International Journal of Control Theory and Applications

ISSN : 0974-5572

Volume 10 • Number 28 • 2017

Speaker Independent Recognition of Spoken Punjabi Alphabets, Vowels and Auxiliary Signs

Puneet Mittal[a] and Navdeep Singh[b]

[a]*CSE & IT Department, BBSBEC, Fatehgarh Sahib. Email: pmittal.cse@gmail.com*
[b]*Department of Computer Science, Mata Gujri College, Fatehgarh Sahib. Email: navdeep_jaggi@yahoo.com*

Abstract: Automatic speech recognition of Punjabi alphabets, vowels and auxiliary signs has been explored in this paper. Being a tonal language having acoustic similarity of letters makes the recognition of Punjabi alphabets and dependent vowels very difficult. Presence of pitch contour changes the meaning of the word depending on the way it sounds. This paper aims to develop a model for recognizing the Punjabi alphabets, vowels and auxiliary signs spoken by the user. For developing the acoustic model pronunciation dictionary has been prepared and training is done using CMU sphinxtrain tool. For this speech data of 25 speakers (13 females and 12 males) is considered. After that the developed model is tested on speech data of 9 speakers (5 females and 4 males). The observed recognition accuracy for male and female data is 81.6% and 84% respectively.

Index Terms: Alphabet Recognition, Automatic speech recognition, Punjabi language, CMU tools.

1. INTRODUCTION

Automatic speech recognition (ASR) is a process in which human speech is converted into text. ASR is being widely used in the development of smart devices, smart voice user interfaces, applications for people with disabilities, military, robotics and various other fields. Various spoken languages are currently under investigation and models have been developed for popular languages like English[17], Chinese[22], Japanese[30], Arabic[9], [29], [36] etc. Indian languages have been slightly neglected for ASR. Significant work is being done in Hindi[32], Bengali[15] and Tamil [34] languages.

Punjabi[21], [25] is one of the popular languages of India. Being a tonal language having acoustic similarity of letters makes the recognition of Punjabi alphabets and dependent vowels very difficult. Presence of pitch contour changes the meaning of the word depending on the way it sounds. There are three types of tones: low, high and level[24], [26]. The low tone is characterized by lowering the voice below the normal pitch and then rising back in the following syllable. In high tone the pitch of the voice rises above its normal level falling back at the following syllable. The level tone is carried by the remaining words. Tones are not represented by any letters or symbols in the Gurmukhi script. Letters ਘ, ਝ, ਢ, ਧ and ਭ are of special significance in the context of

tone. Letters like ਦ and ਧ, ਕ and ਖ, ਜ and ਝ, ਡ and ਢ are often misinterpreted while speaking them. Accurate recognition requires system to perform phonetic distinction with ability to recognize the letter even if there is minimal sound/ phonetic difference. This paper deals with speech based recognition of Punjabi alphabets, vowels and auxiliary signs.

The paper is organized as follows, section II describes the literature survey, section III presents brief description of Punjabi language, section IV explains Punjabi Speech Recognition System, section V discusses experimental results and section VI provides conclusions.

2. LITERATURE REVIEW

Various researchers have contributed their work for the recognition of spoken alphabets. Cole et. al., [13] developed English Alphabet Recognizer (EAR) to perform speaker-independent recognition of letters spoken in isolation. They developed routine for signal processing that transforms the digitized speech into useful representations, for locating segment boundaries, for feature measurements on the speech segments, and for classification of the letter. The system was trained from 120 speakers on every token of each letter. When tested on a new set of 30 speakers, the performance of the system was 95% and 96% respectively for two different tokens of each letter.

Karnjanadecha et. al., [20] proposed Signal Modeling for High Performance Robust Isolated Word Recognition system for English language. Accuracy achieved by the system for speaker independent alphabet recognition was 97.9%.

Cosi et. al., [14] designed and tested a telephone bandwidth speaker-independent continuous digit recognizer which was completely based on an artificial neural network and got 99.92% word accuracy rate and a 92.62% accuracy rate for sentence recognition. They developed and implemented speaker recognition system based on Hidden Markov Model (HMM)[37].

Loizou et. al., [23] developed recognizer for English alphabets that incorporates a series of new approaches to tackle the problems associated with the confusions occurring between the stop consonants in the E-set and the confusions between the nasals. One feature representation and two subspace approaches are proposed for improved stop consonant discrimination and nasal discrimination. A telephone alphabet recognizer was also developed using context-dependent HMM's having 91.7% correct letter recognition with 1.1% letter insertions.

Fanty et. al., [16] developed English alphabet recognition system with telephone speech using auditory model analysis, explicit segmentation and speech knowledge. The recognition system classified letters of the English alphabet produced by any speaker over telephone lines at 89% accuracy for spelled alphabets and retrieves names from a list of 50,000 with 91% first choice accuracy.

Thakuria et. al., [33] proposed Automatic Speech Recognition of BODO Alpha digits using HMM. They performed ten different observations on GU_Bodo corpus database (Bodo alphabets and digits). They achieved accuracy of 90.60%, 70.17% and 81.20% respectively for correct digit recognition, alphabet recognition and mixed alphabets and digits recognition in the noisy environment using mixed training and testing subsets.

Adam et. al., [8] proposed Spoken English Alphabet Recognition with Mel Frequency Cepstral Coefficients and Back Propagation Neural Networks for the E-set and for 26 English alphabets. They found that one of the confusable set is called the E-set letters which consist of the letters B, C, D, E, G, P, T, V and Z. By adjusting these parameters they managed to achieve 62.28% and 70.49% recognition rate for E-set recognition under speaker-independent and speaker- dependent conditions respectively.

Satori et. al., [28] trained model for Amazigh language. Corpus consisted of Amazigh_Alphadigits having 10 digits and 33 Amazigh alphabets. They used CMU Sphinx tools for developing the model and recognition. The system showed performance of 92.89% when trained using 16 Gaussian Mixture Model.

Kumar and Singh [39] proposed an Automatic Spontaneous Speech Recognition for Punjabi Language Interview Speech Corpus trained with 1227 Punjabi words and 461 Punjabi sentences and achieved recognition accuracy of 98.8% for Punjabi words and 98.6% for Punjabi sentences.

Major Contribution: The contribution of previous researchers remained limited to recognition of trained words only. Their systems remained unable to recognize those words which are not stored in the speech corpus during training. Our work focuses on recognition of Punjabi alphabets, vowels and auxiliary signs so that words can be formed from these constituents. As words are made from alphabets, vowels and auxiliary signs that are limited in number, their recognition may help in recognition of all possible words with limited training. Such type of model can assist for applications like teaching, automated phone book, directory to retrieve information like names, addresses etc.

3. PUNJABI LANGUAGE

Punjabi language is based on the principle of "one sound one symbol". Still there are few symbols which have very similar sound and are difficult to differentiate. Usually non Punjabi speakers mispronounce them. Punjabi is spoken as a native language, second language, or third language by about 30 million people in India. The Punjabi language is written in Gurumukhi[18] and Shahmukhi[31] scripts, making it one of the relatively few languages written in more than one script. In India, Punjabi is written in the Gurmukhi script in offices, schools, and media. Gurumukhi is considered as the standard script for Punjabi whereas Shahmukhi script is used in Pakistan. In this paper Gurmukhi script is explored. The Gurmukhi script has 41 letters including 38 consonants and 3 basic vowel sign bearers. There are 10 clear vowel signs and 3 auxiliary signs. Out of 10 vowels Mukta (ਮੁਕਤਾ) is invisible.

4. PUNJABI SPEECH RECOGNITION SYSTEM

This section describes formation and development of Punjabi speech recognition system using CMU Tools. There are two modules in system development: training and testing. Figure 1 and Figure 2 show the training and testing modules respectively.

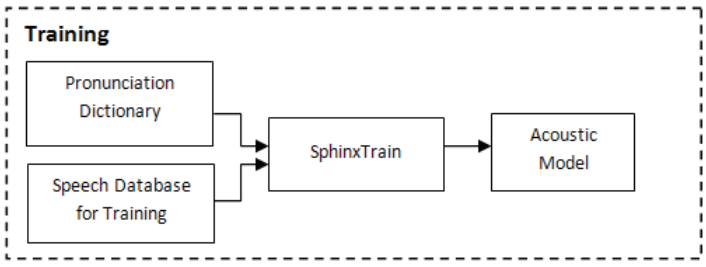

Figure 1: Training Module

Speech Database and Pronunciation Dictionary are pre-requisites for training. During training, CMU SphinxTrain tool builds acoustic model from the pronunciation dictionary and speech database. This acoustic model is used during testing of the system. Testing starts with the creation of language model using language modelling toolkit. Then, speech database (different from training) is created for testing. These language model and speech database along with the acoustic model and pronunciation dictionary created during training are given to the decoder for evaluation of accuracy and word error rate.

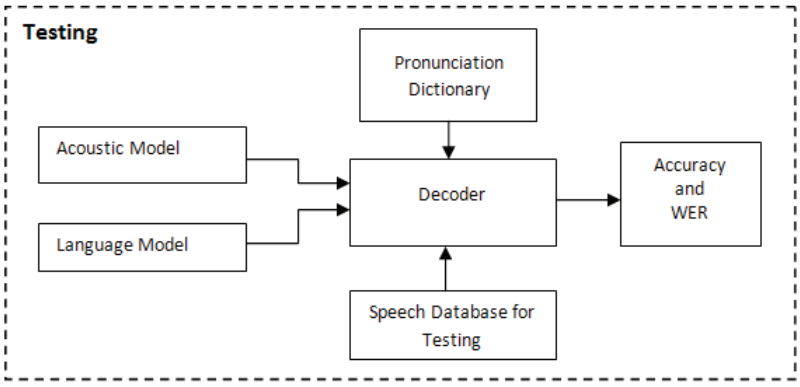

Figure 2: Testing Module

Various components of training and testing modules of Punjabi speech recognition are detailed below:

1. *Speech database creation:* Two speech databases are collected from 34 Punjabi speakers who are native of Punjab (16 males and 18 females) for training and testing of the system. The corpus consisted of 41 Punjabi letters, 9 vowels and 3 auxiliary signs with a total of 53 recordings per speaker. Files are recorded in .wav format with sample rate of 16 kHz, 16 bit mono. Speech Recordings are collected in environment with normal noise conditions. The distribution of speech databases for training and testing is shown in Table 1 having the Punjabi corpus description.

Table 1
Punjabi Corpus Description

Pbi_alphabet_corpus (Training)	Number of tokens	53 (41 alphabets, 9 dependent vowels, and 3 auxiliary signs)
	Number of speakers	25 (13 females, 12 males)
	Total number of recordings	1325
Pbi_alphabet_corpus (Testing)	Number of tokens	53 (41 alphabets, 9 dependent vowels, and 3 auxiliary signs)
	Number of speakers	9 (5 females, 4 males)
	Total number of recordings	477

2. *Pronunciation dictionary:* Pronunciation dictionary contains all Punjabi alphabets, vowels and auxiliary signs followed by phone sequence of each alphabet. Pronunciation of each alphabet is written along with the alphabet. Table 2, 3 and 4 show their pronunciation in the Gurmukhi script.

3. *Acoustic model:* Acoustic model represents the relationship between the recorded speech and the phonemes that make up the speech. In this paper, CMU Sphinxtrain[35] is used to train the model. Sphinxtrain is HMM based technique, in which words are represented as sequence of phonemes and each phoneme is modelled as a sequence of HMM states. Acoustic model is developed using speech recordings of Punjabi from the pronunciation dictionary and speech database. Speech features are extracted from the recordings by removing the unwanted and redundant speech signals, and stored in .mfc file. From each speech recording, sequence of feature vectors are extracted and computed using the Sphinxtrain tool. The engine looks into the Punjabi phonetic dictionary which maps each alphabet, vowel and auxiliary sign with the phoneme sequence. The acoustic features considered for this system consist of first and second derivatives of 13 dimensional Mel Frequency Cepstral Coefficients (MFCC). The window size of 25 ms and frame shift of 10 ms is considered.

Table 2
Pronunciation of Gurmukhi Letters

Gurmukhi Alphabet	*Punjabi Pronunciation*	*English Pronunciation*	*Gurmukhi Alphabet*	*Punjabi Pronunciation*	*English Pronunciation*	*Gurmukhi Alphabet*	*Punjabi Pronunciation*	*English Pronunciation*	*Gurmukhi Alphabet*	*Punjabi Pronunciation*	*English Pronunciation*
ੳ	ਊੜਾ	Oo'rhaa	ਚ	ਚੱਚਾ	chach'cha	ਤ	ਤੱਤਾ	tat'taa	ਯ	ਯੱਯਾ	yay'yaa
ਅ	ਐੜਾ	Ai'rhaa	ਛ	ਛੱਛਾ	chhachh'chhaa	ਥ	ਥੱਥਾ	thath'thaa	ਰ	ਰਾ ਰਾ	ra'raa
ੲ	ਈੜੀ	Ee'rhee	ਜ	ਜੱਜਾ	jaj'jaa	ਦ	ਦੱਦਾ	dad'daa	ਲ	ਲੱਲਾ	lal'laa
ਸ	ਸੱਸਾ	sas'saa	ਝ	ਝੱਝਾ	jhaj'jhaa	ਧ	ਧੱਧਾ	dhad'daa	ਵ	ਵਾ ਵਾ	vav'vaa
ਹ	ਹਾ ਹਾ	haa'haa	ਞ	ਞੱਞਾ	Njan'njaa	ਨ	ਨੱਨਾ	nan'naa	ੜ	ੜਾ ੜਾ	rhar'rhaa
ਕ	ਕੱਕਾ	Kak'kaa	ਟ	ਟੈ ਕਾ	tain'kaa	ਪ	ਪੱਪਾ	pap'paa	ਸ਼	ਸ਼ੱਸ਼ਾ	shash'shaa
ਖ	ਖੱਖਾ	khakh'khaa	ਠ	ਠੱਠਾ	thath'thaa	ਫ	ਫੱਫਾ	phaph'phaa	ਖ਼	ਖ਼ੱਖ਼ਾ	kha'khaa
ਗ	ਗੱਗਾ	gag'gaa	ਡ	ਡੱਡਾ	ddad'daa	ਬ	ਬੱਬਾ	bab'baa	ਗ਼	ਗ਼ੱਗ਼ਾ	gag'gaa
ਘ	ਘੱਘਾ	ghag'ghaa	ਢ	ਢੱਢਾ	dhad'daa	ਭ	ਭੱਭਾ	bhab'baa	ਜ਼	ਜ਼ੱਜ਼ਾ	Zaz'zaa
ਙ	ਙੰਙਾ	Ngan'ngaa	ਣ	ਣਾ ਣਾ	nhaa'nhaa	ਮ	ਮੱਮਾ	mam'maa	ਫ਼	ਫ਼ੱਫ਼ਾ	faf'faa
									ਲ਼	ਲ਼ੱਲ਼ਾ	lal'laa

Table 3
Pronunciation of Auxiliary Signs

Auxiliary Sign	*Pronunciation in Punjabi*	*Pronunciation in English*
ੱ	ਅਧਕ	Adhak
ਂ	ਬਿੰਦੀ	Bindī
ੰ	ਟਿੱਪੀ	ṭippī

Table 4
Pronunciation of Dependent Vowels

Vowel	*Pronunciation*	*Pronunciation in English*
ਾ	ਕੰਨਾ	Kannā
ਿ	ਸਿਹਾਰੀ	Sihārī
ੀ	ਬਿਹਾਰੀ	Bihārī
ੁ	ਔਂਕੜ	auṅkaṛ
ੂ	ਦੁਲੈਂਕੜ	dulaiṅkaṛ
ੱ ਂ	ਲਾਂ ਵਾਂ	lāṃvāṃ
ੱ ਂ	ਦੁਲਾਂ ਵਾਂ	dulāṃvāṃ
ੋ	ਹੋੜਾ	hōṛā
ੌ	ਕਨੌੜਾ	kanauṛā

4. *Language model:* Language model is a probability distribution over sequence of words. It is used for searching the correct word sequence by estimating the likelihood of the word based on previous words. Punjabi Language model is developed for 41 alphabets, 9 dependent vowels and 3 auxiliary signs. In this paper, CMU - Cambridge statistical language modeling toolkit[12] has been used to develop language model for Punjabi.

5. *Decoder:* In testing, the decoder matches the text recognized from the recordings with the actual text and calculates the accuracy. Pocket Sphinx [3], [12] decoder has been used for testing. It is a lightweight decoder, specially designed for hand held devices like mobile phones, tablets etc.

5. RESULTS

In the previous section acoustic model has been trained using 13 female and 12 male speech recordings. For testing purpose 5 female and 4 male speech recordings are considered as described in Table 1. The Pocketsphinx decoder matches the text recognized from the recordings with the actual text and calculates the accuracy. The accuracy of the system is represented by word error rate (WER). Low value of WER means high accuracy of system to recognize the speech. Table 5, Table 6 and Table 7 show the accuracy of each alphabet, dependent vowel and auxiliary sign respectively. Word error rate is calculated from the testing corpus. WER is calculated as:

$$\mathbf{WER} = \frac{\mathbf{I} + \mathbf{S} + \mathbf{D}}{\mathbf{N}} \tag{1}$$

where, I is number of insertions, S is number of substitutions, D is number of deletions and N is total number of words. Total WER observed is 17.4% and accuracy is 82.6%. Average accuracy of Punjabi alphabets, dependent vowels and auxiliary signs is 79.4%, 93.8%, 96.3% respectively.

Table 5
Accuracy of alphabets

Alphabet	*Accuracy (%age)*	*Alphabet*	*Accuracy (%age)*	*Alphabet*	*Accuracy (%age)*	*Alphabet*	*Accuracy (%age)*
ਓ	100	ਚ	66.7	ਤ	44.4	ਯ	100
ਅ	44.4	ਛ	100	ਥ	66.7	ਰ	88.9
ਏ	100	ਜ	100	ਦ	100	ਲ	100
ਸ	88.9	ਝ	77.8	ਧ	55.6	ਵ	77.8
ਹ	88.9	ਞ	88.9	ਨ	100	ੜ	100
ਕ	55.6	ਟ	100	ਪ	77.8	ਸ਼	100
ਖ	88.9	ਠ	88.9	ਫ	44.4	ਖ਼	66.7
ਗ	88.9	ਡ	77.8	ਬ	88.9	ਗ਼	33.3
ਘ	77.8	ਢ	100	ਭ	66.7	ਜ਼	77.8
ਙ	55.6	ਣ	55.6	ਮ	88.9	ਫ਼	100
						ਲ਼	33.3

Table 6
Accuracy of Dependent vowels

Dependent Vowel	*Accuracy (%age)*	*Dependent Vowel*	*Accuracy (%age)*
ਾ	88.9	ੱ	100
ਿ	100	ੂ	88.9
ੀ	100	ੇ	88.9
ੁ	88.9	ੈ	100
ੋ	88.9		

From the Table 5, we can observe that alphabets ਅ, ਤ, ਫ, ਗ਼, ਲ਼ have accuracy below 50%. Originally Punjabi had only 35 alphabets from ੳ to ੜ. Alphabets ਸ਼, ਖ਼, ਗ਼, ਜ਼, ਫ਼, ਲ਼ were added to accommodate sounds from Sanskrit and Persian[38] literature. Their sounds are similar to their original alphabet of Gurmukhi as ਸ and ਸ਼, ਖ and ਖ਼, ਗ਼ and ਗ਼, ਜ and ਜ਼, ਫ and ਫ਼, ਲ and ਲ਼. Phonologically there is little difference in their pronunciation. Because of this ਫ, ਗ਼, ਲ਼ are showing accuracy below 50%. ਅ is having sound similarity with ੜ, and most of the time it is misrecognized as ੜ. ਤ sounds similar to ਥ and ਦ and is being misrecognized often. Table VIII shows the alphabets that were misrecognized more than 4 times while testing. The alphabets ਕ and ਫ show multiple recognitions as compared to others which show not more than two recognitions. So, there is a need to improve model for these alphabets. Decoder needs to be fine tuned for these alphabets specifically.

Table 7
Accuracy of Auxiliary signs

Auxiliary Sign	*Accuracy (%age)*
◌ਂ	100
◌ੰ	100
◌ੱ	88.9

Table 8
Misrecognized Alphabets

Alphabet	*Misrecognized alphabet*
ਅ	ੜ, ਢ
ਕ	ਘ, ਦ ,ਖ਼, ਘ
ਝ	ਢ
ਚ	ਝ
ਣ	ਨ, ੜ
ਤ	ਦ, ਧ
ਥ	ਛ, ਠ
ਫ	ਖ, ਕ, ਫ਼, ਹ, ਥ
ਭ	ਬ, ਦ, ਪ
ਗ਼	ਗ
ਲ਼	ਲ, ਨ

A. Recognition Accuracy based on Gender

The recognition accuracy of 84% was observed for female speakers and 81.6% accuracy was observed for male speakers. Table 9 and Table 10 show accuracy for female and male speakers respectively. Accuracy in case of male speakers is less than the female speakers.

Speaker's pronunciation abilities, health, communication speed, environmental noise, microphone quality etc. may cause degradation of recognition accuracy. To increase the accuracy the recordings should be as superior as possible.

The accuracy of proposed model for Punjabi alphabets (79.4%), vowels (93.8%) and auxiliary signs (96.3%) is better than speech recognition for alphabets of BODO Alpha digits using HMM [33] and model for English alphabets [8] having accuracy of 70.17% and 62.28% respectively.

Table 9
Recognition Accuracy for Female Speakers

Speaker	*Total No. of tokens correctly recognized*	*Total No. of tokens spoken*	*Accuracy Percentage*
Female Speaker 1	48	53	91
Female Speaker 2	53	53	100
Female Speaker 3	40	53	75.5
Female Speaker 4	32	53	60
Female Speaker 5	49	53	92

Table 10
Recognition Accuracy for Male Speakers

Speaker	*Total No. of tokens correctly recognized*	*Total No. of tokens spoken*	*Accuracy Percentage*
Male Speaker 1	43	53	81.1
Male Speaker 2	43	53	81.1
Male Speaker 3	44	53	83
Male Speaker 4	43	53	81.1

6. CONCLUSION

In this paper automatic speech recognition system for Punjabi language is proposed. CMU tools like Sphinx Train, statistical language modeling toolkit are used to build the acoustic model and language models, while Pocket Sphinx is used to build decoder for system testing. The average accuracy of the system comes out to be 82.6%. Accuracy of alphabets, vowels and auxiliary signs come out be 79.4%, 93.8%, 96.3% respectively. Accuracy for female and male speakers is also analyzed. It is observed that accuracy for female speakers (84%) is more than that of male speakers (81%). Few alphabets ਅ, ਤ, ਫ, ਗ਼, ਲ਼ have accuracy below 50%. This is due to the fact that few alphabets are phonologically similar in structure. Overall accuracy of the acoustic model is good for Punjabi alphabets, vowels and auxiliary signs. As the speech recordings were collected in informal conditions, Speaker's pronunciation abilities, health, communication speed, environmental noise, microphone quality etc. may cause degradation of recognition accuracy. To increase the accuracy recordings may be done under controlled conditions.

REFERENCES

[1] T. B. Adam, and M. D. Salam, "Spoken English Alphabet Recognition with Mel frequency Ceptral Coefficients and Back propagation Neural Networks," International Journal of Computer Applications, 2012, Vol. 42, No. 12, pp. 21-27.

[2] F. Biadsy, P. J. Moreno, and M. Jansche, "Google's Cross-Dialect Arabic Voice Search," in Proc. ICASSP, 2012, pp. 4441-4444.

[3] (2015) Buliding application with pocketsphinx. [Online] Available: http://cmusphinx.sourceforge.net/wiki/tutorialpocketsphinx

[4] (2016) CMU Sphinx. [Online]. Available: http://cmusphinx.sourceforge.net/

[5] (2016) CMUCLMTK Development. [Online]. Available: http://cmusphinx.sourceforge.net/wiki/cmuclmtkdevelopment

[6] R. Cole, M. Fanty, Y. Muthusamy, and M. Gopalakrishnan, "Speaker Independent recognition of Spoken English Letters," in Proc. IJCNN, 1990, Vol. 2, pp. 45-51.

[7] P. Cosi, J. Hosom, and A. Valente, "High Performance Telephone Bandwidth Speaker Independent Continuous Digit Recognition," in Proc. ASRU, 2001, pp. 9-13.

[8] B. Das, S. Mandal and P. Mitra, "Bengali speech corpus for continuous automatic speech recognition system" in Proc. COCOSDA 2011, 2011, pp. 51-55.

[9] M. Fanty, R. A. Cole, and K. Roginski, "English alphabet recognition with telephone speech," Advances in Neural Information Processing Systems CA Mateo: Morgan Kaufmann, 1992, Vol. 4, pp. 199-206.

[10] (2015) Google voice search. [Online]. Available: http://www.google.co.in/insidesearch/features/voicesearch/

[11] (2015) Gurmukhi Alphabet – Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/Gurmukhi_alphabet

[12] D. Huggins-Daines, M. Kumar, A. Chan, A. W. Black, M. Ravishankar, and A. I. Rudnicky, "Pocketsphinx: A Free, Real-Time Continuous Speech Recognition System for Hand-Held Devices," in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, 2006, Vol. 1, pp. 185-188.

[13] M. Karnjanadecha, and Z. Zahorian, "Signal Modeling for High-Performance Robust Isolated Word Recognition," IEEE Transactions on Speech and Audio Processing, 2001, Vol. 9, No. 6, pp. 647–654.

[14] S. S. Khaira, Punjabi Bhasha Viyakarn Ate Bantar (Punjabi), Punjabi University, Patiala, 2011.

[15] X. Liu, J. L. Hieronymus, M. J. F. Gales, and P. C. Woodland, "Syllable language models for Mandarin speech recognition: Exploiting character language models," Journal of Acoustics Society of America, 2013, Vol. 133, No. 1, pp. 519-528.

[16] P. C. Loizou, and A. S. Spanias, "High-Performance Alphabet Recognition," IEEE Transactions on Speech and Audio Processing, 1996, Vol. 4, No. 6, pp. 430–445.

[17] (2015) Online Punjabi Teaching. [Online]. Available: http://www.learnpunjabi.org/intro1.asp#consonants

[18] (2015) Punjabi Language Encyclopedia Britannica Online. [Online]. Available: https://www.britannica.com/topic/Punjabi-language

[19] (2016) Punjabi language. [Online]. Available: http://en.wikipedia.org/wiki/Punjabi_language

[20] (2015) Punjabi/Gurmukhi/Tone – Wikibooks, open books for an open world. [Online]. Available: https://en.wikibooks.org/wiki/Punjabi/Gurmukhi/Tone

[21] H. Satori, and F. ElHaoussi, "Investigation Amazigh speech recognition using CMU tools," International Journal of Speech Technology, 2014, Vol. 17, pp. 235-243.

[22] H. Satori, M. Harti, and N. Chenfour, "Introduction to Arabic Speech Recognition Using CMUSphinx," in Proc. International Symposium on System Information and Communication Technologies, 2007, pp. 1-4.

[23] M. Schuster, and K. Nakajima, "Japanese and Korean Voice Search," in Proc. ICASSP, 2012, pp. 5149-5152.

[24] (2015) Shahmukhi Alphabet – Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/Shahmukhi_alphabet

[25] K. Sharma, T. V. Prasad, and S. V. A. V. Prasad, "Hindi Speech Enabled Windows Application Using Microsoft SAPI," International Journal of Computer Engineering & Technology, 2013, Vol. 4, No. 2, pp. 425-436.

[26] L. Thakuria, A. Das, P. Acharjee, and P. H. Talukdar, "Automatic Speech Recognition of BODOAlpha digits Using Hidden Morkov Models," International Journal of Advanced Research in Computer Science and Software Engineering, 2013, Vol. 3, No. 11, pp. 1496-1500.

[27] R. Thangarajan, A. M. Natarajan, and M. Selvam, "Syllable modeling in continuous speech recognition for Tamil language," International Journal of Speech Technology, 2009, Vol. 12, pp. 47-57.

[28] (2016) Training Acoustic Model for CMUSphinx. [Online]. Available: http://cmusphinx.sourceforge.net/wiki/tutorialam

[29] R. Walha, F. Drira, H. El-Abed, and A. M. Alimi, "On Developing an Automatic Speech Recognition System for Standard Arabic Language," International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering, 2012, Vol. 6, No. 10, pp. 1138-1143.

[30] L. Rabiner, and B. Juang, "An introduction to hidden Markov models," IEEE ASSP Magazine, 1986, Vol. 3, No. 1, pp. 4-16.

[31] (2016) Painti Akhri - SikhiWiki, free Sikh encyclopedia [Online]. Avilable: http://www.sikhiwiki.org/index.php/Painti_Akhri

[32] Y. Kumar and N. Singh, "Automatic Spontaneous Speech Recognition for Punjabi Language Interview Speech Corpus," International Journal of Education and Management Engineering (IJEME), 2016, Vol. 6, No. 6, pp. 64-73.