International Journal of Information Technology and Knowledge Management
January-June 2011, Volume 4, No. 1, pp. 139-141

GAYATRI – A FAST HINDI TEXT TO SPEECH SYSTEM WITH INPUT SUPPORT FOR ENGLISH LANGUAGE

Danish Chopra*

These days, Indian users prefer to type Hindi sentences in English language. This fact is prevalent all over the web especially on the social networking websites such as Facebook, Orkut etc. and the internet messengers. Users do this because they are used to the English QWERTY keyboard and due to non-familiarity and availability of the Hindi keyboard. In this paper, I have presented a simple yet fast and robust approach to develop a text to speech system for Hindi language called Gayatri. Gayatri is a Hindi text to speech system for such a scenario in which the input is in form of English typed Hindi sentences or words. A Hindi word being typed in English can have many different possibilities according to different users. We consider all of them for each individual word. Gayatri also supports a custom developed Hindi keyboard layout for English QWERTY keyboards to type sentences in Hindi language using the English QWERTY keyboard. Gayatri has been developed for windows platform but can also run on a web based interfaces and windows based mobile phones.
Keywords: Artificial Voice Synthesizer, QWERTY Keyboard, World Wide Web, Comma Separated Values, Phoneme, Diphones

1. INTRODUCTION

Text to speech systems are increasingly becoming an essential component of different type of computing systems. Also known as an artificial voice synthesizer, a text to speech system can produce human voice artificially based on a given string. Developing a text to speech system for a language that can support inputs in other languages can be helpful not only to the users that know that language but are not familiar with its relative keyboard layout but also to international users that do not know that language at all and can hence type in that language using their local language keyboard layout. Gayatri is a Hindi text to speech system for such a scenario in which the input string comes in form of an English sentence or word (all possibilities considered) apart from the regular Hindi sentence or word. This scenario in which a Hindi sentence or word is being type in English is a common practice on the World Wide Web. Users prefer this since they are more familiar and comfortable with an English keyboard layout rather than a Hindi keyboard layout. So the users can type Hindi sentences or words in English language and feed it into the Gayatri TTS engine to be converted to speech. Gayatri also supports a custom developed Hindi keyboard layout for English QWERTY keyboards to type sentences in Hindi language using the QWERTY keyboard.

2. PRESENT WORK

The various steps used to develop the Gayatri Hindi text to speech system are discussed below.

* Department of Computer Science Engineering, DAVIET, Jalandhar, Email: danishchopra@gmail.com

2.1. Approach

Gayatri is based on pre-recorded voice and concatenation of Hindi language diphone transcriptions.

2.1.1. Prerecorded Voice

We record large paragraphs of Hindi words (from any book involving good and commonly used Hindi vocabulary) in a continuous rhythm with little gap between two consecutive words in form of a silence and save them as .ogg audio files on the disk. Saving audio files in .ogg format save us a lot of disk space yet retaining the audio quality. We maintain a large central database containing different tables that act as a lookup repository for the different Hindi words supported by Gayatri. We save the information of a given Hindi word in the database according to its word length and its arrival time in a particular .ogg file. The table format is described in figure 1.0.

Word	Start Time	End Time	Filename

Fig. 1: Table Format for the Database

The different tables are named from 1 to 20. These table numbers signify nothing else but the length of the Hindi words. So if we are looking for a Hindi word of word length 5 in the database, then the search must be performed only in the table 5, providing an efficient and fast way to search the database. If a match is found in the database for that word, we read its entry for its arrival time, end time and filename (.ogg). We now play the .ogg file between the arrival time and end time to speak out the Hindi word (Figure 2.0).

The different tables are named from 1 to 20. These table numbers signify nothing else but the length of the Hindi words. So if we are looking for a Hindi word of word length 5 in the database, then the search must be performed only in the table 5, providing an efficient and fast way to search the database. If a match is found in the database for that word, we read its entry for its arrival time, end time and filename (.ogg). We now play the .ogg file between the arrival time and end time to speak out the Hindi word (Figure 2.0).

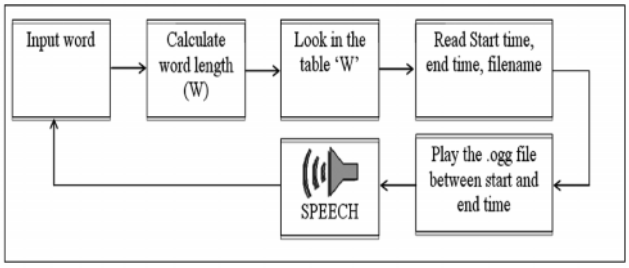

Fig. 2: Gayatri Approach for Pre-recorded Voice

2.1.2. Support for English Language

Each Hindi word is stored in the database in the following way:

1. Actual Hindi Word: This is the actual Hindi word when typed through our custom developed Hindi keyboard.
2. English Word Possibilities: Followed by the actual Hindi word and a comma, English word possibilities are the various possibilities in form of comma separated values of actual Hindi word when typed in through the English keyboard.

Example: ok fiQj vkx;k

ok will be saved as VO, VOH, WO, WOH fiQj – PHIR, FIR vkx;k– AAGYAA, AAGEYA, AAGAYA.

Here, VO, PHIR and AAGYAA are the actual Hindi words typed through the custom developed Hindi keyboard (Figure 3.0) while other words are the English word possibilities.

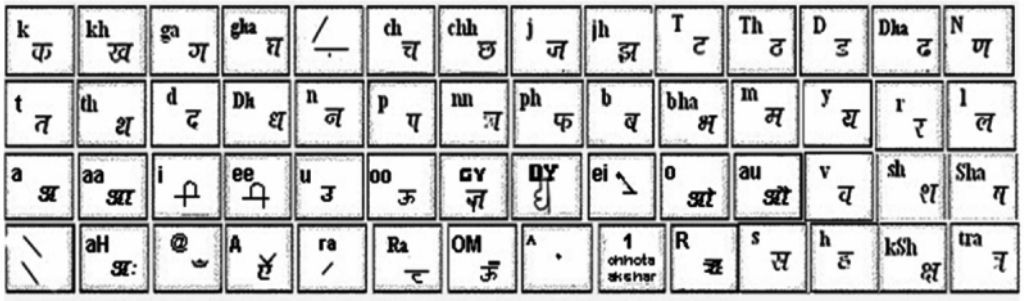

Fig.3: Custom Developed Hindi Keyboard for Gayatri (Case Sensitive)

The Hindi font used is Rovisuam. This keyboard layout is available as a popup application in windows and can be used to type Hindi language using an English QWERTY keyboard layout. The keyboard layout may also be printed and pasted on the actual keyboard.

2.1.3. Concatenation of Words and Diphones

Since all the Hindi words cannot be recorded; we follow a different approach for words that are not available in the database. We follow the approach of concatenation of Hindi words and diphones to form a new word in real-time and play it. Phonemes are the smallest distinct units of sound that are clubbed to form words; words are clubbed to form sentences and so on. A diphone is a pair of phonemes. Concatenation of different words and diphones can result into formation of any word for a particular language.

For example: Hkkjr can be formed from concatenation of Hkk and jr.

We record all the possible Hindi diphones as well as all the possible combinations of 2 and 3 letter Hindi words and store them in the database. We now follow the following approach after taking the input from the user:

1. If the input word has 1 letter, 2 letters or 3 letters, it need not be broken because we have recorded all possibilities containing such words already.

   Examples:

   d d de dy fd dp dn dgk dje

2. If the input word has 4 characters we divide it into a ratio of 2:2

   Example: Hkkjr → Hkk + jr

3. If the input word has 5 characters we divide it into a ratio of 2:3 or 3:2

   ilrD → il + rD or il + lrD

4. If the input word has 6 characters we divide it into a ratio of 3:3 or 2:2:2.
5. If the input word has 7 characters we divide it into a ratio of 3:2:2 or 2:2:3.
6. If the input word has 8 characters we divide it into a ratio of 3:3:2 or 2:3:3 or 3:2:3 And so on.

During concatenation, there will be a point of silence between any two consecutive diphones/words resulting in inaccuracy and non-continuity in playing out the complete word in a continuous flow. This silence is due to inability of the sound player to play one sound just next to another. To eliminate such errors we follow an approach as we call it "Selective Focus". Selective focus makes sure that when the first diphone/word is been played out by the sound player, it must be followed by playing the sound of the last letter of that diphone continuously until the next consecutive diphone/word is concatenated with it and played.

Example: We focus on k "AA" part of Hk "BH" and concatenate jr "RT" with it to produce a continuous and nonstop speech.

3. USE CASE SCENARIOS

Let us consider the following case when the user wants the text to speech system to speak out the following Hindi sentence:

ejk uke nkfu'k g (Equivalent of "Eksjk UkkEk nkfUk'k gS" in Rovisuam font).

The user can type this Hindi sentence in English in one of the following possibilities:

1. mera naam danish hai (best case).
2. meraa naam daanish ha.
3. mera nam danish haye (worst case).

4. CONCLUSION

Currently, Gayatri is able to speak out majority of the Hindi words when typed in English. The challenge is to make the system able to speak out the words in more natural way. Another challenge is to support the various possibilities of Hindi word when typed in English. Future work includes support for standard Hindi keyboard and indic keyboard layouts. The approach can be used to develop text to speech systems that support inputs in other languages. Since, English QWERTY keyboard is the most widely used and accepted keyboard in the world, so typing in English even the local languages becomes a fast, easy and effective method of communication. Gayatri is an effort in such a scenario in which the user type in the Hindi words/sentences in English and it is converted to Hindi speech.

REFERENCES

[1] Mukhopadhyay, A.; Chakraborty, S.; Choudhury, M.; Lahiri, A.; Dey, S.; Basu, A., "Shruti- an Embedded Text-to-speech System for Indian Languages" Software, IEE Proceedings, 153, Issue 2, April 2006, Page(s) 75–79.

[2] Ramakrishnan, A.G. et.al., "Tools for the Development of a Hindi Speech Synthesis System", in 5th ISCA Speech Synthesis Workshop, Pittsburgh, pp. 109-114, 2004.

[3] Hoory, R.; Shaked, N.; Chazan, D., "Building a Speech Database for the Purpose of Speaker Specific Speech Synthesis" Signal Processing, Oct. 1996, 1, 14 Page(s):741 - 744.

[4] S.P. Kishore, Rohit Kumar, Rajeev Sangal, "A Data-driven Synthesis Approach for Indian Languages using Syllable as Basic Unit", in Intl. Conf. on Natural Language Processing (ICON), 2002, pp. 311-316.

[5] Google Online Hindi to English Transliteration Tool, Referenced Online at www.google.com/transliteration.