Biomedical Engineering: Applications, Basis and Communications, Vol. 27, No. 2 (2015) 1550020 (9 pages)
DOI: 10.4015/S1016237215500209

SYLLABLE-BASED SPEECH RECOGNITION USING ELECTROMYOGRAPHY AND DECISION SET CLASSIFIER

Marko Topalović*
University of Kragujevac, Faculty of Engineering, Kragujevac, Serbia

Đorđe Damnjanović
University of Kragujevac, Faculty of Technical Sciences Čačak
Čačak, Serbia

Aleksandar Peulić, Milan Blagojević and Nenad Filipović
University of Kragujevac, Faculty of Engineering, Kragujevac, Serbia

Accepted 16 August 2014
Published 12 December 2014

ABSTRACT

During the speech, contractions of muscles in the speech apparatus produce myoelectric signals that can be picked up by electrodes, filtered and analyzed. The problem of extraction of speech information from these signals is significant for patients with damaged speech apparatus, such as laryngectomy patients, who could use speech recognition based on myoelectric signal classification to communicate by means of the synthetic speech. In the most previously conducted research, classification is performed on a ten word vocabulary which resulted in a good classification rate. In this paper, a possibility for myoelectric syllable based speech classification is analyzed on a significantly larger vocabulary with novel decision set based classifier which is simple, easy to adapt, convenient for research and similar to the way humans think. In order to have a high quality of recorded myoelectric signals, analysis of the optimal position of electrodes is performed. Classification is performed by comparison between syllable combination and whole words. Based on classification rate, words can belong to easy, medium or hard to distinguish group. Results based on generated list of best matching combinations show that decision set analysis of myoelectric signals for speech recognition is a promising novel method.

Keywords: Speech recognition; Electromyography; Decision set; Speech apparatus muscles.

INTRODUCTION

Electromyography (EMG), is the analysis of electrical signals generated by muscle contraction (myoelectric signals), and it is widely used to study various medical conditions[1] and to perform diagnostics.[2–4] Myoelectric signals are also used to facilitate human–computer interaction.[5,6] One of the most important and successful application of EMG is control of myoelectric prosthetics.[7,8] The movement of these artificial limbs is controlled by interpretation of myoelectric signals,[9] which greatly improves quality of amputee life.[10] These success studies lead many researchers into investigation of other possible applications of EMG such is speech recognition by classification of myoelectric signals generated by speech apparatus.[11,12] Similarly to myoelectric

*Corresponding author: Marko Topalović, University of Kragujevac, Faculty of Engineering, Kragujevac, Serbia. E-mail: topalovic@kg.ac.rs

prosthetics which help amputee, EMG signals are used to help people with motor disabilities,[13] such are laryngectomy patients, which have damaged speech apparatus that makes their interaction with other people more difficult.[14] EMG is used for the development of silent speech recognition and thus improving the quality of laryngectomee life.[15,16] EMG is also applied to augment audio speech recognition of healthy subjects in acoustically harsh environments.[17,18] Such noisy conditions can occur in aircraft cockpit or firefighter breathing apparatus.[19,20] These multi-expert classifiers offer increased reliability in comparison to classic speech recognition techniques.[21]

Most of the EMG classification research done so far was performed on very small vocabulary of tested words which is the main problem that we addressed in this paper. Commonly 10 word vocabulary that consists of numbers one to ten is used. Although 10 words might be sufficient for a fighter pilot to issue commands to the plane, for normal communication with patients much larger vocabulary is required. Since people use about 3000 different words in their day to day speech, recording them one by one into classifier is a lengthy and difficult endeavor. This setback is overcome by phoneme based classifiers which use phoneme signals which are joined into words and used for classification.[22,23] Using syllables for speech recognition is a novel approach.[24–28]

In this paper, we presented syllable based EMG classifier which is a compromise between accuracy of word based and flexibility of phoneme based classification. Since syllable combinations are used, the number of words in classifier vocabulary is much greater which decreases the classification rate. To compare words with syllable combination and to classify them based on their likelihood, we used decision set classifier, which is a new method used for EMG classification. The decision sets that we used can be considered as a derivation of fuzzy sets. Fuzzy logic is widely used to make decisions most adequate to the current situation, but it is also used in signal processing.[29,30] Biological signals are recognized with fuzzy logic as well.[31] Signals stored in classifier are used to create arrays of sets. These sets are used to assess tested signals. The fact that there are arrays of these sets differs them from the classic notion of fuzzy sets, so we find term decision sets more adequate.

So far, there have been two approaches for the state of the art EMG classifiers bio-mimicking methods such are artificial neural networks (ANN) classifiers,[32–40] and statistical methods like Hidden Markov Model (HMM)[41] or support vector machines (SVM).[41–46]

The method that we developed has some similarities with ANNs which mimic functionality of nervous system using computer simulation of neurons. Most neural networks have input layer which receives information, output layer which presents result of neural network analysis and a hidden layer which processes the information. In our classifier, input signal is divided into segments and every segment is compared with values from knowledge base (KB). These segments correspond to input neurons and they transfer degree of membership value as an input to the summing functions which correspond to output neurons. This representation means that our classifier does not have equivalent to hidden layer of neural networks. Also, signals are compared with every possible combination of syllables which would correspond to very simple interconnection between neurons in neural network (including connecting a syllable with itself for example for word "mama"). In order to avoid comparing input signal with meaningless combinations, entire dictionary must be added to the program so that only real syllable combinations are tested in classification process. If each syllable is considered as neuron, this would create more realistic and scarcer interconnection pattern. In pattern recognition, implementation of ANN supervised learning is used, which means that for the input EMG signal of tested word and for the output textual or synthetic speech representation of the word, neural network classifier must determine weights of neuron connections. In our classifier, EMG signal values for syllables are stored in KB and degree of membership (which correlates to neuron connection weight) is calculated and every time signal is analyzed for every segment of the signal.

Another state of the art statistical method for speech classification SVMs use training examples which are divided into two categories and represented as points in space with a gap between these two categories as wide as possible. When new example is added, based on it's position in space it can belong to one or another category. Since this is binary classification method, in order to use it as a classification tool for speech recognition, this method must be expanded into multiclass SVM (MSVM).[45] This means that every multiclass problem is considered as a large number of binary classification problems. One versus all approach in MSVM determines the highest output function for all cases and selects it as a classification result. Our classifier similarly compares sums of degrees of membership and creates list of best matches. The main disadvantage of SVM approach is its computer requirements due to a need for a large training data,[42] so our syllable based classifier offers re-use of recorded syllable signals for comparison with tested words. Fuzzy nature of our classifier offers results which are not strict like yes/no but more flexible which accommodates for uncertainties, noise and similarities which are more present in EMG speech recognition in comparison to the audio speech recognition.

MATERIALS AND METHODS

EMG signals produced by muscle contractions during speech are picked up on various regions of the subject's face, depending on spoken phoneme. We determined that vowels are more distinguishable than consonants. Consonants are best distinguished in neck region (Fig. 1 (a)), while the vowels are best distinguished on lower jaw (Fig. 1(b)). Syllables are made of both vowels and consonants, so we used the neck position as it gives better results in distinguishing syllables.

For the signal acquisition, we used low power equipment based on MSP430F1232 microcontroller produced by Texas instruments® (Fig. 2).

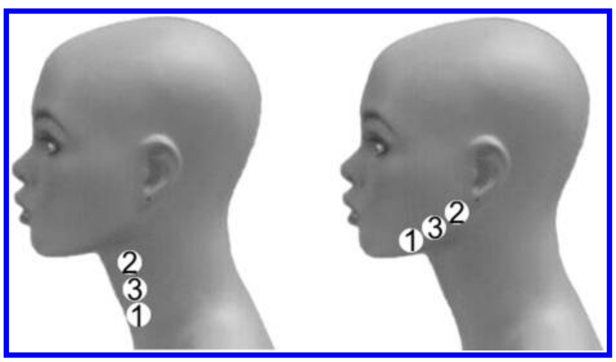


Fig. 1 Positions for electrode placement: **(A)** Neck position and **(B)** jaw position.

Three electrodes connected to this controller are placed on the test subject's neck (Fig. 1(a)). Electrodes 1 and 2 measured voltage between them when the EMG signal occurs, while electrode 3 is placed between them and it was used to cancel noise. The further noise reduction was done by application of Butterworth filter. After the syllable signals were recorded and filtered, they were inputted into the classifier. Later, the same subject spoke several words, which were also recorded, filtered and inputted in classifier for analysis.

EMG syllable signals, after processing, have on average 1200 time discretized values. To further cancel noise influence, and to make process practicable in scope of current computer resources, signal is divided into segments, for each segment signal average is calculated. Averaging reduces noise, but it also means that some information is lost in the process, so in light of that tradeoff, 50 segments are chosen as an optimum value (Fig. 3).

Syllable signal information is stored in the KB. The process of storing syllable information starts with classifier opening filtered syllable signal, reading it line by line, creating an array of time discretized values. This original array is reduced to 50 segment array by averaging signal values for each segment (Fig. 4(A)).

Classifier then compares the new segmented tested signal with every syllable already stored in the KB and creates a list of best matches. The classifier can store several signals of the same syllable, because user can pronounce the same syllable faster or more slowly. When tested syllable signal is compared to the particular syllable in KB, classifier calculates average value for every segment.

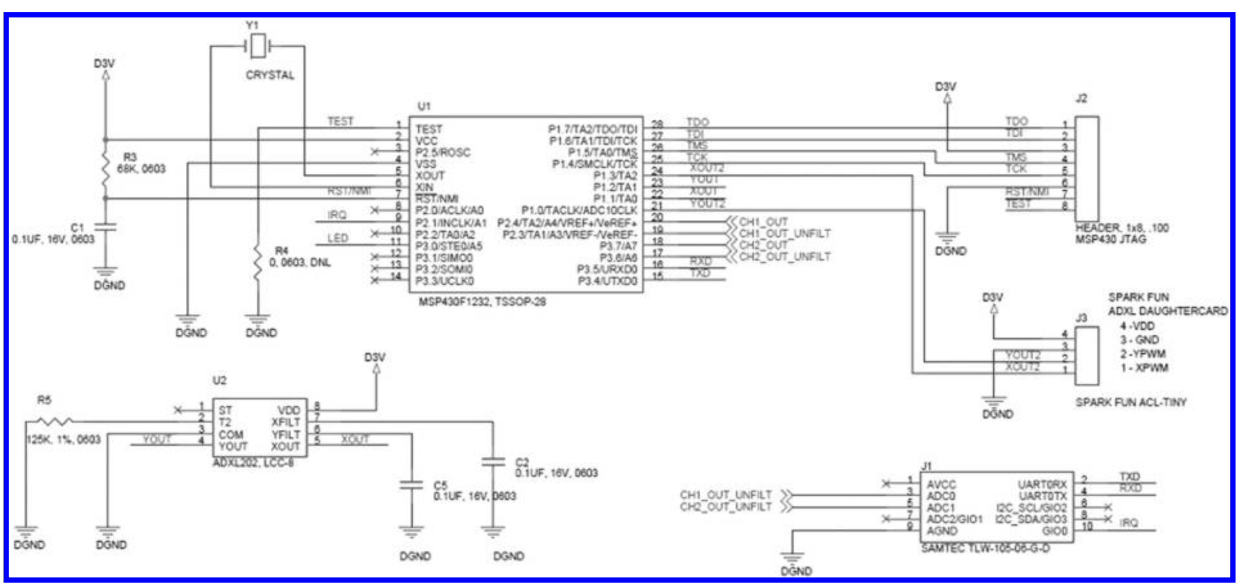


Fig. 2 MSP430F1232 based controller.

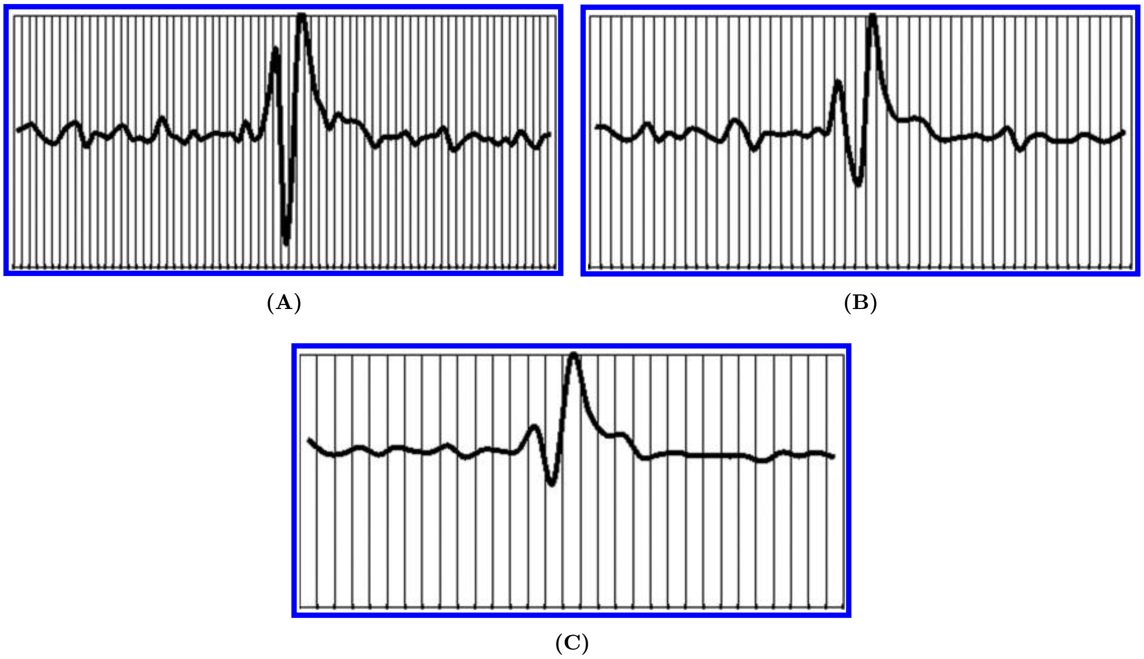


Fig. 3 Influence of number of segments **(A)** 70, **(B)** 50 and **(C)** 30.

If tested signal is not in the KB, user (test subject, classifier operator, or patient) can type in new syllable and save it in KB (Fig. 4(B)).

The same as loud noise can interfere with sound of someone's voice, other myoelectric signals can interfere with signals generated by speech apparatus. People can grade certainty of what they heard from "not sure" to "pretty sure". To emulate the way humans think, we used the decision set logic for our classifier. Decision set classifier creates a list of 10 words which syllable signal combinations have most resemblance with tested signal. With no required knowledge of high level mathematics, its simple design makes it highly adoptable, a perfect tool for research in EMG signal processing. Elements of decision sets have degrees of membership defined by graduation, which means that everything can be graded by the degree of truth membership, ranging from $0 =$ false to $1 =$ true (Fig. 5).

Decision set can be defined as pair (U, m) where U is a set and $m : U \to [0, 1]$ is membership function of the decision set (U, m). For each $x \in U$ the value of membership function $m(x)$ is called grade (degree) of membership of x in (U, m). If $m(x) = 0$, x is not included in set (U, m). If $m(x) = 1$, x in fully included in set (U, m). In the third case we have $0 < m(x) < 1$. Decision sets used for EMG classification are shown on Fig. 5(A).

For a description of tested signal, we used 3 decision sets with 2 rectangular and one triangle membership function. If tested signal value is within the 50–150% scope of the KB signal value, membership function is a triangle with height of 1, and it can be described as medium valued (MV) (Table 1). If tested signal value is between 0% and 55% of KB signal value membership function is a rectangle with height of 0.1, and it can be described as small valued (SV). If tested signal value is above 145% of KB signal value membership function is a rectangle with height of 0.1, and it can be described as large valued (LV).

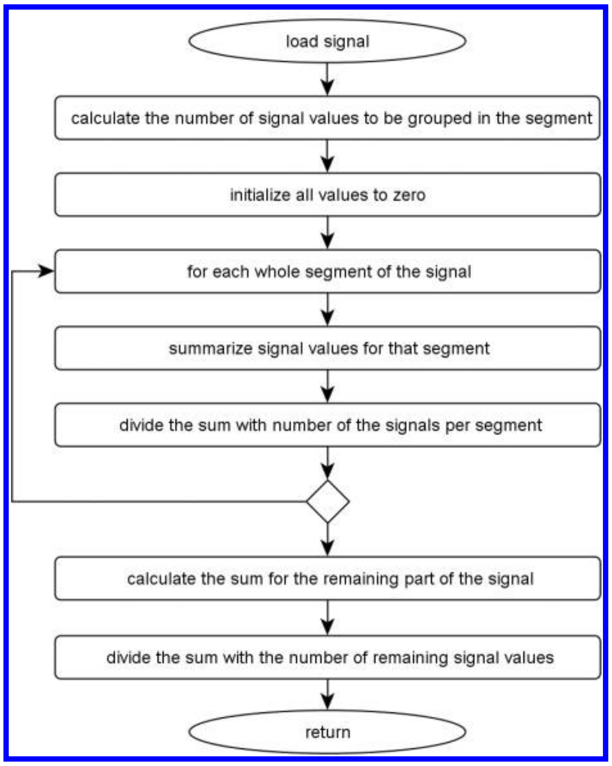
On these sets we can perform the union operation $m_{\mathrm{SV}\cup\mathrm{MV}\cup\mathrm{LV}} = \max(m_{\mathrm{SV}}, m_{\mathrm{MV}}, m_{\mathrm{LV}})$ so when tested signal falls in SV or LV category membership value is 0,1. For MV category (triangle membership function) degree of membership is calculated using the test signal segment value and average KB values (Fig. 5(B)).

There are two cases for MV category: Tested input signal is lesser than the average value in KB (left half of the triangle) and second case when tested input signal is greater than average value in KB (Fig. 5(B)). For lesser value, degree of membership can be calculated as

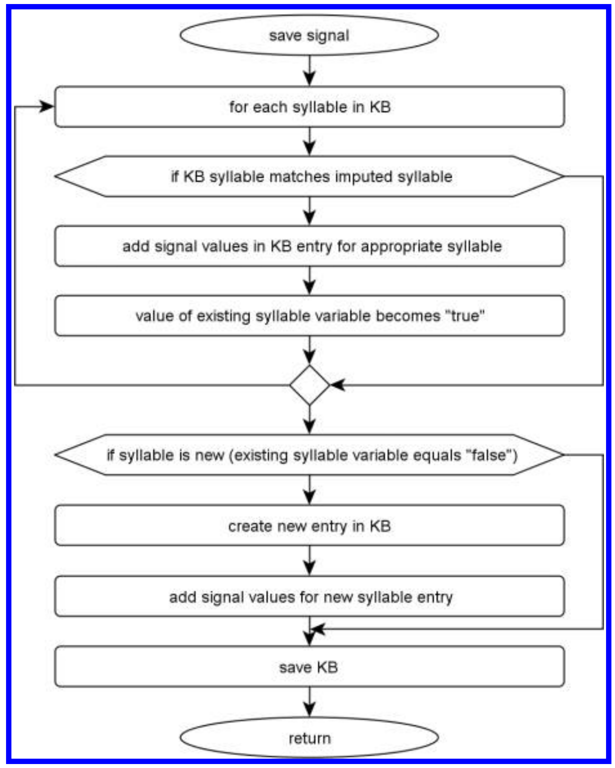
$$m(\nu_{ti}) = \frac{\nu_{ti} - \nu_{\min}}{\nu_{\mathrm{avg}} - \nu_{\min}} \tag{1}$$

where ν_{ti} represents value of tested input signal, ν_{avg} is average value of signals for that segment from KB and $\nu_{\min}$ represents the minimum value for triangle membership function (in this case 50% of the average value). For the greater value degree of membership can be calculated as

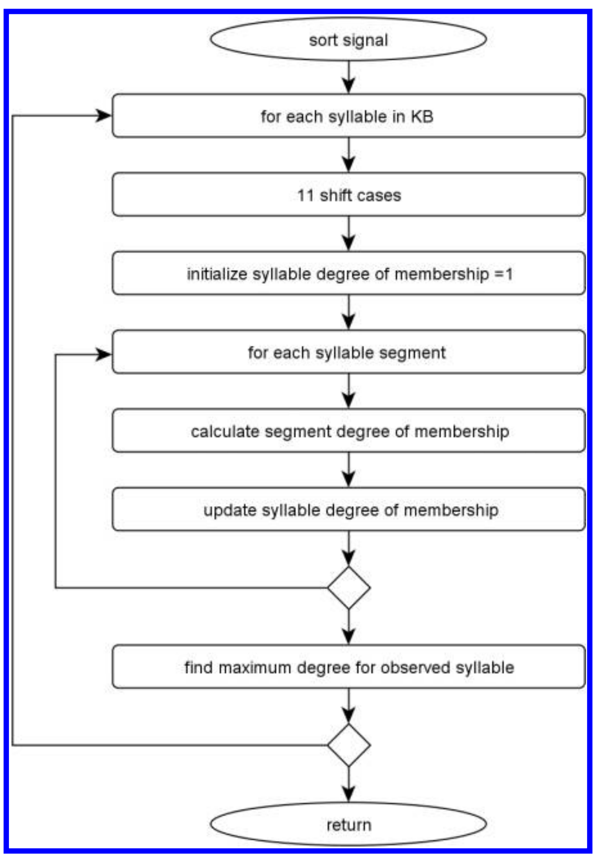
$$m(\nu_{ti}) = \frac{\nu_{\max} - \nu_{ti}}{\nu_{\max} - \nu_{\mathrm{avg}}} \tag{2}$$



(A)

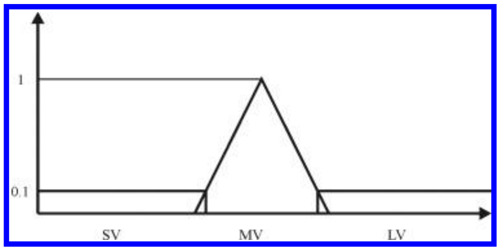


(B)

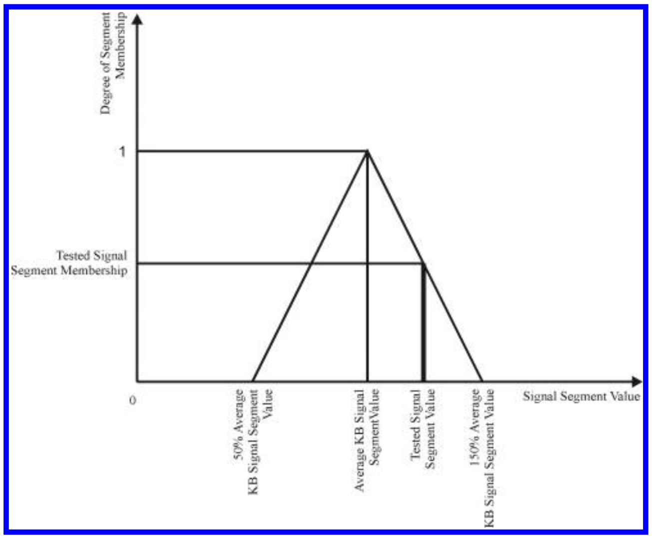


(C)

Fig. 4 Algorithms: **(A)** Reducing input signal into segments, **(B)** input of syllable signal into KB, and **(C)** determination of overall degree of membership for every syllable.



(A)



(B)

Fig. 5 Fuzzy logic implementation: **(A)** Decision sets for EMG signal classification, and **(B)** calculation of degree of segment membership.

where $\nu_{\max}$ represents the maximum value for triangle membership function (in this case 150% of the average value).

The overall degree of membership for that particular syllable is obtained by multiplying all degrees of segment membership. Having product of 50 real numbers less than 1 results in overall membership function very close to 0. Nevertheless, if the tested signal is similar to KB signal, degrees of segment membership are larger than degrees of segment membership for very distinct signals. Therefore, for more similar signals, the overall degree of membership is larger (even though it is still number close to 0).

Because the test subject cannot pronounce syllables at the same time in regard to beginning of signal recording, phase shift is taken into account (Fig. 6).

Table 1. Decision Set Functions for EMG Signal Classification.

Tested Signal Category	Degree of Membership	
Small value signal (SV)	Small degree (SD)	$m(x) = 0.1$
Medium value signal (MV)	Meduim degree (MD)	$0.1 < m(x) < 1$
Large value signal (LV)	Small degree (SD)	$m(x) = 0.1$

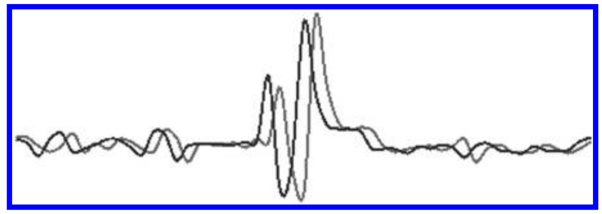


Fig. 6 Phase shift of syllable signal.

For every signal in KB, comparison is performed 11 times and the highest overall degree of membership is assigned for that syllable. 11 comparisons include original tested signal without phase shift and 10 shifted signals which are shifted 1–5 segments forward or backward (Fig. 4(C)).

Once KB is filled with syllable signal information, classifier can analyze word signals. For word classification, the procedure is similar to previously described classification of syllables with a few distinctions. First, classifier creates an array of all possible syllable combinations by joining syllable signals. Number of combinations is equal to the number of syllables with an exponent of the number of syllables in a word. Tested word signal is reduced to 50 segment array and so are syllable combinations. Calculation of the overall degree of membership is the same as for the syllable classification. List of best matching syllable combinations is obtained by sorting overall syllable combinations using language-integrated query (LINQ).

RESULTS

Tested word which is inputted into classifier is compared to all syllable combinations of all syllables saved in KB. Classification accuracy greatly varies from word to word depending on numerous factors such as the difference between first and second syllable, a number of similar syllables to the syllables that comprise tested word, the connection between syllables,[39] the distinctiveness of syllables. In Table 2, classification results are given for the case of 11 syllables in KB which make 121 combinations. Some words have very dinstinctive signals (which is similar to combination of syllable signals that makes them have much greater overall degree of membership while some others words lack this distinctiveness and therefore their overall degree of membership is very low). For any given tested word, classifier displays 10 syllable combinations with the highest overall degree of membership. We observe some characteristic cases (Table 2). For tested word "mama" (in Serbian mother) "ma-ma" syllable combination was on the top of the list. For test word "macka" (in Serbian cat) "ma-cka"

Table 2. Classification Results.

Word		Syllable Combination	Percentage of Sum of all Overall Degrees of Membership (%)
"mama" (mother)	1	mama	15.16523311736
	2	ckako	13.83012679668
	3	ckama	13.75651137424
	4	kama	4.239479717647
	5	kako	4.005676058007
	6	maje	3.547225178866
	7	koma	3.297460260103
	8	jeje	3.093384034208
	9	maja	2.505058700428
	10	jeko	1.886801904893
"macka" (cat)	1	maja	87.82029914485
	2	jeje	5.171563551503
	3	macka	3.084648427361
	4	mala	0.840089519638
	5	ckaja	0.626002341495
	6	kuja	0.391911235745
	7	koma	0.382426900711
	8	mama	0.192735954764
	9	jeja	0.187667730975
	10	maka	0.177965325259
"jaje" (egg)	1	maja	45.95760576238
	2	mala	27.27694368970
	3	kucka	10.63807402424
	4	koja	4.893464592166
	5	macka	3.253079233892
	6	kuko	1.691108961448
	7	jama	1.084150650928
	8	jaje	0.850407113628
	9	jaja	0.774863246454
	10	jeje	0.659389403791

combination was third on the list, these are some of the easiest to distinguish words. Example of medium distinguishable word is "jaje" (in Serbian egg) for which "ja-je" combination was in the eighth position. Members of hard to distinguish group are words such as "kola" (car), "kuca" (house), or "slika" (picture, painting). Their appropriate syllable combinations were not among first ten in likelihood list. Similar words are also high on classification list, for instance, we have combinations that have one same syllable as the tested word and one different. For example, for tested word "mama" similar combinations are "ckama", "kama", "maje", "maja" which all were high on likelihood list (Table 2).

DISCUSSION

Classification results presented in Table 3 show that as the number of syllables in the knowledge base increases, accuracy of classification decreases. This was expected because adding a new syllable increases number of combinations, and every syllable combination has some

Table 3. Sensitivity and Accuracy Results.

		Accuracy for Different Number of Syllable Combinations				
Word	Sensitivity Coefficient (%)	9 (%)	25 (%)	49 (%)	81 (%)	121 (%)
Mama	15.68	44.96	30.36	16.63	15.40	15.17
Macka	38.19	93.19	74.35	74.00	59.40	3.08
Jaje	3.00	36.89	34.05	29.00	3.92	0.85
Kola	0.44	5.97	3.93	1.02	0.50	0.00
Tata	71.72	71.72	0.73	0.09	0.01	0.00

degree of similarity to the right combination that we wish to be on the top of the list. If a knowledge base is expanded with syllables that are not similar to the syllables of tested word, classification accuracy remains high. On the other hand, adding similar syllables results in a sharp drop of accuracy. Syllable based classification faces issue of proper connection between syllable signals.[39] When a person speaks whole words, transition between syllables is smooth and fast, while when syllable signals are recorded there is no transition segment, so in some cases combination of two "wrong" syllables is more similar to the tested word than a combination of the right syllables. Another issue with this method is that the accuracy results do not reflect similarity between tested signal and the combination of syllable signals, but are calculated based on the share of particular syllable combination sum of membership degrees in total sum for all combinations. For example, we cannot say that word "mama" is 70% similar to "ma-ma" combination. We can create a list of best matches (Table 2) and say that for tested word "mama" there is a 15% chance that "ma-ma" was spoken, 14% for "cka-ko" combination and so on. Sensitivity is the relation between changes in output and changes in the input of a system. In this case, changes of input are hard to measure quantitatively, because it is not possible to obtain difference percentage for signals. To test the sensitivity of classifier, we used phase shifting of input signal (which could be easily measured) and observed changes of the output sum of degree membership (Table 3). There were great differences in sensitivity coefficients for different words, but these sensitivity coefficients do not have a strong correlation with accuracy. If the signal is more sensitive to phase shifting, then this signal is more distinguishable and therefore accuracy should be greater, but there are many other factors (like the previously described connection between syllable signals) which could negate this property.

Based on classification results, tested words can belong to easily distinguishable words, medium distinguishable words or hard to distinguish words. Difference in classification rate comes primarily from two factors: One is the electrode position which has a great influence on the quality and dissimilarity of the myoelectric signals and the other is the nature of phoneme generation. Position of electrodes effect signal noise that is picked up alongside EMG signals generated by speech apparatus. Since EMG signals are used for classification, audio noise is irrelevant to classification rate (which is the reason EMG classification is very suitable for augmenting audio speech recognition in acoustically harsh environments in aircraft cockpit or firefighter breathing apparatus), but every other muscle in the neck region (not belonging to speech apparatus) also produces EMG signals. Every head movement can produce spikes in EMG signals which could interfere with speech apparatus signals. If the noise signals are recorded while the head is stationary and while the subject is not talking, these signals could be subtracted later when speech recognition takes place. This procedure on the other hand would increase computational cost with no significant increase in accuracy. Application of Butterworth filter and averaging of signals (Fig. 3) are sufficient tool to cancel the influence of noise on EMG signals. Decision set technique based on fuzzy logic is envisioned to be robust and to offer list of best matches containing the right word even if the noise spike interferes with tested signal. The noise would decrease the overall degrees of membership of tested signal, but similarity between speech apparatus signal and combination of syllable signals stored in KB would still put it high on likelihood list.

Depending on phoneme point of origin, phonemes can be more similar or more distinct in comparison to other phonemes. These two factors result in some syllables being comprised of more distinct phonemes than the others and therefore words comprised of those syllables have a greater classification rate. The decision set syllable based classifier needs improvements in order to have a better classification rate for medium and low distinguishable words. Our future work will be focused on signal acquisition and processing with the goal of obtaining more distinguishable signals. The capabilities of syllable based decision set classifier will be improved by creating the logic which will govern joining of syllable signals by removing excessive ends of signal recordings. Also weight functions will be added which will give importance to signal segments based on their deviation from the average signal value.

CONCLUSIONS

Implementation of decision sets in development of EMG signal classifier is a novel approach which has several advantages in comparison to traditional classification methods. Decision set classifier is simple to implement, easy to augment by adding logic that could improve the classification rate. It can be used to test the influence of signal segmentation and averaging, electrode position or any other signal characteristics on classification rate. It is simple, but computer efficient nature means it can be used and modified by people without a high level of mathematics knowledge with common PC and simple microcontroller. Since decision set classifier is the new concept, there is a lot of space for improvements and further research. Syllable based classification in EMG speech recognition yields great flexibility in classifier training and gives possibility to compare all words that are comprised of syllables saved in the classifier knowledge base.

ACKNOWLEDGMENTS

Research presented in this paper was supported by Ministry of Education, Science and Technological Development of the Republic of Serbia, Grant No. TR 32036.

REFERENCES

1. Fuglsang-Frederiksen A, The role of different EMG methods in evaluating myopathy, *J Electromyogr Kines* **16**:586, 2006.
2. Drost G, Stegeman, Van Engelen BGM, Zwarts M, Clinical applications of high-density surface EMG: A systematic review, *J Electromyogr Kines* **16**:586, 2006.
3. Phinyomark A, Phukpattaranont P, Limsakul C, Feature reduction and selection for EMG signal classification, *Expert Syst Appl* **39**:7420, 2012.
4. Subasi A, Yilmaz M, Ozcalik HR, Classification of EMG signals using wavelet neural network, *J Neurosci Methods* **156**:360, 2006.
5. Rezazadeh IM, Firoozabadi SM, Hu H, Golpayegani MRH, A novel human–machine interface based on recognition of multi-channel facial bioelectric signals, *Australas Phys Eng Sci Med* **34**:497, 2011.
6. Mariunas M, Pośka A, Shshok A, Influence of speech transmitting methods on velocity of the understanding of deaf-mute by sensor organs, *Acta Bioeng Biomech* **1**:59, 1999.
7. Chan A, Kwok E, Bhuanantanondh P, Performance assessment of upper limb myoelectric prostheses using a programmable assessment platform, *J Med Biol Eng* **32**:259, 2012.
8. Skolimowski J, Winiarski S, Anwajler J, Skolimowska B, Barczyk K, Dudek K, Bioelectric activity of selected muscle groups in people with impingement syndrome, *Acta Bioeng Biomech* **11**:37, 2009.

9. Rafiee J, Rafiee MA, Yavari F, Schoen MP, Feature extraction of forearm EMG signals for prosthetics, *Expert Syst Appl* **38**:4058, 2011.
10. Khushaba RN, Kodagoda S, Takruri M, Dissanayake G, Toward improved control of prosthetic fingers using surface electromyogram (EMG) signals, *Expert Syst Appl* **39**:10731, 2012.
11. Huang CN, Chen CH, Chung HY, The review of applications and measurements in facial electromyography, *J Med Biol Eng* **25**:15, 2004.
12. Orr MC, Lithgow B, Higher order statistical analysis of /x/ in male speech, *Australas Phys Eng Sci Med* **28**:56, 2005.
13. Pinheiro CG, Naves ELM, Pino P, Losson E, Andrade AO, Bourhis G, Alternative communication systems for people with severe motor disabilities: A survey, *Biomed Eng Online* **10**:1, 2011.
14. Schindler A, Mozzanica F, Ginocchio D, Invernizzi A, Peri A, Ottaviani F, Voice-related quality of life in patients after total and partial laryngectomy, *Auris Nasus Larynx* **39**:77, 2012.
15. Fagan MJ, Ell SR, Gilbert JM, Sarrazin E, Chapman PM, Development of a (silent) speech recognition system for patients following laryngectomy, *Med Eng Phys* **30**:419, 2008.
16. Lu BY, Wu HD, Shih SR, Chong FC, Hsueh ML, Chen YL, Combination of frequency and amplitude-modulated model for the synthesis of normal and wheezing sounds, *Australas Phys Eng Sci Med* **34**:449, 2011.
17. Chan ADC, Englehart K, Hudgins B, Lovely DF, Myoelectric signals to augment speach recognition, *Med Biol Eng Comput* **39**:500, 2001.
18. Lee KS, Prediction of acoustic feature parameters using myoelectric signals, *IEEE Trans Biomed Eng* **57**:1587, 2010.
19. Chan ADC, Englehart K, Hudgins B, Lovely DF, Hidden Markov model classification of myoelectric signals in speech, *IEEE Eng Med Biol* **21**:143, 2002.
20. Betts BJ, Binsted K, Jorgensen C, Small-vocabulary speech recognition using surface electromyography, *Interact Comput* **18**:1242, 2006.
21. Chan ADC, Englehart K, Hudgins B, Lovely DF, Multi-expert automatic speech recognition using acoustic and myoelectric signals, *IEEE Trans Biomed Eng* **53**:676, 2006.
22. Scheme EJ, Hudgins B, Parker PA, Myoelectric signal classification for phoneme-based speech recognition, *IEEE Trans Biomed Eng* **54**:694, 2007.
23. Zhou Q, Jiang N, Englehart K, Hudgins B, Improved phoneme-based myoelectric speech recognition, *IEEE Trans Biomed Eng* **56**:2016, 2009.
24. Jansma BM, Schiller NO, Monitoring syllable boundaries during speech production, *Brain Lang* **90**:311, 2004.
25. Cholin J, Levelt WJM, Schiller NO, Effects of syllable frequency in speech production, *Cognition* **99**:205, 2006.
26. Cholin J, Schiller NO, Levelt WJM, The preparation of syllables in speech production, *J Mem Lang* **50**:47, 2004.
27. Kanas VG, Mporas I, Benz HL, Sgarbas KN, Bezerianos A, Crone NE, Joint spatial-spectral feature space clustering for speech activity detection from ECoG signals, *IEEE Trans Biomed Eng* **61**:1241, 2014.
28. Lee KS, Prediction of acoustic feature parameters using myoelectric signals, *IEEE Trans Biomed Eng* **57**:1578, 2010.
29. Mendel JM, Uncertainty, fuzzy logic, and signal processing, *Signal Pr* **80**:913, 2000.
30. Zio E, Popescu IC, Recognizing signal trends on-line by a fuzzy-logic-based methodology optimized via genetic algorithms, *Eng Appl Artif Intell* **20**:831, 2007.
31. Das A, Reddy NP, Narayanan J, Hybrid fuzzy logic committee neural networks for recognition of swallow acceleration signals, *Comput Method Prog Biol* **64**:87, 2001.
32. Dede G, Hüsnü Sazli M, Speech recognition with artificial neural networks, *Digit Signal Process* **20**:763, 2010.
33. Juang CF, Lai CL, Tu CC, Dynamic programming prediction errors of recurrent neural fuzzy networks for speech recognition, *Expert Syst Appl* **30**:6368, 2009.
34. Shahamiri SR, Salim SSB, Artificial neural networks as speech recognisers for dysarthric speech: Identifying the best-performing set of MFCC parameters and studying a speaker-independent approach, *Adv Eng Inform* **28**:102, 2014.
35. Ting HN, Yong BF, Mirhassani SM, Self-adjustable neural network for speech recognition, *Eng Appl Artif Intell* **26**:2022, 2013.
36. Siniscalchi SM, Yu D, Deng L, Lee CH, Exploiting deep neural networks for detection-based speech recognition, *Neurocomputing* **106**:148, 2013.
37. Siniscalchi SM, Svendsen T, Lee CH, An artificial neural network approach to automatic speech processing, *Neurocomputing* **140**:326, 2014.
38. Jorgensen C, Dusan S, Speech interfaces based upon surface electromyography, *Speech Commun* **52**:354, 2010.
39. Schultz T, Wand M, Modeling coarticulation in EMG-based continuous speech recognition, *Speech Commun* **52**:341, 2010.
40. Denby B, Schultz T, Honda K, Hueber T, Gilbert JM, Brumberg JS, Silent speech interfaces, *Speech Commun* **52**:270, 2010.
41. Dileep AD, Sekhar CC, HMM based intermediate matching kernel for classification of sequential patterns of speech using support vector machines, *IEEE Trans Audio Speech Lang Process* **21**:2570, 2013.
42. Solera-Urena R, Garcia-Moral AI, Pelaez-Moreno C, Martinez-Ramon M, Diaz-De-Maria F, Real-time robust automatic speech recognition using compact support vector machines, *IEEE Trans Audio Speech Lang Process* **20**:1347, 2012.
43. Ganapathiraju A, Hamaker JE, Picone J, Applications of support vector machines to speech recognition, *IEEE Trans Signal Process* **52**:2348, 2004.
44. Matsumoto M, Hori J, Classification of silent speech using support vector machine and relevance vector machine, *Appl Soft Comput* **20**:95, 2014.
45. He X, Wang Z, Jin C, Zheng Y, Xue X, A simplified multi-class support vector machine with reduced dual optimization, *Pattern Recogn Lett* **33**:71, 2012.
46. Arifin NA, Tiun S, Predicting malay prominent syllable using support vector machine, *Procedia Technol* **11**:861, 2013.