Sign Language Classification Based on EMG using Time Domain Features Extraction

Adhi Dharma Wibawa, Eko Pramunanto and Ifut Rahayuningsih
Department of Computer Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
adhiosa@te.its.ac.id

Abstract: Sign language has been studied by many researchers in the past. However, most of the evaluation techniques were done based on video data, evaluating it using EMG is new approach. Accuracy in performing the sign motion becoming crucial especially when it is needed to be understood by the public. EMG is one of the most advance tools now a days in recording the muscle activity due to a motion. This study was evaluating some motions in Indonesian sign language by using EMG of myoarm sensors. Twenty sign motions were classified by using Naive Bayes to recognize the 20 letters. The 5 participants were involved in this experiment, time domain features such as MAV (Mean Absolute Value), RMS (Root Mean Square), VAR (Variance) and SSI (Simple Square Integral) were used as the input of the classifier. Each motion was repeated 20 times by the participants. The result showed 79% accuracy of the sign language to be recognized. This result was not optimum yet in term of high accuracy. We opinioned that this is due to the variation of motion from one subject when performing one specific task, besides the accuracy in placement the Myo arm sensor's position. Moreover, variation in duration during performing one motion can also be another issue in accuracy calculation.

Key words: Indonesian sign language, electromyography, time domain features extraction, Naive Bayes, Myo arm band, MAV

INTRODUCTION

For daily communication people with disabilities are very dependent on the use of sign language. Not only it is used for communication with each other among them but also it is useful for public communication such as on television, news etc. Coding sign language from human motion into digital information such as text has been explored by many scientists previously (Pandey and Jain, 2015). This transformation is mainly used for sign language training, translation technique or education. Most of those studies have been developed by using computer vision technology such as kinect infrared camera or video processing (Tewari and Srivastava, 2012; Murthy and Jadon., 2009; Pandey and Jain, 2015; Ibraheem and Khan, 2012). In those researchers, their technique was mainly based on camera for recording and recognizing the sign language movement. Some markers (such as color markers or motion markers) have been implemented for recognition purposes (Ibraheem and Khan, 2012). The limitation of this technique is the input was highly dependent on the surrounding environmental condition. Moreover, the lighting conditions of the room often change unpredictably, so, the camera will be difficult to record and recognize the movement precisely. Due to

that reason, evaluating sign language motion by using muscle activity was an option for building a better sign language system recognition. EMG technology now a days has been developed advancely into wireless and wearable devices. Myo armb and is an EMG electrode in the form of bracelet, capable of capturing muscle activity of hand movements.

Previous studies in evaluating sign language using kinect has been done by Yang (2015). Yang has showed 90.4% accuracy of using kinects, however, when recording the small motion such as fingers motion, this system was prone to errors. Some kinect technology even insensitive with the motion of fingers. Other study in evaluating sign language using Myo arm band has been done by Abreu et al. (2016). Abreu has implemented support vector machine algorithm as its clasifier to recognize the static gesture of sign language and the result was not that satisfying. This study is investigating the potential of developing a better sign language system for Indonesian language by using Myo arm band and time domain feature extraction during static gesture. In this experiment, we also explored the motion of fingers and elaborate the detail aspect by using Myo arm band. We hypothesized that the hand motion of Indonesian Sign language can be classified better by using Naive Bayes

---

Corresponding Author: Adhi Dharma Wibawa, Department of Computer Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia, adhiosa@te.its.ac.id

algorithm with higher accuracy. Some analysis regarding the aspects that influence the accuracy result such as the number of features and which features are dominan in this experiment would also be presented.

MATERIALS AND METHODS

Sign language in general is divided into two parts, namely static gesture and dynamic gesture. Static gestures are sign language movements that do not involve changes in hand shape or arm movement. While the dynamic gesture is a movement of sign language that involves changes in the shape of the hand or arm movement. The static and dynamic gesture difference is shown in Fig. 1.

Electromyography (EMG) is a technique for reading and monitoring the activity of bioelectric signals produced by skeletal muscle. Electromyograph will detect the potential of electrical signal activity from muscle fibers during electrical activity. There are several techniques for measuring EMG signals, one of which is using EMG surface. Myo arm band is one of the tools that can be used to measure EMG signal using surface electrode (Wibawa *et al.*, 2016).

Myo arm band is a wearable device developed in Thalmic Labs that uses EMG sensors and combines with IMU (Inertial Measurement Unit) sensors including gyroscopes, accelerometers and magnetometers to recognize motion (Fig. 2). IMU sensor is used as tracking motion and EMG sensor is used for muscle sensing. The EMG sampling rate sensor on Myo arm band is 200 Hz and 50 Hz on the IMU sensor.

Myo arm band has 8 EMG channels that correspond to a specific muscles in the human forearm muscles. The Myo arm band electrode was positioned circularly as Fig. 3. Configuration of Myo electrode is channel 4 which has a blue marker followed by channel 3 clockwise and channel 5 counterclockwise. The major muscles covered are the extensor digitorum and flexor digitorum muscles. The muscle is the muscle that moves the wrist, the index finger, middle finger, ring finger and pinkie.

Time domain: In EMG signal analysis, there are several methods that can be used as feature capture which are: time domain features, frequency domain features and time-frequency domain features. A feature commonly used in EMG signal classification is Time Domain features (TD). The advantage of using TD is that TD is faster to calculate because TD does not require mathematical transformation. TD features are used as follows:
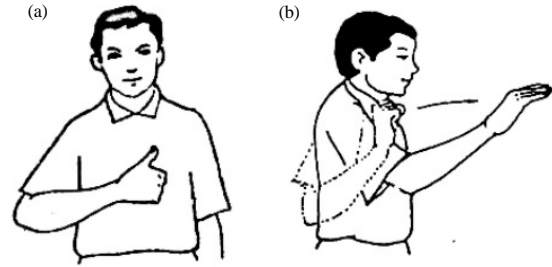

Fig. 1: Static gesture: a) 'Good', dynamic and b) 'Throw'
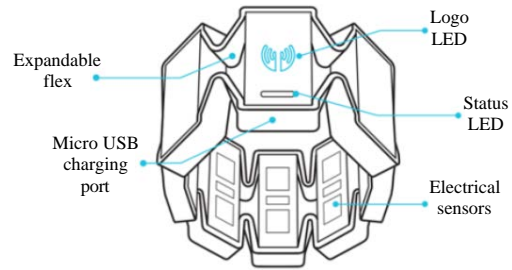

Fig. 2: Myo arm band (Abreu *et al.*, 2016)

Mean Absolute Value (MAV): Is used as an onset detection index especially in EMG surface signals for control of artificial limbs (Irfan *et al.*, 2016). MAV can be calculated using Eq 1:

$$MAV = \frac{1}{N}\sum_{i=1}^{N} |x_i| \qquad (1)$$

Root Mean Square (RMS): This feature is associated with constant force and contraction non-fatiguing (Irfan *et al.*, 2016). The RMS equation can be seen in Eq. 2:

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \qquad (2)$$

Simple Square Integral (SSI): This parameter is defined as an energy index (Irfan *et al.*, 2016) whose value can be calculated using Eq. 3:

$$SSI = \sum_{i=1}^{N} x_i^2 \qquad (3)$$

Variance of EMG (VAR): Is another power index (Irfan *et al.*, 2016). Can be calculated using Eq. 4:

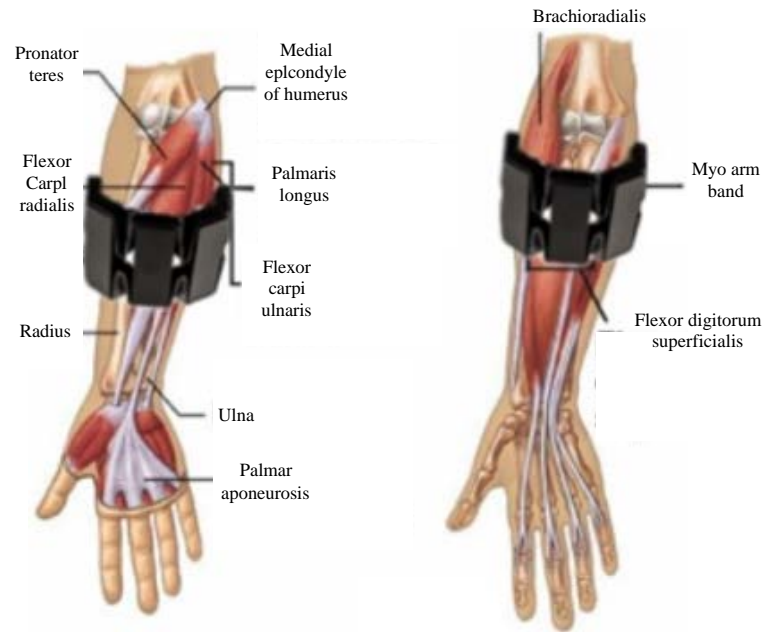$$VAR = \frac{1}{N-1}\sum_{i=1}^{N} x_i^2 \qquad (4)$$

Fig. 3: Forearm muscle

Naive Bayes classifier: The Naive Bayes method is a method used to predict probability. While Bayes classification is a statistical classification that can predict a member class of probability. The concept of the classification of Naive Bayes or so-called Naive Bayes classifier assumes that its classification is based on the effect of a class attribute value whereas the attribute given is free of other attributes or is independent. The main feature of this method is a very strong (Naive) assumption of the independence of each condition/event where it is assumed that each sample attribute (sample data) is mutually exclusive according to the class attribute. The Naive Bayes equation can be formulated as:

$$P(C_k \mid x) = \frac{P(C_k)P(x \mid C_k)}{P(x)} \qquad (5)$$

Where:

x = The vector of the feature
C_k = The class probability k

The equation can be written simply:

$$Posterior = \frac{Prior \times Likehood}{Evidence} \qquad (6)$$

It is assumed that independence is formulated as:

$$P(x_i \mid C_k, x_1, ..., x_{i-1}, x_{i+1}, ..., n) = P(x_i \mid C_k) \qquad (7)$$

Then for all i, it can be simplified to be:

$$P(C_k \mid x_i, ..., x_n) = \frac{P(C_k)\prod_{i=1}^{n} P(x_i \mid C_k)}{P(x_1, ..., x_n)} \qquad (8)$$

With P $(x_i, ..., x_n)$ is constant then P (C_k) dan P $(x|C_k)$ can be estimated using Maximum a Posteriori or (MAP). This MAP simplification is called Naive Bayes.

Data acquisition: The number of participants involving in this study was 5 people. The motion conducted in this study was 20 types of motions taken from the Indonesian sign language system. Respondents performed 20 moves for one type of movement. The data retrieval time for each motion sample was about 5 sec. The EMG signal was taken using a tool called Myo arm band that was positioned on the forearm muscles. At the time of installation of myo arm band in the forearm muscles, it is necessary to consider the muscles under the electrode. This needs to be done to reduce EMG signal noise from undesired muscles. The movement is recorded in the start-gesture-end sequence that can be seen in Fig. 4.

The start and end conditions are in the resting position. Gesture is the basic movement of a word gesture. Every move starts from start to end and then the data is saved into CSV file. The new motion retrieval starts from a restart state until the end motion and will be saved as a new CSV file. The motion data are then sent to the PC via. a Bluetooth connection.

The EMG signal obtained at recording was derived from the entire area given the electrode. Since, each motion has a specific characteristic (such as the speed of motion, the initial condition and motion stability), this has

caused the signal pattern of each channel was not the same, even when the motion was done by the same subjects. A high signal indicates the presence of muscle contraction activity as shown in Fig. 5. A specific threshold was implemented to define which channel is active and which one is not.
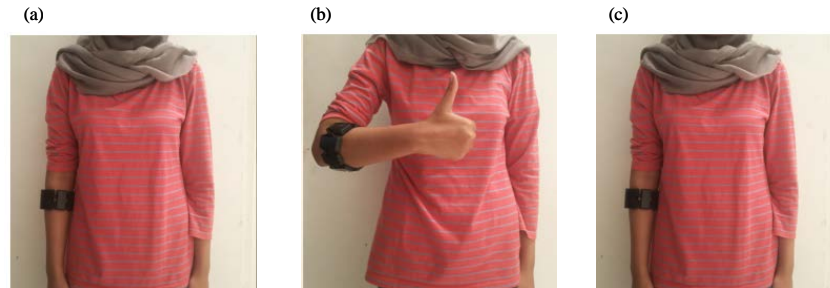

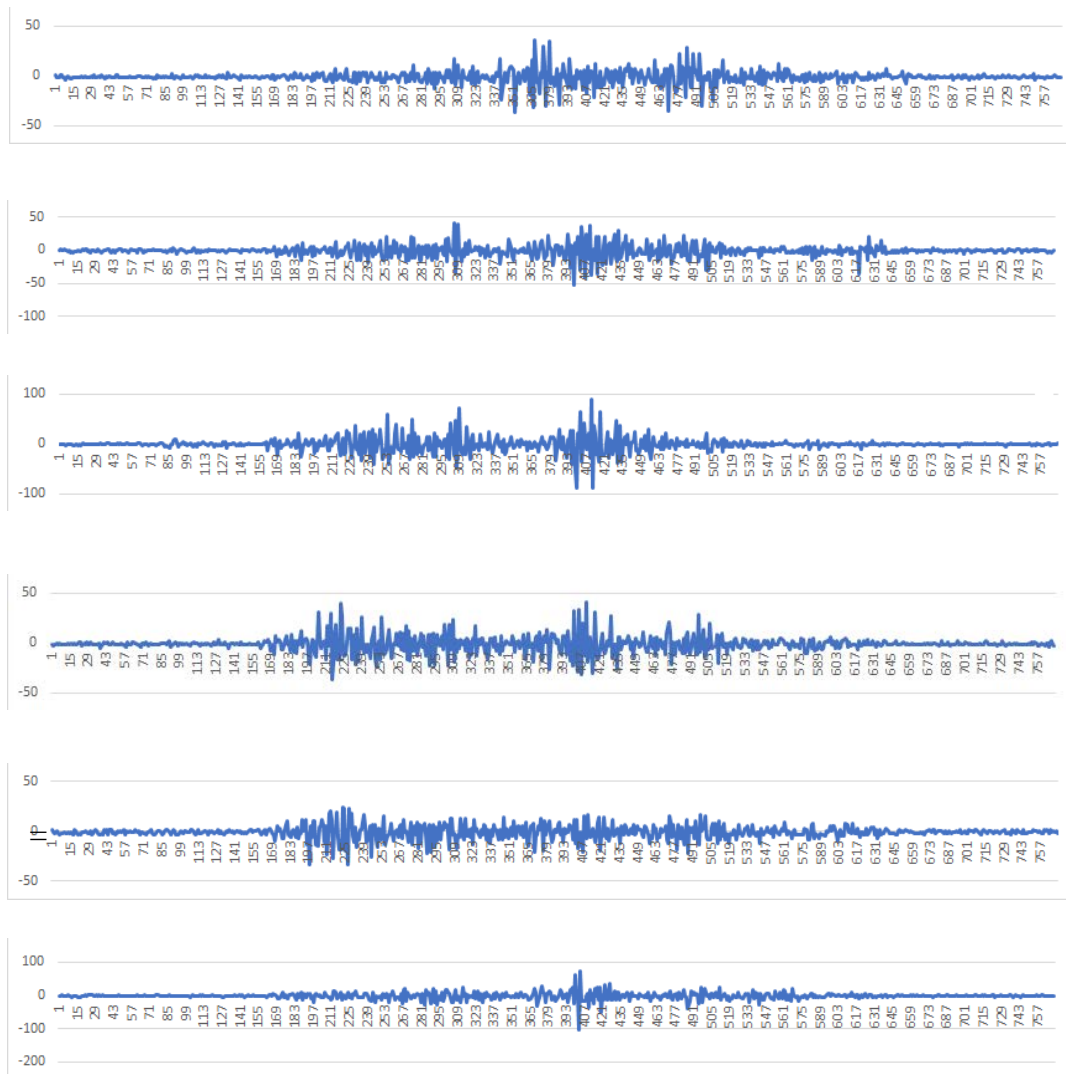Fig. 4: Movement recording sequence; a) Start; b) Gesture and c) End
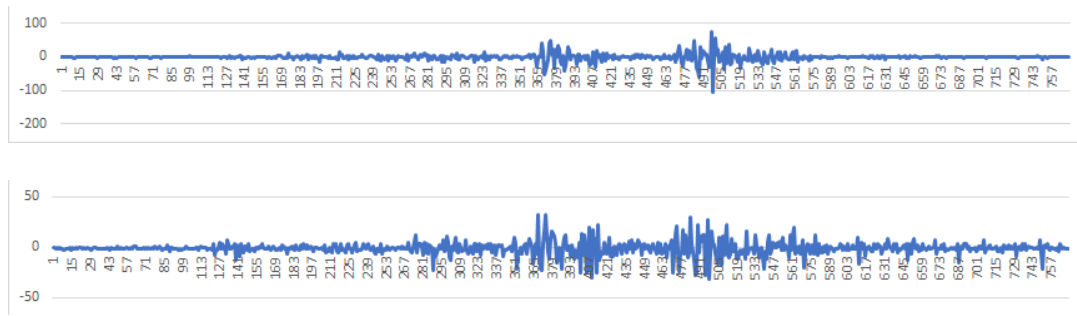

Fig. 5: Continue

Fig. 5: Visualizing RAW EMG signals every channel (1-8)

Pre-processing and feature extraction: The raw data of EMG signal recording contains positive and negative values. The negative values need to be transformed into an absolute value, so that, the threshold then can be applied (Abreu *et al.*, 2016). In order to clear the low amplitude noise that still participates in the reading of the data, a specific threshold setting was required. This threshold can also be used to detect the start and end of the movement (the onset and offset). In this study, the threshold value used was 20 µV (Wibawa *et al.*, 2016). Signals that have values below 20 µV will be zeroed. The raw positive EMG was then being extracted by using time-domain feature extraction to get the input for classification process. The method used for feature extraction were: MAV, RMS, VAR and SSI values. In the same time Myo arm band produces eight signals from the 8 channel EMG sensor. From the eight EMG channels from Myo arm, 4 features would be extracted, so, the output for classification as many as 32 features. In other words, one motion data has created 32 features.

Classification: The classification method used in this research was Naive Bayes. Naive Bayes classifier is a supervised learning method that requires data learning as a representation. Therefore, the data has already been grouped data training and data testing. Data training is used as a representation of knowledge that will be used to predict new data classes that have never existed. Data testing is used to measure to the accuracy of the classifier. The classification steps of the Naive Bayes method was presented below:

Calculation of parameter values: The Gaussian algorithm requires two parameters, namely the mean value (mean) of the class and also the value of the class variable. Suppose the mean value is μ_c and the standard deviation value is σ^c, so that, the variance value is σ_c^2. Standard deviation is used to know the diversity in a data set.

Gaussian calculation: After calculating the probability of class C with the new data v (data testing) or P (x = v|C) using the Gaussian calculation formula. The next calculation was the probability value between classes.

Class predictions: Due to the condition that in this study, the data are multiclass then the next value of arg max (arguments of maxima) was obtained by calculating the multiplication of all probabilities resulted from each class. The largest value obtianed from the calculation between the class would be defined as the result of class prediction. For this case this value will be used to determine the meaning of motion.

RESULTS AND DISCUSSION

In this research, the level of accuracy was calculated by using Eq. 9:

$$\text{Accuration} = \frac{\text{Count of true movement}}{\text{Total movement}} \times 100\% \quad (9)$$

Before classification, the data input was separated by data training and data testing. The comparison of the amount of data used for this class was 50:50, meaning that 50% of data will be used as training and the rest data will be used as testing. In this study, we evaluated the accuracy of the classification method by using two stages. In the first stage, training data will be used as training data and as test data at once. While on the second stage, the testing data will use new totally data that have not been learned before. The result of classification can be seen in Fig. 6.

The accuracy obtained in the classification of trained data was greater than the accuracy of new data, that was 91% compared to 79%. We can

Tabel 1: Comparison of the amount of training and testing data

Trial calculation	Number of data	
	Training	Testing
1	200	1800
2	400	1600
3	600	1400
4	800	1200
5	1000	1000
6	1200	800
7	1400	600
8	1600	400
9	1800	200
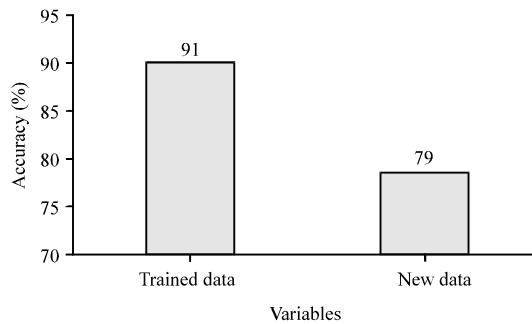
Fig. 6: Comparison of accuracy testing by using trained data and untrained data/new data
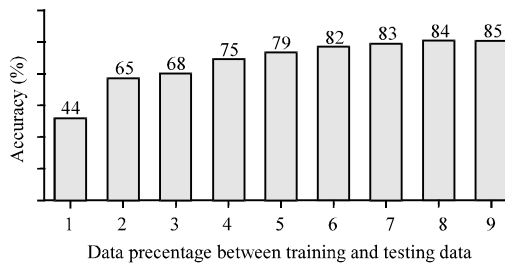
Fig. 7: Results of accuracy on the comparison of the amount of training data and data testing

evaluate here that even for the trained data, the accuracy reached maximum 91%. We opinioned that this could be caused by the variation of motion during each performance in each subject. In this study, we also simulated the accuracy of classification by applying different percentage of data being used as the training data and the testing data.

From a total of 2000 samples data (20 motion, repeated 20 times from 5 subjects), it was divided into training data and data testing as in Table 1. The result of the accuracy simulation has been showed in Fig. 7.

The results of accuracy of each classification can be observed in Fig. 7. In the classification 40 samples showed the lowest accuracy, 44%. The greatest accuracy

Table 2: Accuracy results using 3 features

Experiments	Parameters			Accuracy (%)
1	MAV	RMS	VAR	84
2	MAV	RMS	SSI	81
3	MAV	VAR	SSI	82
4	RMS	VAR	SSI	82

Table 3: Accuracy results using 2 features

Experiments	Parameters		Accuracy (%)
1	MAV	RMS	84
2	MAV	VAR	85
3	MAV	SSI	80
4	RMS	VAR	83
5	RMS	SSI	81
6	VAR	SSI	82

Table 4: Accuracy results using 1 feature

Experiment	Parameter	Accuracy (%)
1	MAV	81
2	RMS	84
3	VAR	82
4	SSI	80

was obtained when the amount of training data is 360 samples which is 85%. The sample size of each class is 18 samples. So, there are 18 variations of data held in each data class.

Better accuracy continues to increase significantly starting from the data training amounted to 40 samples up to 200 samples. Further accuracy increased constantly by 1% in each experiment. From this experiment, we obtained that the accuracy was better when the training data was greater than the amount of testing data. We also found that the more variety of data, the level of accuracy was also increasing. We found that this is because the function of training data is as a representation of knowledge that will be used to predict new data class.

Accuracy tests on the effect of features: Based on our results of classification on the ratio of training and testing data 60:40, using four features, namely MAV, RMS, VAR and SSI, we obtained 82% accuracy. With this result we run some tests to evaluate the effect of number of features in increasing or decreasing the accuracy of data classification. The goal of this test was to analyze which features that likely contributed the most in obtaining high accuracy of classification using Naive Bayes or vice versa. The results of the classification using 4 features will be compared with the result of classification using the same number of data comparisons but with features that are <4 features. In the first experiment, the classification was done by removing one feature and then viewed the effect of the features on the accuracy result. In the second experiment, the classification was performed by

eliminating two features. Furthermore, in the third experiment, the classification was performed using one feature only. Then the accuracy results will be compared with the accuracy results using four features. The results of the tests are shown in Table 2-4.

Table 2 is the result of classification accuracy, using three features. In the classification using three features the highest accuracy was obtained 84% by using MAV, RMS and VAR features. The result of this accuracy is higher when compared to the accuracy result using 4 features that have an accuracy of 82%. It can be concluded that the removal of SSI features can improve accuracy. The MAV and RMS parameters do not have a large effect on the classification. The parameters that showed the greatest influence was VAR feature which means that if this feature is removed from classification, it can result in decreasing the accuracy. Table 3 is the result of classification accuracy using two features. On the classification using two features the highest accuracy, using MAV and VAR feature was 85%. The result of this accuracy is higher when compared to the accuracy result using 4 features that have an accuracy only 82%. From experiment 1-6 (Table 3) it can be seen that the use of SSI features in all experiments on average has decreased the accuracy. While the use of the VAR feature on all experiments on average can improve the accuracy. This result confirmed well our previous finding that the parameter that has the greatest influence in increasing accuracy is the VAR feature. The removal of the VAR feature resulted a decreased accuracy.

Table 4 is the result using only one feature. In the first experiment, the classification was performed using MAV feature yielded 81% accuracy. In the second experiment, the classification was reached 84% using the RMS feature. In the 3rd experiment, the classification was done by using VAR feature and yielded 82% of accuracy. In the 4th experiment, the classification performed using the SSI feature obtained 80% accuracy results. The order of accuracy level from the greatest tested feature was: RMS, VAR, MAV and SSI with result 84, 82, 81 and 80%. We found that actually by applying just one feature, the system can classify the data with high accuracy.

CONCLUSION

However, we opinioned that this result depends on the number of motions and type of motions that will be implemented in sign language classification. For future works, more motions should be tested in the experiment,

so that, better understanding in sign language classification technique then can be obtained. Based on the analysis and discussion in the previous section, we obtained the following conclusions: in order to get higher accuracy in sign language classification, the number of training data used for classification should be ≥50% of the total sample data. In our result the accuracy level was 85% at the highest. On the classification using three features, obtained the highest success rate was 84% by using MAV, RMS and VAR features and the lowest success rate was obtained by using features of MAV, RMS and SSI that yielded 81% accuracy. On the classification using two features, we obtained the highest accuracy using MAV and VAR features was 85% and the lowest accuracy results using MAV and SSI features was 80%.

The accuracy of the classification using one feature alone was quite high, it was 84%. The order accuracy from the highest to the lowest was feature: RMS, VAR, MAV and SSI with result 84, 82, 81 and 80%. Applying several features in sign language classification can cause in bigger computation cost and time. Choosing the best format for a specific ideal condition need to be tested and calculated, besides considering the type of motions and the number of motion, so that, the process of translating sign language can be achieved effectively and fast.

REFERENCES

Abreu, J.G., J.M. Teixeira, L.S. Figueiredo and V. Teichrieb, 2016. Evaluating sign language recognition using the Myo armband. Proceedings of the 2016 18th International Symposium on Virtual and Augmented Reality (SVR), June 21-24, 2016, IEEE, Gramado, Brazil, ISBN:978-1-5090-4149-7, pp: 64-70.

Ibraheem, N.A. and R.Z. Khan, 2012. Survey on various gesture recognition technologies and techniques. Intl. J. Comput. Appl., 50: 38-44.

Irfan, M., W. Caesarendra and M. Ariyanto, 2016. [The classification study of seven hand motion signal electromyography (EMG) using the pattern recognition method (In Indonesian)]. J. Teknik Mesin, 4: 307-316.

Murthy, G.R.S. and R.S. Jadon, 2009. A Review of Vision Based Hand Gestures Recognition. Int. J. May 25, 2019Inf. Technol. Knowledge Manage., 2: 405-410.

Pandey, P. and V. Jain, 2015. Hand gesture recognition for sign language recognition: A review. Intl. J. Sci. Eng. Technol. Res., 4: 464-470.

Tewari, D. and S.K. Srivastava, 2012. A visual recognition of static hand gestures in Indian sign language based on kohonen self-organizing map algorithm. Intl. J. Eng. Adv. Technol., 2: 165-170.

Wibawa, A.D., N. Verdonschot, J.P.K. Halbertsma, J.G.M. Burgerhof and R.L. Diercks *et al.*, 2016. Musculoskeletal modeling of human lower limb during normal walking, one-legged forward hopping and side jumping: Comparison ofmeasured EMG and predicted muscle activity patterns. J. Biomech., 49: 3660-3666.

Yang, H.D., 2015. Sign language recognition with the kinect sensor based on conditional random fields. Sens., 15: 135-147.