ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

Android Application Development for Syllable Based Tamil Text To Speech Synthesizer

M.Aishwarya[1], Shankar[2]

M.Tech, Department of C.S.E, Bharath University, Chennai, Tamil Nadu, India[1]

Research Scholar, Department of C.S.E., Anna University, Chennai, Tamil Nadu, India[2]

ABSTRACT: In the field of Natural Language Processing aims at translating an input Tamil sentence into an equivalent is spoken Tamil translation of the sentence. This brings together two major domains in NLP,i.e, machine translation and TTS(Text to Speech). Natural Language Processing is the field of computer science that deals with interactions between computer and human languages. In this paper, a hybrid approach of translation called Universal Networking Language(UNL) is employed following which a Text-to-Speech system is used that gives a neutral emotionless voice as the output. Emotion-based prosody is added to this neutral voice to give a rich, humanized voice output. Universal Networking Language (UNL) is a declarative formal language that is used to represent semantic data extracted from natural language texts.

KEYWORDS: Natural Language Processing, TTS (Text To Speech), Universal Networking Language

I.INTRODUCTION

Natural language processing is a field that lies at the intersection between artificial intelligence and linguistics. It is primarily concerned with the interactions that occur between computers and human languages. NLP is within the domain of human-computer interaction. The predominant challenge in NLP involves the creation of a medium through which computers assume the ability to comprehend natural language and hence derive meaning from natural language input. Computational linguistics is a subdomain under NLP, which focuses specifically on language related work such as language modeling and representation. Machine Translation is one such field of computational linguistics which involves the development of a system for translation of given text from a source language to the target language.Translation has been in existence ever since written literature started gaining traction. This was primarily due to the fact that knowledge in one language could easily be propagated to other languages, and hence to other cultures with the help of translation. In modern times, the preponderance of literature has made it mandatory to automate the process of translation which would otherwise be tedious if done manually. Thus various machine translation approaches were developed over the past decades. we have developed a system that is capable of translating a Tamil sentence to Tamil speech with emotionless using Universal Networking Language (UNL) as an intermediate. The translated sentence is given to a Text-to-Speech system and then finally prosody is added to give an output voice that is humanized. Since the final output is voice, it can be of great help to visually challenged people and also to illiterate.

II.LITERATURE SURVEY

Synthesized speech can be created by concatenating part of recorded speech which is stored in a database. The mainly significant qualities of a speech synthesis system are naturalness and Intelligibility [1]. Naturalness expresses how intimately the output sounds like human speech, whereas intelligibility is the easiness with which the output is understood. The main function of text-to-speech (TTS) system is to convert an arbitrary text to a spoken Waveform. TTS is one of the major applications of NLP. TTS Synthesizer is a computer-based system that should be understood any text clearly whether it was established in the computer by an operator or scanned and submitted to an Optical Character Recognition (OCR) system. Single text-to-speech (TTS) system for Indian languages viz Tamil. text-to-speech (TTS) system based on the Concatenative synthesis approach. This conversion involves text processing and speech generation processes. These processes have the connections to use the linguistic theory, models of speech

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

production, and the acoustic-phonetic characterization of the language. Font Characters Mapping is required for font characters of an Indian language to represent vowel and consonant alphabet. Text analysis is the task of identifying words in the text.Text normalization includes "Token Identification" which is the task of identifying special symbols, numbers and "Token to Words" which convert the identified tokens to words for which there is a well-defined method of pronunciation. In a text-to-speech system, spoken utterances are automatically produced from the text.Tamil text-to-speech (TTS) system based on the Concatenative synthesis approach. The TTS can be a voice for those people who cannot speak. The TTS system can be useful for the SMSs, Web pages, News, any articles, and the Microsoft office tools and so on. Text-to-speech system based on Concatenative synthesis needs well-arranged speech corpus. The quality of synthesized speech waveform depends on up on the number of realization of various units present in the speech corpus. A good quality microphone should be used to avoid noise in speech wave file. In text-to-speech, the accuracy of the system is calculated in the ways of naturalness and intelligibility.

III. PROBLEM DESCRIPTION

With a Tamil sentence as an input, the MT system should translate it to a correct and equivalent English sentence with no loss in meaning.The sentence should also be voiced with the emotion conveyed in it. When a paragraph is given as input, it should translate the paragraph sentence by sentence.

IV. EXISTING SYSTEM

Windows and Linux based Application only available Requires domain analysis on removing the ambiguity of relations in language processing. No consideration about complex and compound Tamil Sentences.

A. TEXT-TO-SPEECH SYSTEM

Text to Speech scheme is a major kind of speech synthesis application and that is hardened to generate a vocal sound version of the text. The TTS system produces the spoken voice from pre-recorded database. The following measures are used to determine the quality of the system.

B. INTELLIGIBILITY

One is the use of speech intelligibility testing in the assessment of hearing disorders and the other has used these tests to assess speech synthesis and speech transmission systems.

C. NATURALNESS

The synthesized speech should have the quality of speech like the human. The naturalness of speech means, approaching the human levels of speech with reliability. The following Fig. 1 illustrates the elemental manner of Text-to-Speech system for Natural Languages. The conversion of given text into corresponding speech output. This mapping of text into speech is performed in two phases usually called as high – level and low-level synthesis. The high-level phase 2014 International Conference on Recent Trends in Information Technology is also called as Text analysis, where the input text is transcribed into a phonetic representation with the linguistic rules in that language [1]. Low-level phase is called as phonetic analysis. The phonetic analysis is trying to map the Phoneme with the pre-recorded speech using unit selection speech synthesis algorithm to increases its naturalness.

V. ARCHITECTURE

A.POS Tagger

The input sentence is given to this rule based POS tagger which returns the part of speech of every word in the sentence. In Tamil there are ‘vallinammigumidam’ where a vallinamei (ì, î, ò, ô) will be added to the inflected word, they are scanned from right to left. Once those vallinamei(if any) is removed, the inflections are used to determine the POS tags. A customised tagset tailored to fit the UNL relations and Enconversion module is created and used by this tagger. These tags are used in Morphological analysis and root word extraction.

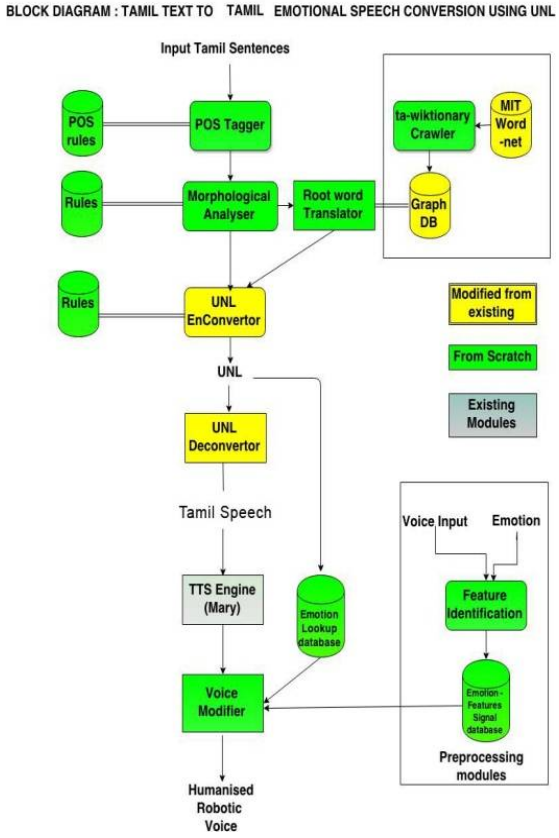
ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017



B.Morphological Analyser

Tamil is a morphologically rich language. A number of words can be formed from a root word by inflections. But only the root word will be mapped for translation and features like tense, gender and plurality can be extracted from the inflections which will be used in Deconversion to target language. A rule based morphological analyser has been developed for this purpose which takes inflected word with POS tag as input and it returns the root word which is the 'Universal Word' (UW) in the UNL representation.

C.Emotion-FeaturesDatabase Construction

This database consists of all features that have been extracted from the voice samples. The voice samples are obtained from 3 female speakers. Since the output of the "Mary" Text to Speech Synthesis system which has a female voice that need to be modified, the voice samples are taken from female speakers.The voices are recorded in a fairly quiet room so as to avoid any external noise getting recorded. Then the sentences are read into Audacity, an audio analysis software in order to distinguish between the word boundaries clearly and to split them into audio files corresponding to each word. If excess noise has been recorded, then using the noise removal feature of Audacity, the audio sample can be made relatively noise free.

D.Voice Modification

The output of the text to speech system is modified to add the emotion and prosody. The emotion features database is consulted and the corresponding values of various parameters (like fundamental frequency, intensity, duration etc) are calculated. Then the robotic voice is modified based on the values of these parameters to produce the humanised voice.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

VI.PROPOSED SYSTEM

A.TEXT ANALYSIS

NLP component is responsible for the text analysis phase.

TTS system converts the text into phonemes. The words, which are not present in the database, there is a need of syllabification.

B.TOKENIZATION

Tokenization is the process of breaking upstream of text into words, symbols called tokens. In each sentence, there are many words are separated out. The written text can be segmented using whitespace as delimiter (“!”,” ”,”?”,”.”). This process is called tokenization. The Tamil input text is accepted as the string in the text field using Netbeans. Arial Unicode MS is set as a font type to accept the string in Tamil. Based on the delimiter the input sentence is tokenized into words.

C.SYLLABIFICATION

A syllable is the unit of language, which consists of an interrupted portion of sound when a word is pronounced. The given input checked with the speech corpus, whether the speech exists. If the word does not exist, the word is segmented into syllables using syllabification rules. The segmented units

are in the form of c m1 v c m2, where c is consonant, v is vowel and m1, m2 may vary from 0 to n. The following rules are used for Text syllabification process, which gives the syllabified units for the TTS process.

E.g. 1

வான்ேமகம் | கருைமயாக | உள்ளது (3 units)

வான் | ேமகம் |கருைமயாக | உள்ளது (4 units)

Based on these rules the text is split into syllables and the location of the corresponding sound files

VII.BLEU Score

BLEU stands for Bilingual Evaluation Understudy.It is an n-gram co-occurrence based MT evaluation system that verified how close the MT translated text is to, to a human translated reference. N-gram precision in BLEU is computed as follows:

$$Pn = \frac{\sum_{Cef \in Candidates} \sum_{n\ gram \in C} Count_{clip}(n\text{-}gram)}{\sum_{Chef \in Candidates} \sum_{n\ gram \in C} Count_{clip}(n\text{-}gram)}$$

Where Count clips(n-gram) is the maximum number of n-grams co-occurring in a candidate translation. To prevent very short translations that try to maximize their precision scores, BLEU adds a brevity penalty, BP, to the formula:

$$BP = \begin{cases} 1 & \text{if } jcj > jrj \\ e^{(1-|r|-c|)} & \text{if } jcj \ jrj \end{cases}$$

Where jcj is the length of the candidate translation and |r| is the length of the reference translation. The BLEU formula is then written as follows

$$BLEU = BP \exp \sum_{n=1}^{N} wn \log pn$$

The weighting factor wn, is set at N1.

The BLUE score comparison of our System to that of the Google translator is shown in Figure 6.1.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

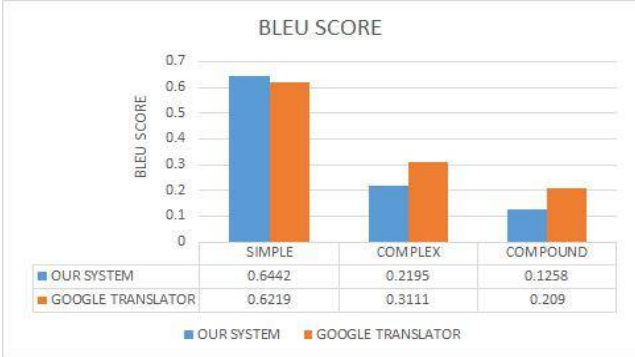


Figure 6.1 BLEU scores: Our system vs Google translator for the syllables are written in a text file.

IX.CONCLUSION

This system translates Tamil text to Tamil text with UNL as interlingual and then voices the Tamil text out with emotions.In the case of a morphologically rich language such as Tamil, the use of an interlingua is better for translation.A custom built morphologically analyzer is used to extract noun, tense and gender from the Tamil text. A POS Tagger was also created to classify the words into noun, verb, object, a converter adverb, and adjective. Based on the POS tags and morphological structures rules are devised to convert the Tamil text to Upconverter. Then the emotion in the sentence is identified from the UNL and based on the identified emotion, the TTS output for the translated Tamil Speech is modified to incorporate prosody on it.

REFERENCES

[1] G. G. Chowdhury, "Natural language processing," *Annual review of information science and technology*, vol. 37, no. 1, pp. 51–89, 2003.
[2] K. W. Church and R. L. Mercer, "Introduction to the special issue on computational linguistics using large corpora," *Computational linguistics*, vol. 19, no. 1, pp. 1–24, 1993.
[3] J. Lehrberger and L. Bourbeau, *Machine Translation: Linguistic characteristics of MT systems and general methodology of evalua-tion*, vol. 15. John Benjamins Publishing, 1988.
[4] L. Dugast, J. Senellart, and P. Koehn, "Statistical post-editing on systran's rule-based translation system," in *Proceedings of the Sec-ond Workshop on Statistical Machine Translation*, pp. 220–223, 2007.
[5] P. Koehn, F. J. Och, and D. Marcu, "Statistical phrase-based trans-lation," in *Proceedings of the 2003 Conference of the North Amer-ican Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1*, pp. 48–54, 2003.
[6] V. Kumar, R. Sridhar, S. Narayanan, and S. Bangalore, "Enriching spoken language translation with dialog acts," in *Proceedings of the 46th Annual Meeting of the Association for Computational Linguis-tics on Human Language Technologies: Short Papers*, pp. 225– 228, 2008.
[7] D. Chiang, "A hierarchical phrase-based model for statistical ma-chine translation," in *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, pp. 263–270, 2005.
[8] A. Ramanathan, J. Hegde, R. M. Shah, P. Bhattacharyya, and M. Sasikumar, "Simple syntactic and morphological processing can help english-hindi statistical machine translation.," in *IJCNLP*, pp. 513–520, 2008.
[9] K. Bar, Y. Choueka, and N. Dershowitz, *An Arabic to English example-based translation system*. PhD thesis, 2007.
[10] Y. Lepage and E. Denoual, "Purest ever example-based ma-chine translation: Detailed presentation and assessment," *Machine Translation*, vol. 19, no. 3-4, pp. 251–282, 2005.
[11] J. G. Carbonell, S. Klein, D. Miller, M. Steinbaum, T. Grassiany, and J. Frey, "Context-based machine translation," pp. 19–28, 2006.
[12] H. Uchida, M. Zhu, and T. Della Senta, *Universal Networking Lan-guage*. 2005.
[13] P. Kumar, *UNL Based Machine Translation System for Punjabi Language*. PhD thesis, Tharapur University, Patiala, 2012.
[14] D. Schwarz, "Data-driven concatenative sound synthesis," in *Pro-ceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-00), Verona, Italy,2009*.
[15] R. Carlson, T. Sigvardson, and A. Sjölander, "Data-driven formant synthesis," *KTH, Stockholm, Sweden, Progress Report*, vol. 44, pp. 121–124, 2002.
[16] J. Stark, B. Lindblom, and J. Sundberg, "Apex an articulatory syn-thesis model for experimental and computational studies of speech production," *TMH-QPSR*, vol. 2, pp. 45–48, 1996.
[17] K. Tokuda, T. Yoshimura, T. Masuko, T. Kobayashi, and T. Ki-tamura, "Speech parameter generation algorithms for hmm-based speech synthesis," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, pp. 1315–1318, 2000.
[18] R. J. McAulay and T. F. Quatieri Jr, "Computationally efficient sine wave synthesis for acoustic waveform processing," June 26 1990. US Patent 4,937,873.
[19] A. W. Black and K. A. Lenzo, "Flite: a small fast run-time synthesis engine," in *4th ISCA Tutorial and Research Workshop (ITRW) on Speech Synthesis*, 2001.
[20] S. Sangeetha and S. Jothilakshmi, "Syllable based text to speech synthesis system using auto associative neural network prosody prediction," *International Journal of Speech Technology*, vol. 17, no. 2, pp. 91–98, 2014.