INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2022, VOL. 68, NO. 1, PP. 161–166
Manuscript received August 29, 2021; revised January, 2022. DOI: 10.24425/ijet.2022.139856

Effect of Time-domain Windowing on Isolated Speech Recognition System Performance

Ananthakrishna Thalengala, Anitha H and Girisha T

Abstract*—Speech recognition system extract the textual data from the speech signal. The research in speech recognition domain is challenging due to the large variabilities involved with the speech signal. Variety of signal processing and machine learning techniques have been explored to achieve better recognition accuracy. Speech is highly non-stationary in nature and therefore analysis is carried out by considering short time-domain window or frame. In the speech recognition task, cepstral (Mel frequency cepstral coefficients (MFCC)) features are commonly used and are extracted for short time-frame. The effectiveness of features depend upon duration of the time-window chosen. The present study is aimed at investigation of optimal time-window duration for extraction of cepstral features in the context of speech recognition task. A speaker independent speech recognition system for the *Kannada* language has been considered for the analysis. In the current work, speech utterances of *Kannada* news corpus recorded from different speakers have been used to create speech database. The hidden Markov tool kit (HTK) has been used to implement the speech recognition system. The MFCC along with their first and second derivative coefficients are considered as feature vectors. Pronunciation dictionary required for the study has been built manually for mono-phone system. Experiments have been carried out and results have been analyzed for different time-window lengths. The overlapping Hamming window has been considered in this study. The best average word recognition accuracy of 61.58% has been obtained for a window length of 110 msec duration. This recognition accuracy is comparable with the similar work found in literature. The experiments have shown that best word recognition performance can be achieved by tuning the window length to its optimum value.

***Keywords*—Hidden Markov model (HMM), Isolated speech recognition (ISR) system, *Kannada* language, Mono-phone model, Mel frequency cepstral coefficients (MFCC).**

I. Introduction

SPEECH is the general mode of communication between human beings. To be able to establish better human-machine interaction, speech has always intrigued mankind. Many researches have been done to create better speech recognition engine. Speech recognition is the process of translation of speech into textual information. To develop an accurate automatic speech recognition is still a difficult task because of numerous variability present in speech. Few of the important factors which are affecting the recognition precision are rate of speech utterance, speaker to speaker variations, recording environment, background noise, size of vocabulary and so on. Many of the research works have been carried out using European languages such as English, Spanish, and German to build automatic speech recognition system. But the works using the Indian languages are very limited [1] [2] [3] [4] [5]. There are more than 200 scriptable languages in India with 22 official languages identified by the government. The *Kannada* language is one among them. The most of the Indian languages are syllable timed and there exists one to one mapping among its pronunciation and written scripts [6] [7]. Researchers have also recently addressed on the multilingual phone recognition for Indian languages [8] [9]. In our earlier studies, we have attempted speech recognition work on *Kannada* language [10] [11] [12].

The challenges in isolated word recognition task can be viewed as speech feature analysis related at the front-end and pattern recognition at the back-end. The ceptral domain features and hidden Markov models (HMM) have become standards for front-end and back-end respectively for the speech recognition system. The choice of analysis window (time-window length) considered play important role in effectiveness of chosen feature. This study is aimed to bring-out the optimal time-window length for the extraction of cepstral parameters in the context of speech recognition task. In the literature it can be seen that about 2 to 3 pitch duration is taken to be the length of time-window [13] [14] [15]. To obtain better time resolution it is desirable to consider larger analysis window lengths. However, lager length window suffers from poor frequency resolution. The experiments with different window lengths have been performed and recognition results are analyzed.

In the present work, automatic speech recognition (ASR) system has been built using the *Kannada* language speech database. *Kannada* is one of the Dravidian languages of India with the history of more than 2000 years. There are more than million *Kannada* speaking peoples present inside and outside of Karnataka state. *Kannada* is a syllable-timed language having 52 phonemic letters (called "Akshara Maala") which are basically evolved from the *"Kadamba"* script. The alphabets of *Kannada* are grouped into three categories namely, vowels (Swaragalu), consonants (Vyanjanagalu) and letters which are

Ananthakrishna Thalengala is with Department of Electronics and Communication Engineering, Manipal Institute of Technology (MIT), Manipal Academy of Higher Education (MAHE), Manipal, Karnataka State, India (e-mail: anantha.kt@manipal.edu .

Aitha H is with Department of Electronics and Communication Engineering, Manipal Institute of Technology (MIT), Manipal Academy of Higher Education (MAHE), Manipal, Karnataka State, India (e-mail: anitha.h@manipal.edu .

Girisha T is with Department of Electronics and Communication Engineering, Manipal Institute of Technology (MIT), Manipal Academy of Higher Education (MAHE), Manipal, Karnataka State, India (e-mail: grsh.246@gmail.com

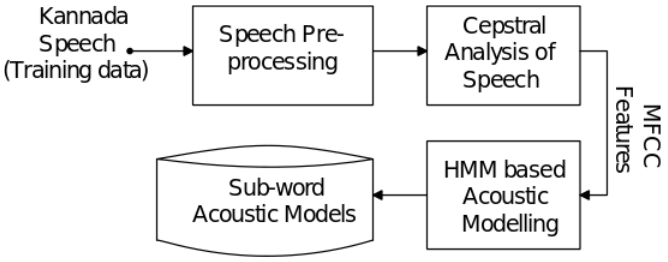
Kannada Speech (Training data)
Speech Pre-processing
Cepstral Analysis of Speech
MFCC Features
HMM based Acoustic Modelling
Sub-word Acoustic Models

Fig. 1. Acoustic model building steps: The training phase.

neither consonant nor vowel (Yogavaahakagalu) [16] [17] [18]. The *Kannada* phones and their corresponding English labels are shown in Table I. The arrangement of alphabets in the language is well structured and is as per the place of articulations. In *Kannada* language, the alpha-syllabary units are very stable and have unique pronunciations which are independent of their occurrence in a word or sentence.

The design of ASR system needs careful attentions to pre-processing, feature extraction and pattern recognition stages. The pre-processing step mainly involve time-windowing and pre-emphasis. The Mel frequency cepstral coefficients (MFFCs) have been successfully used as features in speech recognition task and are considered in this study. The hidden Markov models (HMMs) are proved to be good statistical models to represent the time series events such as speech signal. As a counter-part, neural network based phone recognition models can be found in literature [19] [20]. However due to success of statistical models for speech recognition task, HMMs are considered in this work. In the present work, HTK software has been used to implement and develop the *Kannada* ASR system [21].

II. Speech recognition system

The basic speech recognition model is shown in Figure 1 and Figure 2. ASR system consists of the steps which include, acquisition of speech signal, pre-processing, analysis of cepstral coefficients and recognition of word. Initial three steps represents the front-end of ASR system. HMM is used to build the speech word model and recognizing of words is done using Viterbi decoding algorithm. The pre-processing step make sure better quality of the captured speech data against recording levels and noise. Analysis of the speech processing is done in cepstral domain to extract cepstral coefficients. MFCC coefficients are commonly used in ASR systems to yield a better performance [22]. Pattern classifier is the important building block of the ASR system. This has been achieved by selecting the HMM as a statistical model. The ASR system comprises of two phases namely training (refer Figure 1) and testing (refer Figure 2) phase. In training process, by using the known training data HMM based acoustic models are created and during the testing phase, to calculate the performance accuracy of the ASR system, unknown speech samples are applied. Accuracy of the ASR system is evaluated from the percentage of words recognized.

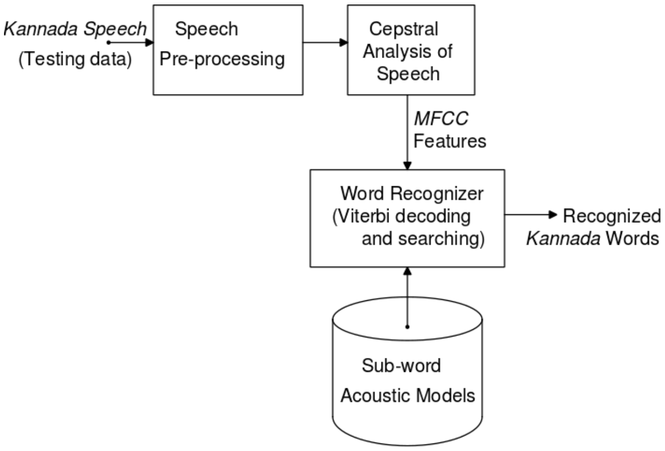
Kannada Speech
(Testing data)
Speech Pre-processing
Cepstral Analysis of Speech
MFCC Features
Word Recognizer (Viterbi decoding and searching)
Recognized *Kannada* Words
Sub-word Acoustic Models

Fig. 2. Speech recognition steps: The testing phase

A. *Speech pre-processing*

The speech signal pre-processing is having pre-emphasis and windowing techniques. A high pass filter is used as a pre-emphasis filter and it flattens the speech signal. Later, speech signal is sliced and segmented into overlapping time frames called framing or windowing. Speech signal is non-stationary in nature and therefore framing (or windowing) play important role in deciding effectiveness of the feature set considered. While choosing the window frame, two considerations are important, length of the window and type of the window. Smaller the length of the window better will be time resolution and lager the length of the window yields good spectral resolution [13]. Typically length of the window of about 2 to 3 pitch periods will be used in the analysis of speech [15]. In the current work, hamming window of different duration are used.

B. *Cepstral analysis of speech*

Feature extraction play an important role in the performance of the speech recognition. Feature extraction process gives different parameters of the speech signal. The present work considers MFCC as a feature vector for the speech recognition task. Human audible capability (both frequency range and intensity range) is basically logarithmic in nature. The Mel-Scale filter bank is used to characterize the human ear perceiveness of frequency. Mel frequency cepstrum represents the short term power spectrum of a sound obtained by applying cosine transform on log of the power spectrum on the nonlinear Mel scale. MFCC features have been successfully used in speech recognition task to yield better results [22]. In the present work, 12-MFCC parameters, single log energy value, 13 delta and 13 delta-delta(acceleration) coefficients are used. So the total length of feature vector becomes 39.

C. *Hidden Markov model*

Hidden Markov Model (HMM) is generally used for acoustical model of speech sounds. It is a Markov process consisting of hidden states. It is a double stochastic process and hidden

TABLE I
PHONES IN *Kannada* LANGUAGE.

Label	***Kannada* Phone**	Label	***Kannada* Phone**	Label	***Kannada* Phone**	Label	***Kannada* Phone**
a	ಅ	Au	ಔ	T	ಟ್	bh	ಭ್
A	ಆ	M	ಅಂ	Th	ಠ್	m	ಮ್
i	ಇ	H	ಅಃ	D	ಡ್	y	ಯ್
I	ಈ	k	ಕ್	Dh	ಢ್	r	ರ್
u	ಉ	kh	ಖ್	N	ಣ್	l	ಲ್
U	ಊ	g	ಗ್	t	ತ್	v	ವ್
ru	ಋ	gh	ಘ್	th	ಥ್	sh	ಶ್
rU	ೠ	ng	ಙ್	d	ದ್	S	ಷ್
e	ಎ	c	ಚ್	dh	ಧ್	s	ಸ್
E	ಏ	ch	ಛ್	n	ನ್	h	ಹ್
ai	ಐ	j	ಜ್	p	ಪ್	L	ಳ್
o	ಒ	jh	ಝ್	ph	ಫ್	kSh	ಕ್ಷ್
O	ಓ	ny	ಞ	b	ಬ್	jn	ಜ್ಞ

state is estimated by using set of processes which gives observation sequence. HMM complete set of model is represented by equation (1).

$$\lambda = (X, Y, \Pi) \tag{1}$$

Where, X is transition state probability matrix, Y denotes output probability matrix and π will be initial state probability vectors [14]. If L denotes the vocabulary size of *Kannada* words considered and to find closest matching between word's observation sequences is given by equation (2) below.

$$O = \{o_1, \quad o_2, \quad . \quad . \quad . \quad . \quad o_M\} \tag{2}$$

Let vector W denote words to be recognized, which are represented by sub-word units is as shown in equation (3) below.

$$W = \{sw_1, \quad sw_2, \quad . \quad . \quad . \quad . \quad sw_N\} \tag{3}$$

Out of L words, the most likely word estimate for a given observation sequence O is given by equation (4) is as follows.

$$\widehat{W} = \max_{W\epsilon L} [P(W/O)] \tag{4}$$

By using Baye's rule equation (4) above can be re-written as given in equation (5) below.

$$\widehat{W} = \max_{W\epsilon L} [P(O/W)P(W)] \tag{5}$$

Where, $P(W)$ is the prior probability of a specific model. For any word model $W, P(O/W)$ is the probability of observation sequence O. The above said mathematical steps can be summarized into training phase and testing phase as described by the following two steps.

1) Separate HMM models has been created for each phones in the vocabulary. This step is considered as the training phase.
2) Second step is considered as testing phase; here HMM models are utilized to recognize every unidentified word in a given test database. In the recognition phase, Viterbi decoding algorithm is applied to get best matching word.

The available data set has been grouped and are used for training and testing of the system, which is given in detail under section III-A. The recognition results are quantified interms of recognition accuracy and system performance has been evaluated.

III. Implementation of ASR system using HTK

Hidden Markov Tool Kit (HTK version 3.4.1) is one of the benchmark software tool in speech recognition and is used to implement ASR system for *Kannada* language [21]. The implementation process consists primarily preparation of data, coding of data, creating acoustic model by utilizing HMM and evaluation of ASR system performance. In the data preparing stage, speech signal is recorded in the regulated surroundings and pre-processed. By using pronunciation dictionary which consists of phones and labels corresponding to each words, the vocabulary of the system is described. In the speech analysis step (data-coding), cepstral domain features are computed. The appropriate specifications such as window length, window type, frame rate and other additional parameters are set during the speech analysis. In this work, hamming window of different window lengths are used to evaluate the system performance. HMM's for the sub-word units are described and feature vectors are applied to re-estimate and to produce the statistical models. The testing data groups are used in evaluation of recognition accuracy of HMMs.

A. Speech database

Recording of speech signal is done by using good quality audio recording device in a sound proof environment. The recorded speech is therefore has high signal-to-noise (SNR) and require minimum pre-processing steps. The speech is acquired and recorded at 10 KHz sampling, 16-bit PCM, and stored in wave file format. In the current work, speech corpus is created by recording the *Kannada* broadcast news voiced by twelve distinct speakers of which, five female and seven male speakers. The recorded database is having 4921 *Kannada* words in which 735 words are distinct (vocabulary size). So the dictionary size becomes 735 words. The segmentation of word is manually carried out. The available speech database has been divided in to four groups for the experimental purpose. The four data groups named G1, G2, G3 and G4 consists of 1244, 1233, 1225 and 1219 *Kannada* words respectively. The well-known hold-out method has been adapted for choosing the data sets for training and testing. One of the four group is used for testing, while other three groups are for training the system. The *Kannada* language phone sets are depicted by English alphabets shown in the Table I.

Another task in database development is dictionary building. The simple phone dictionary (also known as pronunciation dictionary) is considered in this study. The phone dictionary for the chosen *Kannada* words are obtained by representing each word as a sequence of phones. A few examples of *Kannada* words represented as series of phones are shown in Table II. The standard phone dictionary is not available for *Kannada* language and therefore dictionary preparation is also a contribution in this work.

B. Feature analysis

Block diagram of the HTK speech coding stage is shown in Figure 3. The 'HCopy' command in HTK extracts MFCC parameters for the given speech input as per the given specifications. The specifications mainly consists of input speech sampling rate, window type, length of window, frame rate, and feature details. Each feature vector consists of one log energy value, 12 MFCC values, 13 first derivative values (delta coefficients), and 13 second derivative values (delta - delta coefficients). So the total size of output feature vector becomes 39. The length of analyzing window (window duration in msec) is kept different for each experiment and results are analyzed based on this. The window length varying between 20 msec and 150 msec are considered in this study. The details of the window lengths considered in the study are given in section IV. The speech parameterization is carried out for the entire available speech database, which include training and testing data samples. The feature vectors obtained from the training data samples are used to construct the HMMs whereas feature vectors obtained from the testing data samples are used for the evaluation of models.

TABLE II
Example: Phone sequence for *Kannada* words.

Kannada Word	**English Meaning**	**Sequence of Phones**
ಸಹಕಾರ (sahakara)	Help	s a h a k A r a
ಮೂಲಕ (moolaka)	Through	m U l a k a
ಸಾವಿರ (savira)	Thousand	s A v i r a
ಅವರು (avaru)	They	a v a r u
ನೂತನ (noothana)	New	n U t a n a

C. Generating HMM

The acoustical models are basically stochastic models which define the sub-word unit. The HMMs have been proved to be good statistical models to represent time-sequence events such as speech units [23] [24]. The acoustic models for *Kannada* phones are built using HMMs. Each phone in *Kannada* language is therefore represented by HMM and such acoustical models are known as mono-phone models. Monophone models are developed with the assumption that each phoneme is an independent acoustic unit and has no acoustic coupling with adjacent phonemes. However, this is not true in a continuous utterances of words as there exists co-articulation affect. So, the designed acoustical models in this study are not

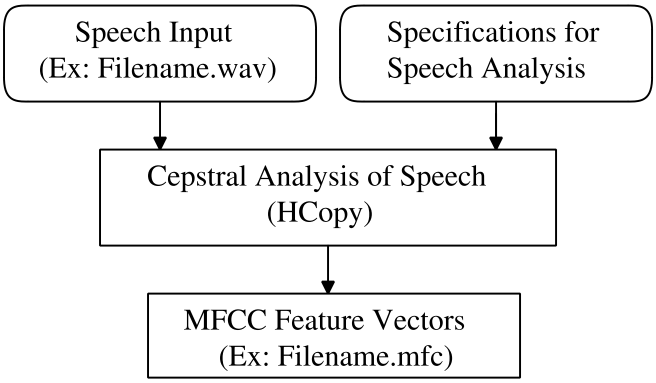

Fig. 3. Speech coding steps in HTK

optimized to explore acoustic coupling between the adjacent phones. But the objective is to obtain the optimal time-window length to be considered to extract cepstral features for acoustic modeling.

The HMMs are designed with five state with initial and end states. The middle three states describe the phone considered. We have identified 52 phones for *Kannada* language (refer Table I) and are modeled with 5-state HMM. Each state of HMM is Gaussian distributed and are constructed by using MFCC features from the training speech data. In HTK, the 'HRest' function is used to estimate and re-estimate HMMs.

The recognition performance is evaluated using the testing data set. The Viterbi decoding algorithm is applied to obtain the optimal phone sequence (optimum path) and thereby choosing the matching word from the dictionary. The 'HVite' command in HTK is used to implement the pattern matching. Analysis of result ('HResults' command in HTK) is carried out using word recognition accuracy (expressed in percentage). The word recognition accuracy is defined and elaborated in the next section IV.

IV. Results and Discussions

Performance evaluation of the system is carried-out with the help of word recognition accuracy. The word recognition accuracy of a isolated word recognition system can be defined as given by equation (6) below.

$$Percentage\ of\ Accuracy = \frac{M-N}{M} \times 100\% \qquad (6)$$

Where M represents the total number of words in the test database and N is the number of words misclassified.

The speech database consists of utterances of *Kannada* words taken from news corpus. The database has been divided into four groups to perform the experiments and the details of data grouping are given in section III-A. The experiments are repeated by varying the cepstral analysis time-window length between 20 msec and 150 msec. In every recognition experiment, one group is used for testing the system whereas the other 3 groups are used for training. This procedure is repeated for all the data groups and average word recognition accuracy is obtained. The average recognition accuracies for various analysis window lengths are plotted in Figure 4. The best average word recognition accuracy is obtained when the analysis time-window length is 110 msec. The detailed performance of the system for window length of 110 msec is given in Table III.

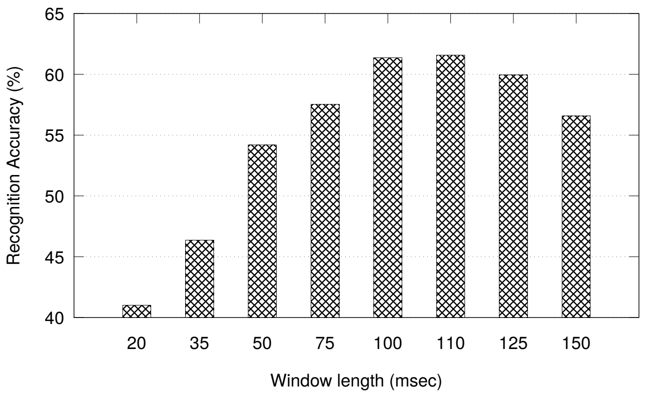

Fig. 4. Performance of ISR system against window length

TABLE III
Performance of ISR system for 110 msec window

Trial No.	Training Groups	Testing Groups	Recognition Accuracy (%)
1	G2, G3, G4	G1	61.25
2	G1, G3, G4	G2	60.42
3	G1, G2, G4	G3	59.51
4	G1, G2, G3	G4	65.14
Average			61.58

It is observed that the recognition performance of the system depends on the analysis time-window length considered. The shorter length window provide better time resolution, whereas larger length window gives better frequency resolution. Here, cepstral analysis of speech segment is performed at the front-end of the ISR system and larger length analysis window is desired for the better representation. Hence the best results are obtained when the window-length of 110 msec duration is considered. And further increase in analysis window length leads to decrease in the performance due to the averaging effect of cepstral features.

The best average recognition accuracy of 61.58% has been obtained which can be comparable with the similar work reported in the literature. The poor results are due to the context independent mono-phone acoustic models considered in the study. Therefore the context independent tri-phone models can be used to further improve the performance. Perhaps the present study is focused on optimizing the time-window length for cepstral analysis.

V. Conclusion

In the present work, HTK toolkit has been used to implement the isolated speech recognition system for *Kannada* language. The main objective of the study is to investigate and optimize the time-window duration for the best performance of speech recognition system. The cepstral features considered in the speech recognition system have been analyzed for time-window length varying from 20 msec to 150 msec. The best word recognition accuracy of 65.14% has been obtained when time-widow length of 110 msec is used. The choice of length of time-window is based on the time resolution and frequency resolution required. The required time and frequency resolution depends on the features considered and intended application. In the current study, it is found that time-window length of 110 msec considered for cepstral feature extraction gives optimal performance for the *Kannada* isolated word recognition system. Hence it is concluded that choice of optimal time-window length is critical to performance of speech recognition system. The analysis can be further extended for speech recognition system based on tri-phone model to obtain higher recognition accuracy.

References

[1] Bharali, S. S., & Kalita, S. K., ”A comparative study of different features for isolated spoken word recognition using HMM with reference to Assamese language”. International Journal of Speech Technology, 18(4), 673-684 (2015). https://doi.org/10.1007/s10772-015-9311-7

[2] Kumar, K., Aggarwal, R. K., & Jain, A., ”A Hindi speech recognition system for connected words using HTK”, International Journal of Computational Systems Engineering, 1(1), 25-32 (2012). https://doi.org/10.1504/IJCSYSE.2012.044740

[3] Thangarajan, R., Natarajan, A. M., & Selvam, M., ”Syllable modeling in continuous speech recognition for Tamil language”, International Journal of Speech Technology, 12(1), 47-57 (2009). https://doi.org/10.1007/s10772-009-9058-0

[4] Dua, M., Aggarwal, R. K., Kadyan, V., & Dua, S., ”Punjabi automatic speech recognition using HTK”, IJCSI International Journal of Computer Science Issues, 9(4), 1694-0814 (2012).

[5] Hegde, S., Achary, K., & Shetty, S., ”Statistical analysis of features and classification of alphasyllabary sounds in Kannada language”, International Journal of Speech Technology, 18(1), 65–75 (2015). https://doi.org/10.1007/s10772-014-9250-8

[6] Panda, S. P., & Nayak, A. K., ”Automatic speech segmentation in syllable centric speech recognition system”, International Journal of Speech Technology, 19(1), 9-18 (2016). https://doi.org/10.1007/s10772-015-9320-6

[7] Thangarajan, R., Natarajan, A., & Selvam, M., ”Syllable modeling in continuous speech recognition for Tamil language”, International Journal of Speech Technology, 12(1), 47–57 (2009). https://doi.org/10.1007/s10772-009-9058-0

[8] Manjunath, K. E., Jayagopi, D. B., Rao, K. S., & Ramasubramanian, V. (2019), ”Development and analysis of multilingual phone recognition systems using Indian languages”, International Journal of Speech Technology, 22(1), 157-168. https://doi.org/10.1007/s10772-018-09589-z

[9] Kumar, C. S., & Mohandas, V. P. (2011), ”Robust features for multilingual acoustic modeling”, International Journal of Speech Technology, 14(3), 147-155. https://doi.org/10.1007/s10772-011-9092-6

[10] Ananthakrishna, T., Maithri, M., & Shama, K., ”Kannada word recognition system using HTK”, In 2015 Annual India Conference, INDICON, New Delhi, India , pp. 1-5, (2015, December). https://doi.org/10.1109/INDICON.2015.7443122

[11] Thalengala, A., & Shama, K., ”Study of sub-word acoustical models for Kannada isolated word recognition system”, International Journal of Speech Technology, 19(4), 817-826, (2016).

[12] Thalengala Ananthakrishna, Kumara Shama, and Maithri Mangalore, ”Performance Analysis of Isolated Speech Recognition System Using Kannada Speech Database”, Pertanika Journal of Science & Technology 26.4 (2018). https://doi.org/10.1007/s10772-016-9374-0

[13] Rabiner, L. R., Juang B. H., & Yegnanarayana B., ”Fundamentals of speech recognition”, Englewood Cliffs: PTR Prentice Hall (2012).

[14] Rabiner, L. R., ”A tutorial on hidden Markov models and selected applications in speech recognition”, Proceedings of the IEEE, 77(2), 257-286 (1989). https://doi.org/10.1109/5.18626

[15] Deller J. R., Proakis J. G. & Hansen J. H. L., ”Discrete Time Processing of Speech Signals”, New York: Macmillan Publishing Company, (1993).

[16] Shridhara, M. V., Banahatti, B. K., Narthan, L., Karjigi, V., & Kumaraswamy, R. (2013, November), ”Development of Kannada speech corpus for prosodically guided phonetic search engine”, In 2013 international conference oriental COCOSDA held jointly with 2013 conference on Asian spoken language research and evaluation (O-COCOSDA/CASLRE) (pp. 1-6). IEEE. https://doi.org/10.1109/ICSDA.2013.6709875

[17] Krishnamurti, B., ”The Dravidian Languages”, Cambridge: Cambridge University Press, (2003).

[18] Steever, S. B., ”The Dravidian languages. London: Routledge Publications, (2015).

[19] Akhmetov, B., Tereykovsky, I., Doszhanova, A., & Tereykovskaya, L. (2018), ”Determination of input parameters of the neural network model, intended for phoneme recognition of a voice signal in the systems of distance learning”, International Journal of Electronics and Telecommunications, 64(4), 425-432. https://doi.org/10.24425/123541

[20] Kumar, R. S., & Lajish, V. L. (2013), ”Phoneme recognition using zerocrossing interval distribution of speech patterns and ANN”, International Journal of Speech Technology, 16(1), 125-131. https://doi.org/10.1007/s10772-012-9169-x

[21] Young S., Evermann G, Gales M., Hain T., Kershaw D., Liu, ”The HTK book (Vol. 2)” Cambridge: Entropic Cambridge Research Laboratory.

[22] Davis, S., & Mermelstein, P., ”Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences”, IEEE Transactions on Acoustics, Speech, and Signal Processing, 28(4), 357-366 (1980). https://doi.org/10.1109/TASSP.1980.1163420

[23] Nilsson, M., ”First Order Hidden Markov Model: Theory and implementation issues”, Technical Report, 2005:02. Blekinge Institute of Technology.

[24] OShaughnessy, D., ”Automatic speech recognition: History, methods and challenges”, Pattern Recognition, 41(10), 2965–2979 (2008).