Perspectives in Science (2016) **8**, 558–560

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/pisc

# Sub-vocal speech pattern recognition of Hindi alphabet with surface electromyography signal☆

**Munna Khan, Mosarrat Jahan***

*Department of Electrical Engineering, Jamia Millia Islamia, New Delhi, India*

Received 15 June 2016; accepted 15 June 2016
Available online 18 July 2016

**KEYWORDS**
Subvocal;
EMG;
Hindi phonemes;
AR coefficients

**Summary** Recently electromyography (EMG) based speech signals have been used as pattern recognition of phoneme, vocal frequency estimation, browser interface, and classification of speech related problem identification. Attempts have been made to use EMG signal for sub-vocal speech pattern recognition of Hindi phonemes क ख ग घ and Hindi words. That provides the command sub-vocally to control the devices. Sub-vocal EMG data were collected from more than 10 healthy subjects aged between 25 and 30 years. EMG-based sub-vocal database are acquired from four channel BIOPAC MP-30 acquisition system. Four pairs of Ag-AgCl electrodes placed in the participant neck area of skin. AR coefficients and Cepstral coefficients were computed as features of EMG-based sub-vocal signal. Furthermore, these features are classified by HMM classifier. H2M MATLAB toolbox was used to develop HMM classifier for classification of phonemes. Results were averaged on 10 subjects. An average classification accuracy of Ka is found to be 85% whereas the classification accuracy of Kha and Gha is in between 88% and 90%. The classification accuracy rate of Ga was found to be 78% which was lesser as compared to Kha and Gha.


## Introduction

Speech production starts from pressure exerted by lungs contraction and expansion. But subvocal speech produced by movement of articulatory muscles and larynx (Armas et al., 2014; Hofe et al., 2013; Denby et al., 2010; Jorgensen and Binstead, 2005; Jorgensen et al., 2003; Brunmberg et al., 2010). Electrical activity produced by movement of articulatory muscles is studied using electromyography (EMG) recording (Wand et al., 2014; Khan and Jahan, 2015; Zhou and Jiang, 2009; Lee, 2008; Heris et al., 2009; Scheme, 2007). The area of research in subvocal word or phonemes recognition has not a very long history; Jorgensen and Binstead (2005), Jorgensen et al. (2003) explored the recognition of subvocal speech using EMG signals recorded from neck area below the chin.

☆ This article belongs to the special issue on Engineering and Material Sciences.
* Corresponding author.
*E-mail address:* mosarrat.jahan@gmail.com (M. Jahan).

http://dx.doi.org/10.1016/j.pisc.2016.06.019

Wavelet based features were extracted from six subvocally spoken words and these isolated words were classified using neural network classifier. They reported the accuracy rate of 92% for six control words (stop, go, left, right, alpha, omega). Heris et al. (2009) suggested to identify the vocal fold disorders or diseases on the basis of feature extracted from wavelet sub-band. Two features energy and Shannon entropy were extracted from different sub-band of wavelet coefficients. Genetic algorithm, support vector machine (SVM) and k-nearest neighbour's (KNN) were applied to find discrimination of features. Highest classification accuracy of 91% was obtained by SVM classifier. Nonlinear features proved in finding the disorders of vocal fold.

The research paper presents a novel approach on subvocal Hindi phonemes pattern recognition based on surface electromyography. Electrical signals recorded from the area of neck below the chin by means of surface electrodes (Ag-AgCl) (Khan and Jahan, 2015). Recorded EMG data for subvocal speech are classified to get pattern recognition of phonemes spoken silently without any audible sound. The technique may be utilised in design and development of assistive devices for the persons who do not have control over speech, as well as for security personnel and military uses.

## Materials and methods

Sub-vocal speech samples are collected from EMG signals. Recording has been done on four channels BIOPAC System with MP30 Acquisition Unit. Ten subjects aged group from 24—35 years were recorded for four Hindi alphabets क ख ग घ . All the participants are informed about the protocol of the recording sessions. During EMG signal recording of subvocal speech, the participants were sitting in a silent room and asked to sign a consent form. Skin areas over the muscles of Participants are cleaned with acetone wetted swabs before applying the conductive electrode gel. Bipolar Ag-AgCl electrodes are placed in the area of neck approximately 30 mm below the chin. Recording of subvocal phonemes are divided into 10 trials. Each trial is recorded for 30 s. EMG signals corresponding to four Hindi phonemes Ka, Kha, Ga, Gha are acquired for subvocal speech with sampling frequency of 1000 Hz. Total 400 data sets are collected for these alphabets. EMG of each phoneme is recorded separately for 30 s in text file format. Artefacts presents in the signal were removed by pre-processing during data recordings. EMG signal recorded for each phoneme was transformed into suitable feature vector by signal processing with autoregressive (AR) and cepstral modelling (Khan and Jahan, 2015). Algorithms have been developed in MATLAB from signal processing to pattern recognition that provides complete performance measure of the subvocal phonemes. The information related to linear parameters of the signal is provided by AR modelling. In AR modelling the current sample of the signal in a data sequence $x(1), x(2) \ldots x(n)$ can be predicted as a linearly weighted sum of '$m$' most recent sample values $x(n-1), x(n-2) \ldots x(n-m)$ and an error term $\varepsilon(n)$. It can be expressed as:

$$x(n) = \varepsilon(n) - \alpha_1 x(n-1) - \alpha_2 x(n-2) - \cdots - \alpha_m x(n-m)$$
$$= \varepsilon(n) - \sum_{k=1}^{m} \alpha_k x(n-k) \quad (1)$$

Table 1 AR coefficients ($\alpha_m$) of subvocal Hindi phonemes for model order six.

| Model order ($m = 6$) | क | ख | ग | घ |
|---|---|---|---|---|
| $m = 1$ | −0.350 | −0.276 | −0.313 | −0.367 |
| $m = 2$ | 0.297 | 0.220 | 0.246 | 0.264 |
| $m = 3$ | 0.066 | 0.129 | 0.132 | 0.223 |
| $m = 4$ | 0.446 | 0.442 | 0.443 | 0.553 |
| $m = 5$ | 0.135 | −0.211 | −0.127 | −0.234 |
| $m = 6$ | 0.538 | 0.440 | 0.428 | 0.504 |

where $x(n)$ = output of the signal, $\varepsilon(n)$ = input of the signal; $\alpha_m$ = AR coefficients of $m$th model order.

Transfer function of Eq. (1) can be written as:

$$H(s) = \frac{X(s)}{\varepsilon(s)} = \frac{1}{1 + \sum_{k=1}^{m} \alpha_k s^{-k}} \quad (2)$$

The cepstrum of the signal can be calculated directly using Eqs. (3) and (4).

$$w_1 = -\alpha_1$$

$$w_n = -\alpha_n - \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) \alpha_k w_{n-k} \quad \text{for} \quad 1 < n \le m \quad (3)$$

$$w_n = -\sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) \alpha_k w_{n-k} \quad \text{for} \quad n > m \quad (4)$$

## Results and discussion

Total 400 data sets of EMG signals are collected from ten subjects for analysed of subvocal Hindi phonemes. There are 100 data sets of each phoneme are recorded from ten participants. Two feature AR coefficients and cepstral coefficients of model order six are extracted for each Hindi phoneme. Six AR coefficients $\alpha 1$—$\alpha 6$ are computed for each phoneme as shown in Table 1. Similarly six cepstral coefficients $w1$—$w6$ are calculated directly corresponding to AR coefficients as tabulated in Table 2. These coefficients are calculated for another nine subject's data sets. Similar results are obtained, but these results are not displayed due to lack of space. Feature vectors extracted for each phoneme are concatenated to form total features before application of the classifier input. Training and test data has been separated, 60% data sets of total feature sets are used to train the HMM classifier, whereas 40% of total feature sets are used for evaluation purposes. To evaluate the performance of classifier three parameters accuracy, sensitivity and specificity is calculated using confusion matrix.In order to evaluate the accuracy of various phonemes for ten subjects are calculated by confusion matrix and results were averaged as shown in Fig. 1. The percentage of average classification accuracy rate is found to be 85.43%, 78.67%, 75.24% and 80.53% for Ka, Kha, Ga, and Gha respectively. Analysing Fig. 2, we conclude that classification accuracy achieved

**Table 2** Cepstral coefficients *w*1—*w*6 for subvocal Hindi phoneme (क ख ग घ ).

| Hindi phonemes | *w*1 | *w*2 | *w*3 | *w*4 | *w*5 | *w*6 |
|---|---|---|---|---|---|---|
| क | 0.07952 | 0.42030 | 0.1838 | 0.27868 | 0.21897 | 0.054513 |
| ख | 0.43440 | 0.07505 | 0.2462 | 0.53844 | 0.22557 | 0.17557 |
| ग | 0.08518 | 0.37582 | 0.1891 | 0.25993 | 0.20431 | 0.09102 |
| घ | 0.09745 | 0.38943 | 0.12841 | 0.28938 | 0.18088 | 0.07821 |

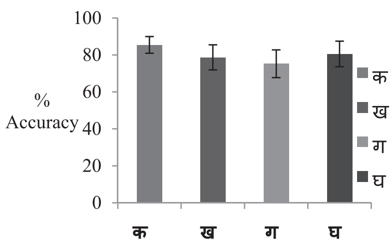


**Figure 1** Classification accuracy rate for subvocal phonemes averaged on 10 subjects.

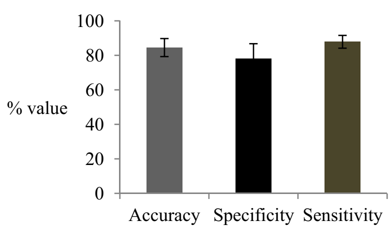


**Figure 2** Performance evaluation of features of subvocal EMG phoneme.

by HMM classifier for subvocal Hindi phonemes are discriminant. It is also observed that classification accuracies for phoneme Ka and Kha is similar. Performance of classifier is investigated by the confusion matrix in terms of accuracy. There is no single classifier that works best on all given features. Therefore various tests have been performed to compare the performance of classifier. The accuracy was also affected by model training and variability in the data.

## Conclusion

This paper explores the analysis of subvocal EMG signal. EMG signals of subvocal Hindi phonemes are processed to extract two features for recognition of phonemes. AR coefficients and cepstral coefficients feature vector have been passed through hidden Markov model (HMM) classifier for classification of phonemes. The results indicated that the cepstral coefficients feature with HMM classifier gives accuracy rate of 85.56%. Therefore HMM classifier proved to be important in recognition of subvocal Hindi phonemes. In future we develop a subvocal phoneme based EMG speech recogniser for security purpose.

## References

Armas, W.D., et al., 2014. Vocal frequency estimation and voicing state prediction with surface EMG pattern recognition. Elsevier Speech Commun. 63 (64), 15—26.

Brunmberg, J.S., et al., 2010. Brain—computer interface for speech communication. Elsevier Speech Commun. 52, 367—379.

Denby, B., et al., 2010. Silent speech interfaces. Elsevier Speech Commun. 52, 270—287.

Heris, H.K., et al., 2009. Optimal feature selection for the assessment of vocal fold disorders. Elsevier Comput. Biol. Med. 39, 860—868.

Hofe, R., et al., 2013. Small-vocabulary speech recognition using a silent speech interface based on magnetic sensing. Elsevier Speech Commun. 55 (1), 22—32.

Jorgensen, C., Binstead, K., 2005. Web browser control using EMG based sub vocal speech recognition. In: Proc. International Conference on System Sciences, pp. 294c—301c.

Jorgensen, C., et al., 2003. Sub auditory speech recognition based on EMG/EPG signals, NASA Ames research centre under the CICT/ITSR program.

Khan, M., Jahan, M., 2015. The application of AR coefficients and Burg method in sub-vocal EMG pattern recognition. J. Basic Appl. Eng. Res. 2 (10), 813—815.

Lee, K.S., 2008. EMG-based speech recognition using hidden Markov models with global control variables. IEEE Trans. Biomed. Eng. 55 (3), 930—940.

Scheme, E.J., 2007. Myoelectric signal classification for phoneme based speech recognition. IEEE Trans. Biomed. Eng. 54, 694—699.

Wand, M., et al., 2014. Tackling speaking mode varieties in EMG-based speech recognition. IEEE Trans. Biomed. Eng. 61 (10), 2515—2526.

Zhou, Q., Jiang, N., 2009. Improved phoneme-based myoelectric speech recognition. IEEE Trans. Biomed. Eng. 56, 2016—2023.