while standing)	-	Hidden Markov Models (HMMs)
Gjoreski et al. [15]	Accelerometers	Waist, chest right thigh, right ankle	Acceleration	Postures (lying, sitting, standing, on all fours, sit on ground, going down, standing up) and Falls (tripping, falling slowly, falling from chair slowly, falling from chair quickly)	Raw acceleration values, length of the acceleration vector, orientation angles for each axis, movement detection attributes, statistical attributes for each axis and the length of the acceleration vector: mean, root mean square, standard deviation,	Nave Bayes, SVM, J48, Random Forest

3 Implementation Details

In order to be able to investigate our hypothesis using the gesture recognition accuracies for each device position, we build a gesture recognition approach (Fig. 1) to record and label eight simple gestures (Fig. 2): waving left, waving right, waving up, waving down, drawing a square, drawing a triangle, and drawing a circle clockwise and counter-clockwise with the right hand in the air in front of the body. Our gesture recognition approach is split into four parts, which we describe in more detail in the following sections: *sensor data collection*, *preprocessing*, *feature extraction* and *feature normalization*. For sensor data collection, a smartwatch (LG G Watch) and a gesture armband (Thalmic Labs' Myo armband) are chosen. For preprocessing, feature extraction, normalization, and classification, R[c]is chosen due to its comprehensive data analysis capabilities and high amount of available classifiers and filters.

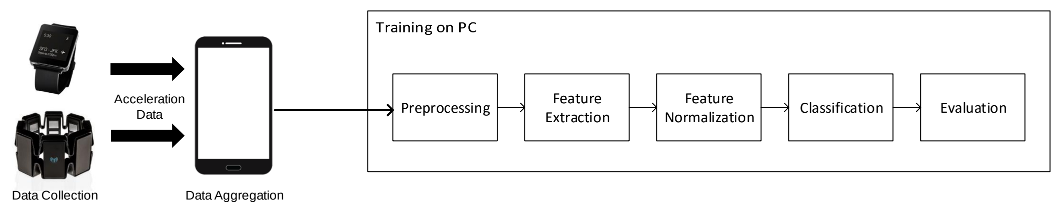


Fig. 1. Overview of our gesture recognition system.

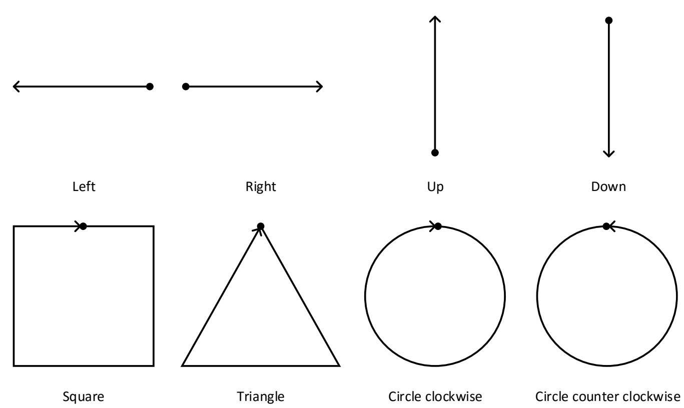


Fig. 2. The eight basic gestures we use for our gesture recognition approach based on previous work [5, 4, 6].

[c]`https://www.r-project.org/`

3.1 Sensor Data Collection

Table 2. The set of features that was used as basis for our evaluation.

No.	Feature	Description
1	Mean Magnitude	The mean over all magnitudes.
2	Minimum Magnitude	The lowest of all magnitudes.
3	Maximum Magnitude	The highest of all magnitudes.
4-6	Range	The acceleration value range for each axis.
7-9	Mean	The mean value of the acceleration values for each axis.
10-12	Mean Absolute Deviation	The mean over all absolute deviations for each axis.
13-15	Root Mean Square	The mean over all squared accelerometer values for each axis.
16-18	Variance	The variance of the accelerometer values for each axis.
19	Mean Variance	The mean over the accelerometer value variances.
20	Magnitude Variance	The variance of the magnitude values.
21-23	Standard Deviation	The standard deviation of the accelerometer values for each axis.
24	Mean Standard Deviation	The mean over the accelerometer value standard deviations.
25	Magnitude Standard Deviation	The standard deviation of the magnitude values.
26-28	Peaks	The number of times an acceleration value is higher than the mean of this axis.
29	Magnitude Peaks	The number of times a magnitude is higher than the Mean Magnitude value.
30	Mean Peaks	The mean over the acceleration value peaks for each axis.
31-33	Mean Peak Values	The mean over all peak values for each axis.
34-36	Correlation Coefficients	The correlation coefficient between two axes.
37-39	Energy	The energy of the accelerometer values for each axis.
40-42	Frequency Range	The frequency range for each axis.
43-45	Mean Frequency	The mean over the frequency values for each axis.
46-48	Frequency Peaks	The number of times a frequency value is higher than the Mean Frequency of that axis.
49	Mean Frequency Peaks	The mean over the frequency peaks for each axis.
50	Gesture Duration	The number of acceleration values for a gesture.
> 51	Accelerometer Data	Every 50th normalized and centred acceleration value. The exact number of acceleration values depends on the gesture length.

To record gestures with the Myo armband and the LG G Watch, an Android application is developed which runs on smartphones with Android version 4.3 or higher. To collect acceleration data from the two input devices, a Bluetooth connection to the smartphone and its application is required. Using the Myo SDK, the connection to the Myo armband is established using Bluetooth 4.0 and the accelerometer data from the Myo is delivered via a callback method in the smartphone application.

For the Android Wear smartwatch, a dedicated application is implemented, which receives a start/stop command for data recording from the smartphone. The collected data is then sent back to the mobile phone application via the established Bluetooth connection. The acceleration data from both devices is collected in a data structure until the gesture is finished, whereupon the raw acceleration data for all three axes and from both devices is stored in two csv-files on the smartphone's internal storage. The Myo armband delivers its data with a fixed sampling frequency of 50 Hz, the sampling frequency of the Android Wear smartwatch was set to 50 Hz in the application, but can slightly vary due to the Android-internal sensor implementation.

3.2 Preprocessing

Preprocessing consists of two steps to a) smooth acceleration data and b) normalize the length of all gestures. As smoothing function, a running median of odd span is used with a width of 11 samples per window, which corresponds to 220 ms.

For the gesture length normalization, the maximum length of all recorded gestures of a device is chosen, to avoid loss of information for single gestures. The acceleration samples of each gesture are interpolated to this maximum length, which is fixed to 303 samples for the Myo armband and 424 samples for the LG G Watch for the first variant of execution, to 243 and 223 samples for the second variant, and to 248 samples for the Myo and 231 samples for LG G Watch for the third variant of the gesture execution.

3.3 Feature Extraction and Selection

Based on several existing works on gesture recognition with three-axis accelerometers [5, 16, 17], a set of features proven to be useful for gesture recognition is chosen. The basic features are calculated from a window that contains all accelerometer data belonging to one single gesture, which means that the window length is 303 respectively 424 for the Myo and the LG G Watch for the first variant of the gestures, 243 respectively 223 for the two devices for the second and 248 respectively 231 for the Myo and the LG G Watch for the third variant of gesture execution.

The feature set with at least 51 features is chosen as the basis for the feature extraction and therefore calculated for each device and setup (Tab. 2). In addition to that, the best feature combination for each device and setup is calculated by applying the recursive feature elimination algorithm provided by R, which tries every combination of the available features and evaluates it with a Random Forest classifier with a 10-fold cross-validation with 10 repetitions. With this approach, we are able to reduce the amount of features from 77 to 45 for the LG G Watch and from 68 to 42 for the Myo armband for the first gesture variant, from 65 to 49 and from 65 to 48 for the second variant of gesture execution and from 65 to 55 for the LG G Watch and from 65 to 63 for the Myo armband for the third variant of the gestures. The features selected for each variant are listed in table 3.

3.4 Feature Normalization

In order to get better results and more accurate recognition accuracies especially for the k-Nearest Neighbour classifier (which is impacted by data that is not normalized), the whole feature set calculated during the feature extraction is additionally centered and scaled to the mean of the data being 0 and the standard deviation being 1.

4 Experiment

4.1 Participants

To test our hypothesis with the developed gesture recognition system described above, we conducted a comparative study with eight voluntary participants (2 female, 6 male). Their ages varied from 22 to 55 years (*mean*=32.25; *SD*=11.95) and all of them were right-handed.

Table 3. The set of the selected features for each device and gesture variant that was used for our evaluation, based on the full feature set listed in table 2.

Gesture Variant	Device	Feature No.
Variant 1	Myo	2-4, 6, 7, 10, 12, 13, 16, 18, 20, 21, 23, 25-28, 30, 31, 34-37, 40, 43, 45, 48-50 and four acceleration readings from the x and z axis starting from the second value and the first five acceleration readings from the y axis.
	Wear	5, 6, 10-12, 14-18, 21-23, 31, 34-36, 38-40, 42, 44-46, 49, 50 and six acceleration reading from the x axis starting from the third value, seven acceleration readings from the y axis and six acceleration readings from the z axis starting from the second value.
Variant 2	Myo	1-4, 6, 7, 10-13, 16-26, 29-31, 34-37, 40, 43-50 and five acceleration readings from the x axis starting from the first value and three acceleration readings from the y and z axis starting from the second value.
	Wear	5-18, 21-23, 26, 29-42, 44-46, 49, 50 and four acceleration readings from each axis starting at the second value.
Variant 3	Myo	1-26, 28-46, 48-50 and the first five acceleration readings of each axis.
	Wear	1, 4-27, 29, 31-42, 44-46, 49, 50 and four acceleration readings from each axis starting at the second value.

None was familiar with gesture recognition or with the eight gestures. Five of the participants were IT-students or programmers, three participants had no IT-background.

4.2 Apparatus

As gesture recording devices, the latest version of Thalmic Lab's Myo armband was used and as Android Wear smartwatch a LG G Watch W100 with Android Wear version 5.1.1 was chosen. For the Android smartphone application, Android Studio 1.4 was used as development environment. A Google Nexus 5x smartphone with Android version 6.0 was used as gesture recording device and for the whole machine learning toolchain, RStudio version 0.99.903 with R version 3.2.3 was used.

4.3 Procedure

The participants were asked to perform the eight simple, dynamic gestures in three different variants with both devices mounted on the same arm simultaneously. At the beginning of each session, the gestures and how they should be performed (starting and end position, palm always facing down) was explained to each participant. For the data collection, the participants were equipped with the Myo armband and the LG G Watch on the right arm, and they were told which gesture to perform next and when to start.

For the first variant of execution (Fig. 3), the participants used the outstretched arm being held orthogonal to the body and parallel to the floor as starting position and also performed the gestures with completely outstretched arm all the time. The second variant of the gestures was performed by holding the upper arm as still as possible on the side of the body and only moving the lower arm from the elbow downwards. For the third variant, the participants were asked to perform the gestures as naturally as possible keeping the palm downwards, which resulted in this approach being a mediocre between the two preceding ones. A comparison of the gesture variants two and three is shown in figure 4. Each gesture of each variant was then trained and performed 30 times by each of the eight participants, which resulted in a total of 2160 gesture samples per variant and device or 12960 samples in total.

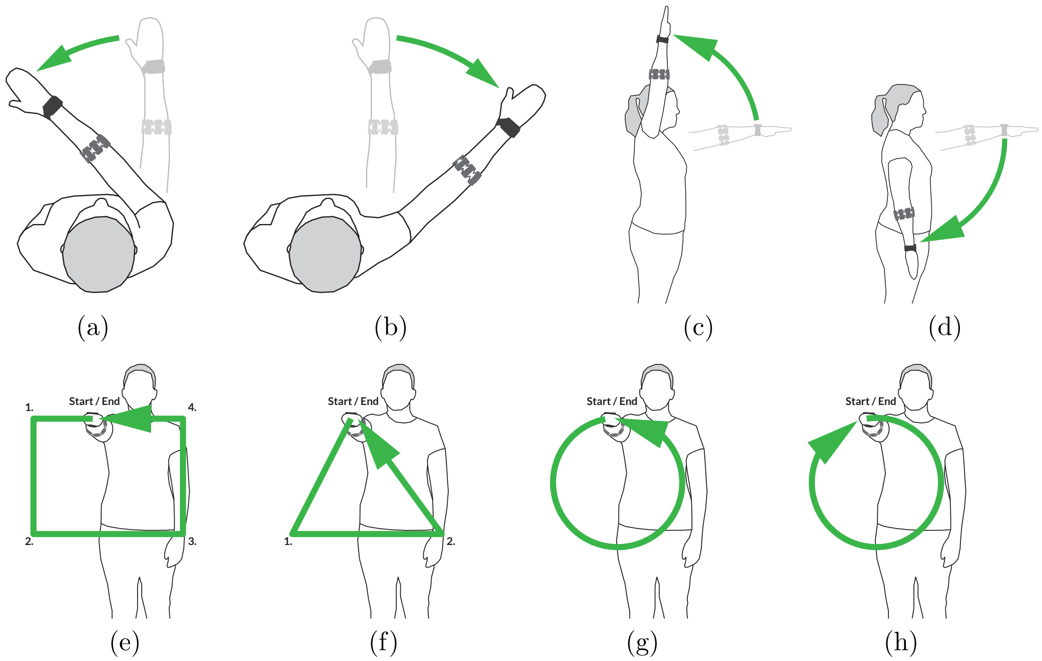


Fig. 3. Visualization of how each gesture was performed for the first variant of execution by the participants: waving (a) left, (b) right, (c) up and (d) down, as well as drawing (e) a square, (f) a triangle, (g) a circle clockwise and (h) a circle counterclockwise.

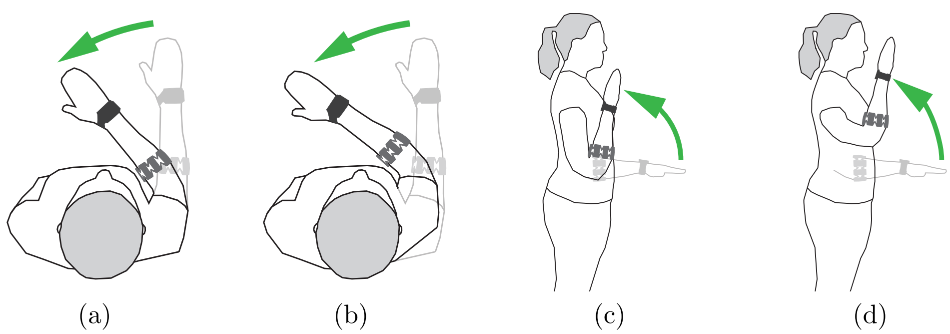


Fig. 4. Visualization of the main differences of the second and third variant of gesture execution. Waving left in variant two (a) is shorter than in variant three (b). Waving up in variant two (c) is only performed with the forearm with the upper arm held as still as possible, whereas in variant three (d), the upper arm is moved upwards too.

4.4 Design

In order to test our hypothesis of the wrist-worn device (LG G Watch) achieving a better recognition accuracy than the one worn below the elbow (Myo armband), we evaluated each variant of the recorded gestures with the evaluation design described below. To calculate the recognition accuracies, we used four different classifiers being k-Nearest Neighbour, J48 decision tree, Random Forest and Naive Bayes. With each classifier, we first evaluated the recognition accuracies using the well known and frequently used 10-fold cross-validation with 10 repetitions. However, this approach lacks in generalizing the results due to using gestures from all participants for training the classifiers, which helps them in adapting to the differences between the participants. Therefore, we secondly conducted a leave-subject-out cross-validation in which a single participant is repeatedly (for each participant) left out from the training data and used for testing only. This approach produces more generalizable results due to testing the trained classifier on the gestures of a person it has never seen before, which would also be the case in real life applications. Moreover, we conducted the training and testing with each classifier with both evaluation approaches for all features (Tab. 2) as well as only for the selected features (Tab. 3).

5 Results

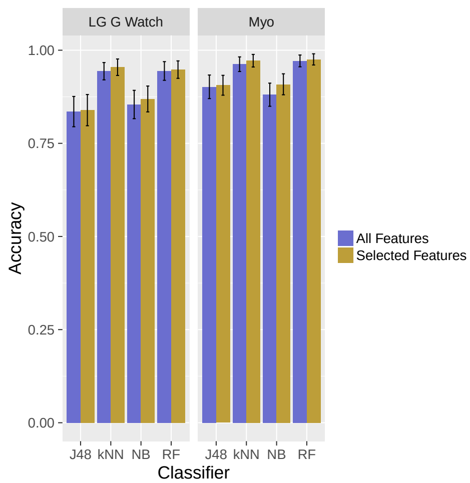
The evaluation and classification of gestures was done with R Studio on a PC. The bar charts (Fig. 5 to 7) show the recognition results for the different gesture variants for each device, classifier, feature set and evaluation type; the error bars represent the standard deviations. In addition to that, the recognition accuracies for each best classifier are shown in table 4. The recognition accuracies for 10-fold cross-validation with 10 repetitions are, as expected, better than for the leave-subject-out cross-validation, and also the deviation of the results is much lower.

Table 4. Recognition accuracies for the best classifier for all gesture variants, devices and feature sets.

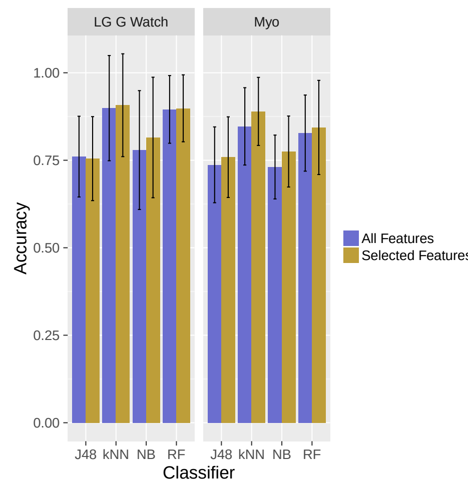
		10-fold cross-validation		Leave-subject-out cross-validation	
Gesture Variant	**Device**	**All Features**	**Selected Features**	**All Features**	**Selected Features**
Variant 1	Myo	97.13%	97.54%	84.69%	88.96%
	Wear	94.43%	95.44%	89.90%	90.73%
Variant 2	Myo	96.58%	97.00%	74.96%	75.80%
	Wear	91.18%	91.94%	80.48%	80.53%
Variant 3	Myo	98.03%	98.17%	81.51%	80.57%
	Wear	91.25%	91.61%	83.33%	84.38%

5.1 Gesture Variant 1

For the first variant of the recorded gestures, the Myo armband achieves between 88.05% and 97.13% mean recognition accuracy for the different classifiers and therefore performs slightly better than the LG G Watch, which achieves between 83.52% and 94.44% with the 10-fold cross-validation and all features. Considering only the selected features (Tab. 3) increases the mean recognition accuracies by approximately 1%, which is a good result for using only about one half of the original features. We also tested the statistical significance for both feature sets (all features and only the selected ones) and the best classifier for each device for the 10-fold

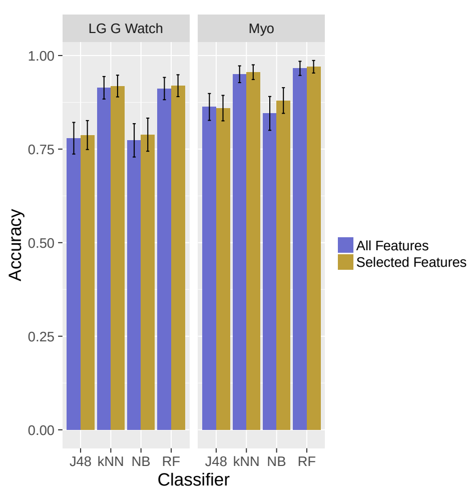


(a)

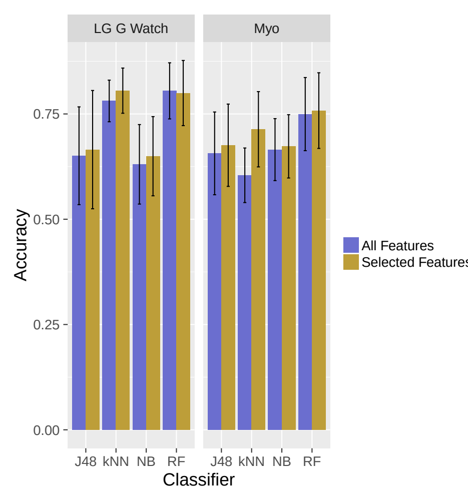


(b)

Fig. 5. The recognition accuracies for the first gesture variant for both devices, each classifier and all features as well as for the selected features. The evaluation was done with (a) 10-fold cross-validation with 10 repetitions and (b) with leave-subject-out cross-validation.

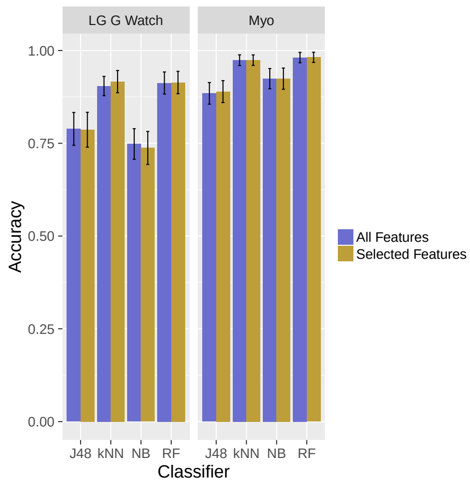


(a)

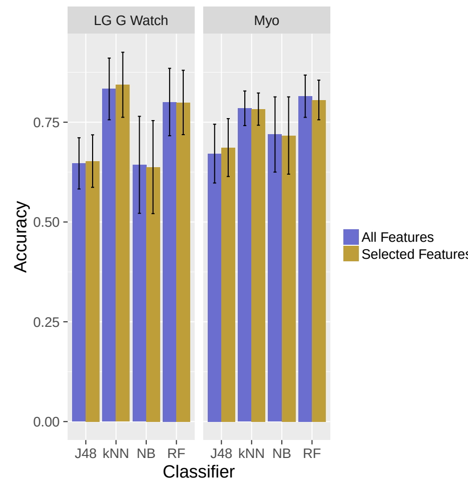


(b)

Fig. 6. The recognition accuracies for the second gesture variant for both devices, each classifier and all features as well as for the selected features. The evaluation was done with (a) 10-fold cross-validation with 10 repetitions and (b) with leave-one-subject-out cross-validation.



(a)



(b)

Fig. 7. The recognition accuracies for the third gesture variant for both devices, each classifier and all features as well as for the selected features. The evaluation was done with (a) 10-fold cross-validation with 10 repetitions and (b) with leave-subject-out cross-validation.

cross-validation. For the setup with all features, the best classifier was the Random Forest for both devices, whereas for the setup with the selected features, the best classifier was the kNN with k=5 for the LG G Watch and the Random Forest for the Myo armband. For both approaches, the Myo armband achieved better recognition accuracies compared to the LG G Watch. The difference between the two devices was proved to be statistically significant (all features: $t99 = 9.6$, $p < 0.001$; selected features: $t99 = 7.9$, $p < 0.001$).

With results for the leave-subject-out cross-validation of the first gesture variant, one can see that the LG G Watch (between 76.04% and 89.89% for all features and 75.74% and 90.73% for the selected features) performs better than the Myo armband (between 73.07% and 84.69% for all features and 75.89% and 88.96% for the selected features) with almost every classifier with both feature sets. However, removing unnecessary features again increased the mean accuracy of the different devices and classifiers (despite one) by at least 0.3% and up to 4.4%. Again, we tested the statistical significance for this type of cross-validation and with both feature sets for the best classifier of each device (kNN with k=5 for the Myo armband and the LG G Watch for all features, kNN with k=9 for the Myo armband and kNN with k=7 for the LG G Watch for the selected features). As there were substantial variations across the participants, the difference between the two devices was not statistically significant with both feature sets for the leave-subject-out cross validation (all features: $t8 = -0.72$, not significant (ns); selected features: $t8 = -0.72$, ns).

Summing up, for the 10-fold cross-validation evaluation approach, we found the Myo armband to perform significantly better than the LG G Watch with both feature sets. However, when using the leave-subject-out cross-validation, no significant difference between the two devices could be found.

5.2 Gesture Variant 2

With the second gesture variant, the Myo armband again performs slightly better than the LG G Watch by achieving between 84.54% and 96.58% mean recognition accuracy compared to accuracies achieved by the LG G Watch, which are between 77.34% and 91.18% with the 10-fold cross-validation and all features. Using only the selected features for the classification increases the mean recognition accuracies for the Myo by up to 3.4% and for the LG G Watch between 0.7% and even 13.9%. Testing the statistical significance for both feature sets and only the best classifier for each device (all features: Random Forest for the Myo armband and kNN with k=5 for the LG G Watch; selected features: Random Forest for both devices), the difference between the two devices was proven to be highly significant (all features: $t99 = 17.19$, $p < 0.001$; selected features: $t99 = 16.36$, $p < 0.001$).

When evaluating the second gesture variant with the leave-subject-out cross-validation with all features, the Myo achieved between 60.44% and 74.96% recognition accuracies whereas the LG G Watch achieved between 63.05% and 80.48%, which is slightly better than the Myo. When using only the selected features for each device, the accuracies could be increased by up to 10.9% for the Myo armband and up to 2.4% by the LG G Watch. When testing the statistical significance for this evaluation approach, we found that there is a significant difference between the two devices when the best classifier (the Random Forest for both, the Myo armband and the LG G Watch) is evaluated with all features ($t99 = -3.24$, $p < 0.05$). However, when using only the selected features and the best classifier for each device, no significant difference between the two devices could be found ($t8 = -1.16$, ns).

Evaluating this gesture variant with the 10-fold cross-validation with 10 repetitions, the Myo armband performed significantly better than the LG G Watch with both feature sets. Besides that, we were also able to prove that the LG G Watch performed significantly better in the leave-subject-out cross-validation approach when used with all features.

5.3 Gesture Variant 3

For the third variant of the recorded gestures when evaluated with the 10-fold cross-validation and all features, the Myo armband achieves between 88.45% and 98.09% mean recognition accuracy for the different classifiers and therefore performs better than the LG G Watch, which achieves between 74.81% and 91.25%. Considering only the selected features increases the mean recognition accuracies only slightly by at most 1%, which is easily explainable due to the fact that only very few features are dismissed during the selection for this gesture variant. For this gesture variant, we again found a highly significant difference between the two devices for the best classifier of the 10-fold cross-validation (the Random Forest for both devices) with all features ($t99 = 21.32$, $p < 0.001$) and also the selected ones ($t99 = 19.88$, $p < 0.001$).

When evaluating the third gesture variant with the leave-subject-out cross-validation, we found that the LG G Watch, which achieved between 64.32% and 83.33% for all features and 63.75% and 84.37% for the selected features, performs better than the Myo armband, which achieved between 67.14% and 81.51% recognition accuracy for all features and 68.65% and 80.57% accuracy for the selected features. Removing the unnecessary features resulted in slightly decreased recognition accuracies (by about 0.1% to 0.9%) for most of the classifiers, but also increased the accuracies for the other classifiers by up to 1.5%. Again, we tested the statistical significance with both feature sets for the best classifier of each device (Random Forest the Myo armband for all and only the selected features, kNN with k=7 for all features and kNN with k=9 for the selected features for the LG G Watch). Also for this variant of the gestures, there were substantial variations across the participants, so the difference between the two devices was not statistically significant with both feature sets (all features: $t8 = -0.52$, ns; selected features: $t8 = -0.98$, ns).

Summing up, we again found the Myo armband to perform significantly better than the LG G Watch with both feature sets when evaluated wit the 10-fold cross validation with 10 repetitions. However, when evaluating with the leave-subject-out cross-validation, no significant difference between the two devices could be found.

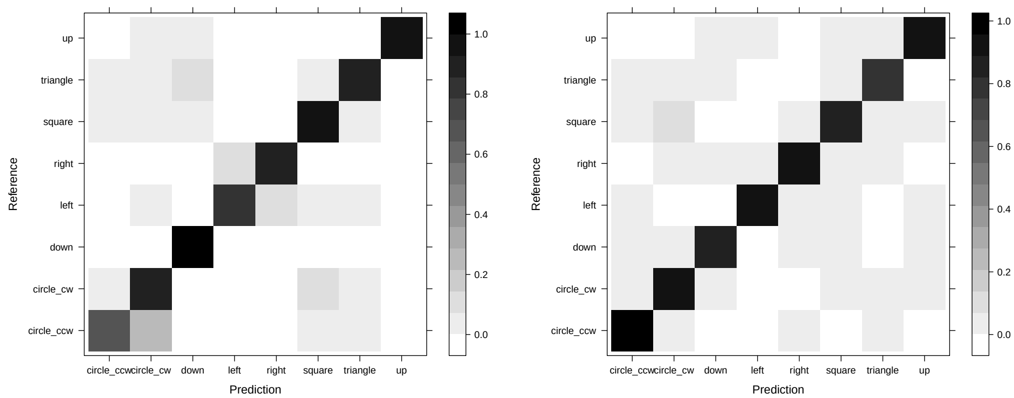
5.4 Discussion

In general, we found that for all three different gesture variants, the Myo armband performed significantly better when evaluated with the 10-fold cross-validation, whereas the LG G Watch performed better in the gallery independent setup with the leave-subject-out cross validation. However, the statistical significance of the watch performing better than the Myo armband in the second evaluation approach could only be proved for one case out of the six presented ones, which is the leave-subject-out cross-validation with all features for the second variant of the recorded gestures. One reason for this result can be found in the evaluation approaches. When using the 10-fold cross-validation for the evaluation, the classifier is tested only with people it has already seen during training. However, when the leave-subject-out cross-validation is used, people that have not been included during the training process and are therefore completely unknown to the classifier, are used to evaluate the classifier. In addition to that, the gesture set we use is quite simple and does not contain any rotational movements, which could also have an influence on the high recognition accuracies but not significant differences between the devices. For this reason, we partly have to reject our hypothesis that there is a statistical significance between the two devices.

Finally, figures 8 to 10 visualize the confusion matrices for the leave-subject-out cross-validation with only the selected features and the best classifiers for each variant of the gestures. One can see that for all three gesture variants, the Myo armband often confused the circle clockwise with the circle counter clockwise and that the LG G Watch performed better, especially in the first variant of the recorded gestures.

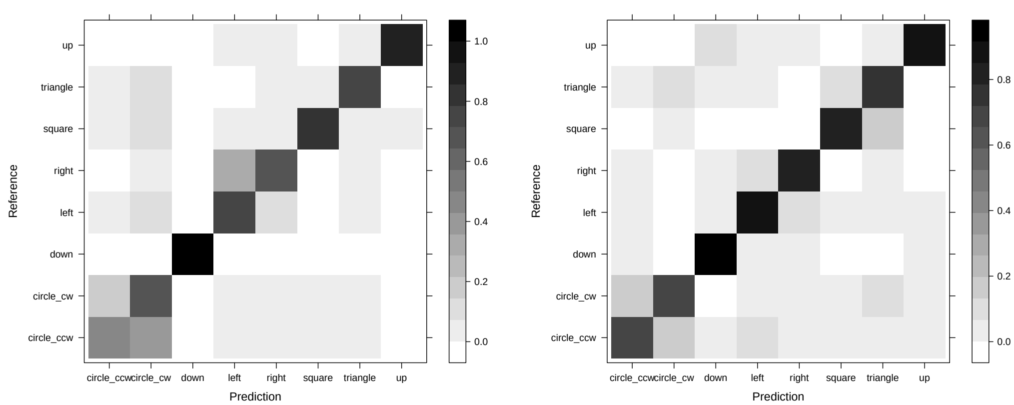
6 Conclusions

In this work, we investigated if gesture recognition accuracies are affected by placing the recording device either at the wrist or below the elbow during performing gestures. We compared the placement of two wearable devices, an LG G Watch worn on the wrist and a Myo armband from Thalmic Labs worn right below the elbow, concerning the recognition of three different variants of dynamic hand gestures. We implemented an Android-based system for simultaneously collecting sensor data from both devices. In a study with eight participants and three different variants of the eight gestures, we collected 2160 gesture samples from each device for each gesture variant. The data was filtered and normalized, features were extracted, relevant features were selected and normalized, and they were classified using k-Nearest Neighbour, J48 decision tree, Random Forest and Naive Bayes classifiers. We



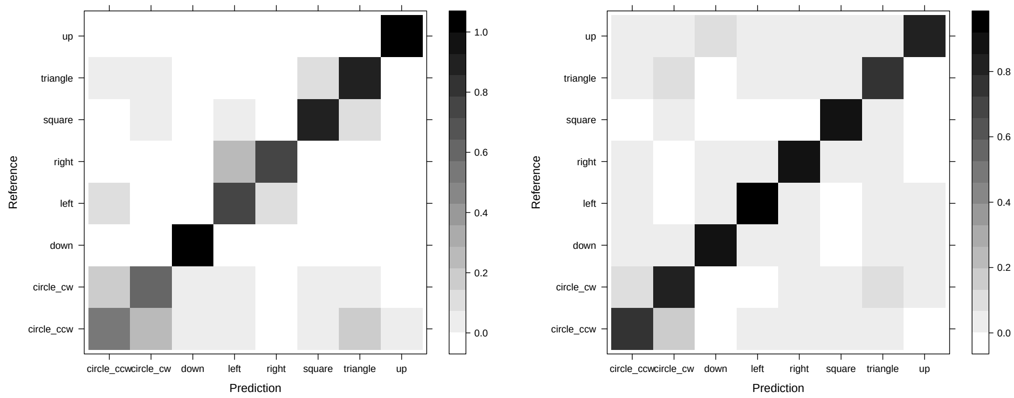
(a) (b)

Fig. 8. The confusion matrices for the first variant of the recorded gestures with leave-subject-out cross-validation for the best classifier of (a) the Myo armband (kNN with k=9) and (b) the LG G Watch (kNN with k=7) with only the selected features.



(a) (b)

Fig. 9. The confusion matrices for the second variant of the recorded gestures with leave-subject-out cross-validation for the best classifier of (a) the Myo armband and (b) the LG G Watch (the Random Forest for both) with only the selected features.



(a) (b)

Fig. 10. The confusion matrices for the third variant of the recorded gestures with leave-subject-out cross-validation for the best classifier of (a) the Myo armband (Random Forest) and (b) the LG G Watch (kNN with k=9) with only the selected features.

evaluated the recognition accuracy with 10-fold cross-validation with 10 repetitions and leave-subject-out cross-validation.

Our assumption was that the wrist-worn device would have a significantly better recognition accuracy due to its bigger action radius. Using the set of selected features only and an evaluation with leave-one-subject-out cross-validation, whose results are more generalizable than those for the 10-fold cross-validation, the recognition accuracies for the LG G Watch were higher than for the Myo armband. Although this supports our hypothesis we had to reject it, as we were not able to prove that the difference between the two devices is significant for all three gesture variants. However, for the second variant of the gestures and the feature set containing all possible features, we found a statistical significance between the two devices.

A possible cause for the non-significance is, that the used gestures – even though we recorded different variants of them – are very simple. Moreover, the gestures were characterized by large arm movements, which is why the acceleration values measured below the elbow were still big enough to distinguish the gestures well. Therefore, an interesting item of future work would be the investigation of recognition accuracies with more complex gestures like less sweeping gestures as well as gestures that include rotations of the forearm.

Acknowledgements

This work has partially been carried out within the scope of the project *InterPlayces* funded by the Austrian Research Promotion Agency (FFG) under contract number 8843 and the scope of *u'smile*, the Josef Ressel Center for User-Friendly Secure Mobile Environments, funded by the Christian Doppler Gesellschaft, A1 Telekom Austria AG, Drei-Banken-EDV GmbH, LG Nexera Business Solutions AG, NXP Semiconductors Austria GmbH, and Österreichische Staatsdruckerei GmbH.

References

1. Lingchen Chen and Feng Wang and Hui Deng and Kaifan Ji (2013), *A Survey on Hand Gesture Recognition*, Proc. of the 2013 International Conference on Computer Sciences and Applications, CSA '13, pp. 313-316
2. J. J. LaViola (1999), *A Survey of Hand Posture and Gesture Recognition Techniques and Technology*, Brown University
3. S. Mitra and T. Acharya (2007), *Gesture Recognition: A Survey*, IEEE T SYST MAN CY C, Vol.37, pp. 311-324
4. J. Kela and P. Korpip and J. Mantyjrvi and S. Kallio and G. Savino and L. Jozzo and D. Marca (2006), *Accelerometer-based Gesture Control for a Design Environment*, Pers. Ubiquit. Comput., Vol.10, pp. 285-299
5. N. Helmi and M. Helmi (2009), *Applying a neuro-fuzzy classifier for gesture-based control using a single wrist-mounted accelerometer*, Proc. of the 2009 IEEE International Symposium on Computational Intelligence in Robotics and Automation, pp. 216-221
6. J. Liu and L. Zhong and J. Wickramasuriya and V. Vasudevan(2009), *uWave: Accelerometer-based personalized gesture recognition and its applications*, Pervasive Mob Comput, Vol.5, pp.657-675
7. L. Porzi and S. Messelodi and C. M. Modena and E. Ricci (2013), *A Smart Watch-based Gesture Recognition System for Assisting People with Visual Impairments*, Proc. of the 3rd ACM International Workshop on Interactive Multimedia on Mobile and Portable Devices, pp. 19-24
8. K. Kefer and C. Holzmann and R.D. Findling (2016), *Comparing the Placement of Two Arm-Worn Devices for Recognizing Dynamic Hand Gestures*, Proc. of the 14th International Conference on Advances in Mobile Computing and Multi Media, pp. 99-104
9. J. Wu and G. Pan and D. Zhang and G. Qi and S. Li (2009), *Gesture Recognition with a 3-D Accelerometer*, Ubiquitous Intelligence and Computing, Vol.5585, pp. 25-38
10. T. Marasovi, V. Papi (2011), *Accelerometer-based gesture classification using principal component analysis*, SoftCOM, pp. 1-5
11. J. Rekimoto (2001), *GestureWrist and GesturePad: Unobtrusive Wearable Interaction Devices*, Proc. of the 5th International Symposium on Wearable Computers
12. X. Chen and T. Grossman and D. J. Wigdor and G. Fitzmaurice (2014), *Duet: Exploring Joint Interactions on a Smart Phone and a Smart Watch*, Proc. of the SIGCHI Conference on Human Factors in Computing Systems, pp. 159-168
13. I. Cleland and B. Kikhia and C. Nugent and A. Boytsov and J. Hallberg and K. Synnes and S. McClean and D. Finlay (2013),*Optimal placement of accelerometers for the detection of everyday activities*, Sensors, Vol.13, pp. 9183-9200
14. D. O. Olguin and A. S. Pentland (2006), *Human activity recognition: Accuracy across common locations for wearable sensors*, Proc. of the 10th IEEE International Symposium on Wearable Computers, pp. 11-13
15. H. Gjoreski and M. Lustrek and M. Gams (2011), *Accelerometer Placement for Posture Recognition and Fall Detection*, Proc. of the 7th International Conference on Intelligent Environments, pp. 47-54
16. T. Marasovic and V. Papic (2011), *Accelerometer-based gesture classification using principal component analysis*, Proc. of the 19th International Conference on Software, Telecommunications and Computer Networks, pp. 1-5
17. Yinghui Zhou and Lei Jing and Junbo Wang and Zixue Cheng (2012), *Analysis and Selection of Features for Gesture Recognition Based on a Micro Wearable Device*, IJACSA, Vol.3