IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 17, Issue 1, Ver. VI (Jan – Feb. 2015), PP 34-43
www.iosrjournals.org

Implementation of English-Text to Marathi-Speech (ETMS) Synthesizer

Sunil S. Nimbhore[1], Ghanshyam D. Ramteke[2], Rakesh J. Ramteke*[3]
[1](Assistant Professor, Department of Computer Science & IT, Dr. B. A. M. University, Aurangabad, India)
[2](Research Scholar, School of Computer Sciences, North Maharashtra University, Jalgaon,, India)
[3](Professor, School of Computer Sciences, North Maharashtra University, Jalgaon, India)

Abstract: *This paper describes the Implementation of Natural Sounding Speech Synthesizer for Marathi Language using English Script. The natural synthesizer is developed using unit selection on the basis of concatenative synthesis approach. The purpose of this synthesizer is to produce natural sound in Marathi language by computer. The natural Marathi words and sentences have been acquired by 'Marathi Wordnet' because all Marathi linguists are referred Wordnet. In this system, around 28,580 syllables, natural words and sentences were used. These natural syllables have been spoken by one female speaker. The voice signals were recorded through standard Sennheiser HD449 Wired Headphone using PRAAT tool with sampling frequency of 22 KHz. The ETMS-system was tested and generated the natural output as well as waveform. The formant frequencies (F1, F2 and F3) were also determined by MATLAB and PRAAT tools. The formant frequencies results are to be found satisfactory.*

Keywords: *Formant Frequency, Natural Synthesizer, Concatenative, LPC, Speech Corpus.*

I. Introduction

Nowadays, Human Computer Interface (HCI) is familiar, and handled by human for increasing the efficiency in various fields. Speech is the most effective way of communication for human beings. The interface between human and computer speech play an important role in day to day life. The new technologies are recently been adopted for effective and user friendly communication into digital technology. However, the English-Text to Marathi-Speech (ETMS) is not available for the commercial purpose. A synthesizer can incorporate a model of the vocal tract for human voice characteristics to create a completely "synthetic" voice output. The quality of a speech synthesizer is referred by its similarity to the human voice and it shall be understandable.

The formants are physically defined as poles in a system function expressing the characteristics of a vocal tract. Therefore, it can be demonstrated clearly their existence. Therefore, a different variety of the formant tracking spectrum can be analyzed and synthesized [1]. The formant frequencies play a vital role for signal classification of speech signals, and therefore, these techniques are reliable for computing and suitable for speech synthesis. The Marathi numerals, vowels, and words speech signals are stored in speech corpus. These speech signals are extracted the first three formant frequencies (F1, F2 and F3) [2]. The most popular two techniques for format frequencies are [A] LPC analysis [B] Cepstral analysis. The experimental work is done by LPC analysis using the MATLAB and PRAAT tool.

The main objective of the present paper is to report design and development of as system which input is in the form of English text and output is corresponding spoken text into Marathi language. The synthetic spoken words are also analyzed and the formant frequencies were determined by tool available in MATLAB and PRAAT. These formant frequencies determine the quality of the spoken synthetic words

The paper is organized as follows; Section-I deals with introduction of Natural Sounding Speech Synthesizer and formant pitch frequencies. Section-II depicts the description of Marathi digit, vowels, consonants and words, written in Devnagari scripts are described. Section-III describes concatenative synthesis approach is based on unit selection. Section-IV describes the methods regarding the determination of formant frequencies using software tools available in MATLAB and PRAAT. Section-V describes the acquisition of Text and Speech Corpus for Marathi Language. The section-VI deals with experimental works and discussion of results. Finally, the paper is concluded in the section-VII.

II. Description of Devnagari Script Language

Devnagari script is used worldwide by millions of people. Marathi, Hindi, Sanskrit Languages are written by using Devnagari script. The Structure and Grammar of Marathi language is similar to Hindi language. Marathi is primarily spoken language in Maharashtra State, (India). It is an official language used in State of Maharashtra. In Maharashtra all students study using the Devnagari script (Marathi) as 1st Language at school level.

Marathi is spoken completely in Maharashtra state which covers in vast geographical area which consists of 36 different districts in different dialects. The Major dialects of Marathi languages are called standard Marathi and Varhadi. The other few sub-dialects are like Ahirani, Dangi, Vadvali, Samavedi, Khandeshi and Malwani. However, Standard Marathi is an official language of Maharashtra. So it is essential to do the research on this domain.

The written digits (0-9), 12-vowels, 34-consonants and some words in English and Devnagari scripts together are shown table-1 [3]. The written text in English can be performed in Marathi in similar way as illustrate in Table-1. For example word 'bharat' in English will be performed as 'भारत' in Marathi.

Table 1. Digits, Vowels, Consonants and Words Written in Marathi and English Script

WRITTEN ENGLISH SCRIPT		WRITTEN DEVNAGARI SCRIPT	
DIGITS	0,1,2,3,4,5,6,7,8,9	**अंक (ANK)**	०, १, २, ३, ४, ५, ६, ७, ८, ९
VOWELS	a, aa, i, ee, u, oo, ae, aae, o, ou, am, aha	**स्वर (SWARAS)**	अ, आ, इ, ई, उ, ऊ, ए, ऐ, ओ, औ, अं, अः
CONSONANTS	ka, kha, ga, gha, nga, cha, chcha, ja, jha, ta, tha, da, dda, nha, tta, ththa, da, dha, na, pa, pha, ba, bha, ma, ya, ra, la, va, sha, sshha, sa, ha, lla, ksha, gnya	**व्यंजन (VANJNAS)**	क, ख, ग, घ, ङ च, छ, ज, झ, ट, ठ, ड, ढ, ण, त, थ, द, ध, न, प, फ, ब, भ, म, य, र, ल, व, श, ष, स, ह, ळ, क्ष, ज्ञ
WORDS	bharat, aajoba, chandichya, zadala. antaralat.	**शब्द (SHABDAS)**	भारत, आजोबा, चांदीच्या, डावपेच, झाडाला, अंतराळात. इ.

III. Concatenative Synthesis Approach

Concatenative Speech Synthesis approach plays an important role to implement the natural synthesized speech for Marathi language. Concatenative synthesis is concatenating the pre-recorded segments to generate the natural speech. Concatenative speech is produce intelligible & natural synthetic speech, usually close to a real voice of person [4, 5]. However, concatenative synthesizers are limited to only one speaker and one voice. The difference between natural variation in speech signals and the nature of the automated techniques are segmenting the waveforms form the audible output. The unit selection synthesis is sub-type of concatenative synthesis approach is details described in next section [6, 7].

Unit Selection Synthesis

Unit selection is the natural extensions of second generation concatenative system. Unit selection synthesis requires large corpus of recorded speech. During corpus of speech creation, each recorded utterances are segmented into the form of digits, vowels, consonants, words and sentences. The segmentation is done using visual representations such as the waveform of speech, pitch track and spectrogram.

An index of the units in the speech database is then created based on the segmentation and acoustic parameters like the fundamental frequency (F0) pitch, duration, position in the syllable, and neighboring phones [8]. In runtime, the desired targeted utterances are determined by the best chain of units from the database (unit selection).

The unit selection provides the extreme naturalness, because a small amount of digital signal processing (DSP) is applied to the recorded speech. DSP often makes recorded speech sound less natural, although some systems use a small amount of signal processing at the point of concatenation to generate the smooth waveform. The unit-selection systems are producing the best output from real human voice. However, it requires large speech database for unit-selection system [9, 10]. The fig.1 represents the process flow of Natural Sounding Speech Synthesizer for Marathi language and is describe below.

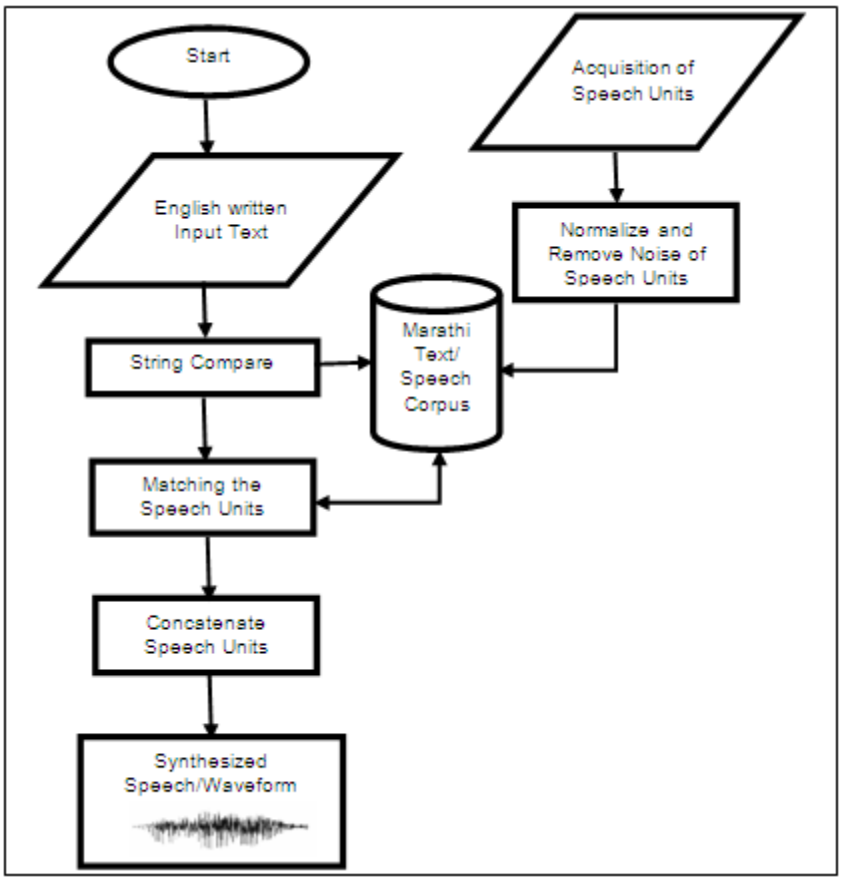

Fig.1. Process Flow of Natural Sounding Speech Synthesizer for Marathi language

Firstly, the input text is taken in the form digits / vowels / consonants / words / sentences. Then corresponding input texts fetched and matching from the speech database. The speech signals are corresponding to the text form speech database. That speech signals were normalized and noise was illuminated through the Audacity, PRAAT and voice activity detection (VAD) algorithm. The units of speech signals were concatenate using the concatenative synthesis approach.

Then system was able to produce the sound related the input text and generate the synthesized speech waveform. The process flow diagram is implementing through the MATLAB tool. Throughout experimental work is carried out to design GUI-based tool for analysis and produce the natural synthetic speech.

IV. Formant Frequency Detection Technique

Formants are defined as the spectral peaks of sound spectrum, of the voice of a person. The speech and phonetics, are acoustic resonance of the human vocal tract is referred as formant frequencies [11]. The vocal tract presents some appropriate pulses are very specious in the spectral of the acoustic signal. These appropriate frequencies constitute the formants of the vocal signal. After calculating the smoothed spectrum, it can extract amplitudes corresponding to the vocal tract resonance. The source vocal tract particularly it peaks corresponding smoothed spectrum to the resonance of the formants. This can be easily obtained by localizing the spectral maxima from frequency bands [12]. The two popular method Cepstral and LPC analysis which have been used for determined the formant frequencies of speech signals.

[A] Cepstral Analysis

In this section a model for formant estimation based on Cepstral analysis is represented. The way of representing the spectral envelope by computing the power spectrum from the fourier transform, by execution of an Inverse Fourier Transform of the logarithm of that power spectrum, and by retaining just the low-order coefficients of this inverse. This overall result is called Cepstrum of the signal. The pitch period is estimated and roundup the log magnitudes are obtained from the Cepstrum. The formants which are estimated from the sharp spectral envelope using the constraints on formant frequency ranges and relative levels of spectral peaks at the formant frequencies. The vocal signal results from the convolution of the source and the contribution of the vocal tract. This technique is designed for separate the barrier of signal components [13].

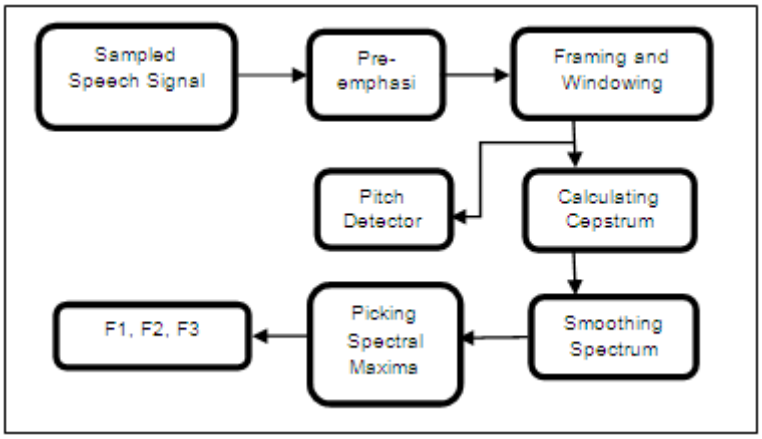

Fig.2. Formant Frequency Estimation Using Cepstral Analysis

The speech signal is represented as;

$$S(n) = g(t) \otimes h(t) \quad (1)$$

Where $\otimes$ denotes convolution, S(n) is the value of the speech signal at the nth point, g(t) and h(t) are contribution of the excitation and vocal tract respectively. The Cepstrum method represents the spectral envelope by computing the power spectrum using Fourier transform of logarithmic of that power spectrum. The Cepstrum method is computed through inverse Fourier transform of the log spectrum. The Cepstrum method expression shown in equation (2).

$$ç(n) = FFT - 1(Log(FFT(s(n)))) \quad (2)$$

Where ç(n) is Cepstrum coefficient of speech signal at the nth point. At this state, the excitation g(t) and the vocal tract shape h(t) are superimposed. It can be separated by conventional signal processing such as temporal filtering. In fact, the low order terms of the Cepstrum contain the information relative to the vocal tract. These two equations are contributed to separate the peak values and peaking the simple temporal windowing.

[B] LPC Analysis

Linear prediction is a good tool for analysis of speech signals. Linear prediction model is the human vocal tract model as treated as infinite impulse response (IIR) system, which produces the speech signal. In speech coding, the success of LPC have been explained by the fact that an all pole model is a reasonable approximation for the transfer function of the vocal tract. All pole models are also suitable in terms of human hearing, because the ear is most sensitive to spectral peaks than spectral valleys. Hence, an all pole model is useful not only because it may be a physical model for a signal, but it is a perceptually expressive parametric representation for a speech signal [14].

The analyzing and estimating the speech signals using formant frequency on the basis of linear predictive coefficient (LPC) technique. The fig.3 shows the LPC based processed step by step. All Cepstral and Linear Predicting Coefficients (12 coefficients) have been computed from pre-emphasized speech signal using 512 points Hamming windowed speech frames [15].

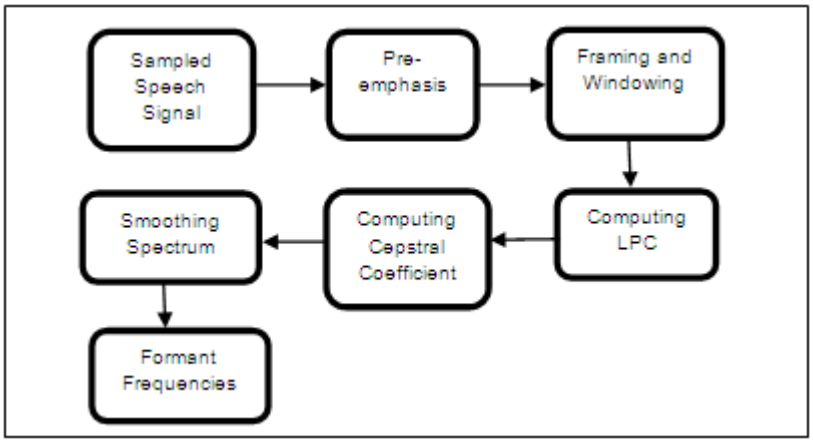

Fig.3. Formant Frequency Estimation Using LPC Analysis

V. Database Acquisition

[A] Text Corpus

The main objective of the text corpora design in this study is to construct a minimum but sufficient speech corpus for concatenative synthesizer system to producing the speaker's natural voice. It is important that created text corpus, words and sentences that frequently appearing in speech corpus which is synthesized naturally and comprehensively.

This research work text corpus is created by 'Marathi Wordnet' because all Marathi linguists are approved the 'Marathi Wordnet' dictionary. The time of selection words and sentences are in order to avoid the repetition of common words, which decreases size of the database, but enhance the overall quality of speech synthesizer. The some natural words with sentences are shown in Table-II. All Marathi words, sentences used in this research work are taken from Marathi Wordnet dictionary, which is consider to be standard in the Marathi language. For present research study around 28580 natural syllables and phonemes are selected in testing phase.

Table 2. Marathi Words and Sentences written in English and Marathi script

Marathi words written in English script	Marathi words in phonetic form	Marathi Sentences written in English Script	Marathi Sentences written in Devnagari script
aachari	आचारी	Aachari plyane varan halvat hota	आचारी पळ्याने वरण हलवत होता.
aajoba,	आजोबा	Aajobansathi nave dhotrache pan aanale	आजोबांसाठी नवे धोतराचे पान आणले.
zadala	झाडाला	Zadala hirvya rangachi pane yetat	झाडाला हिरव्या रंगाची पाने येतात.
antaralat.	अंतराळात	Antaralat khup sury ahet	अंतराळात खूप सूर्य आहेत.
bharat,	भारत	Bharat v Pakistan yanchyatil criketcha samana changlach rangto	भारत व पाकिस्तान यांच्यातील क्रिकेटचा सामना चांगलाच रंगतो.
chandichya,	चांदीच्या	Mithai chandichya varkhat gundalleli hoti	मिठाई चांदीच्या वर्खात गुंडाळलेली होती.

[B] Speech Corpus

The phonetically rich natural Marathi words and sentences are taken form the text corpus. These phonemes are spoken by only female speaker of age 22-year. The concatenative speech synthesis is restricted to only one speaker and one voice for producing natural and intelligible speech signals. These natural words and sentences were acquired through Standard Sennheiser HD-449 Wired Headphone. That speech corpus data acquisition was done in 12' X 10' X 12' lab. The speech corpus was acquired in normal room temperature through PRAAT tool with sampling frequency of 22 KHz. These natural syllables are spoken in continuous rhythm with small gap between two successive words. The speech corpus is divided in the form of Marathi numbers 1-100-319, Vowels-12, Consonants-34, Words-5058 and SentencesX1855 were stored in .wav file format. The size of speech corpus was 1.2 GB.

VI. Experimental Work, Results And Discussion

The formant frequencies have been determined by [A] Cepstral and [B] LPC analysis. Experiment was done through standard Praat and Mat lab Tools. The LPC techniques have been used for estimating the formant frequencies which is denoted as F1, F2, and F3. Various undefined signal have been ignored. For obtaining the formants, it has done the experiment using different values for the prediction order and varying the degree of pre-emphasis. When dealing with windowing speech it need to take into account the boundary effects in order to avoid large prediction errors at the edges. When it defines the area over perform speech at minimization.

The synthesized speech signals which are computed Cepstral with LPC coefficient that smoothing spectrum is denoted as formants F1, F2, and F3. This experimental work it decides some samples are taken for analysis. These samples are described to produce synthesized speech signals. These samples are Marathi Numbers, vowels and words for extracting the formant frequencies on the basis of LPC analysis using Praat and Mat lab Tool.

The string and speech units are matching and comparing from the speech and text corpus. The each speech units are selecting from the speech corpus then concatenate and produce the natural synthetic speech. The Fig. 4 and Fig. 5 indicated the formant pitch track and waveform of Marathi spoken numbers '**पाच**' and '**नऊ हजार आठशे सदुसष्ट**' respectively. Similarly the waveform and Formant Track of Marathi Spoken Vowel '**ओ**'

and Marathi Word '**आचारी**' is shown in Fig.6 and Fig.7 respectively. The table III is obtained result for first three formants frequencies of synthesized speech for Marathi numbers with its duration.

This experiment which have used PRAAT tool for synthesized the speech signals of number, vowels and words. In this order it has estimated the formant frequency (F1, F2, and F3) values are varies from 540 to 2992 and its standard deviation values are varies from 45 to 616 respectively. The LPC based MATLAB tool which have used the for synthesized the speech signals of number, vowels and words. In this order it has estimated the formant frequency (F1, F2, and F3) values are varies from 145 to 831 and its standard deviation values are varies from 43 to 138 respectively. The estimated formant frequencies (F1, F2, and F3) are determined by PRAAT and MATLAB tool is denoted as PT and MT. These results were found to be satisfactory, which is shown in Table-III-VII. The implemented system provides the good accuracy and produces the high quality synthesized speech.

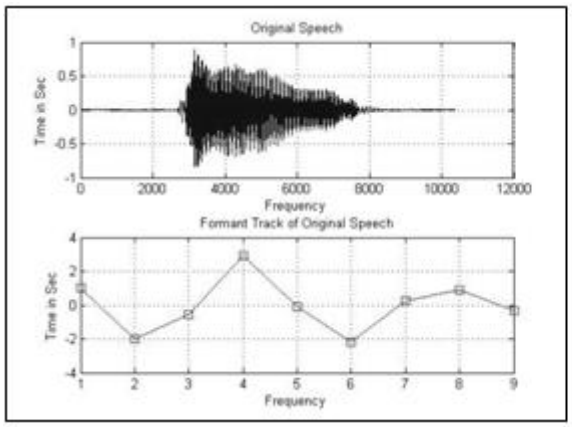

Fig.4. Waveform and Formant Track of Synthesized Speech for Marathi Number '५'

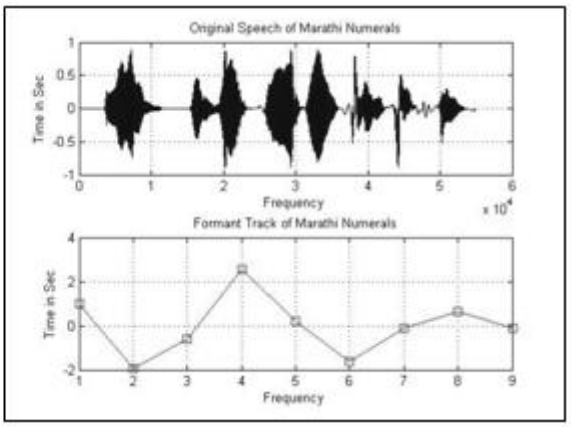

Fig.5. Waveform and Formant Track of Synthesized Speech for Marathi Number '१८६७'

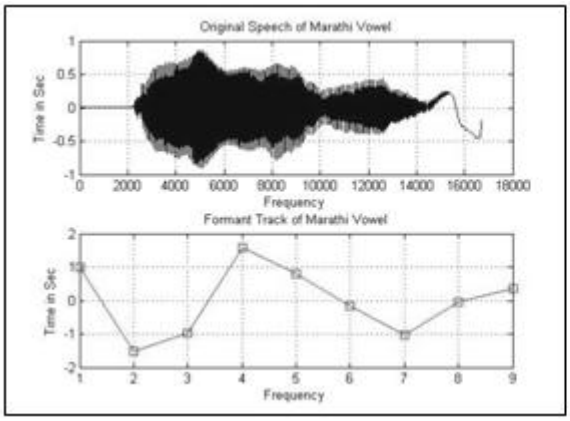

Fig.6. Waveform and Formant Track of Synthesized Speech for Marathi Vowel '**ओ**'

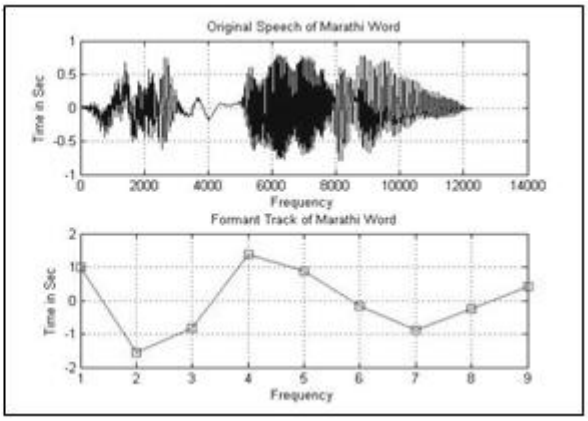

Fig.7. Waveform and Formant Track of Synthesized Speech for Marathi Word '**आचारी**'

Table 3. Estimated Format Frequencies of Marathi Vowels by LPC using PRAAT (PT) Tool.

English Numbers	Marathi Numbers	Marathi Numbers in Phonetic form	Duration in Sec.	F1	F2	F3
8	८	आठ	0.30	596.67	914.37	1773.43
39	३९	एकोणचाळीस	0.69	435.55	974.56	2718.47
44	४४	चौरेचाळीस	0.73	438.74	714.38	2665.99
135	१३५	एकशे पस्तीस	1.30	707.40	1526.45	3237.21
475	४७५	चारशे पंचाहत्तर	1.49	632.48	1200.04	2518.39
838	८३८	आठशे अडोतीस	1.09	431.92	946.94	2722.69
1785	१७८५	एक हजार सातशे पंच्याऐशी	2.08	520.93	1012.98	2668.22
2699	२६९९	दोन हजार सहाशे नव्याण्णव	2.22	459.41	912.24	2640.74
4555	४५५५	चार हजार पाचशे पंचावन्न	2.27	623.78	1034.33	2260.65
5000	५०००	पाच हजार	0.93	607.38	934.06	1996.00
6689	६६८९	सहा हजार सहाशे एकोणनव्वद	2.20	502.87	894.80	2434.45
9867	९८६७	नऊ हजार आठशे सदुसष्ट	2.39	523.82	1193.40	2720.16
Mean				**540.08**	**1021.55**	**2529.70**
Standard Deviation(SD)				**87.89**	**197.03**	**364.94**

Table 4. Estimated Format Frequencies of Marathi Vowels by LPC using PRAAT (PT) Tool.

English Written Marathi Vowels	Marathi Vowels in Phonetic form	Duration in Sec.	F1	F2	F3
a	अ	0.54	570.26	750.73	1098.7
aa	आ	0.71	812.73	995.19	1466.48
i	इ	0.73	288.75	668.44	3108.02
ee	ई	0.74	305.88	648.03	3056.11
u	उ	0.58	335.91	594.73	2717.26
oo	ऊ	0.59	334.59	657.51	2305.26
ae	ए	0.66	445.38	657.55	2231.12
aae	ऐ	0.66	444.89	661.57	2254.37
o	ओ	0.75	494.88	805.54	1532.29
ou	औ	0.68	432.32	664.74	2423.56
am	अं	0.54	496.10	842.48	1608.33
aha	अः	0.50	727.88	947.20	1775.34
Mean			**474.13**	**741.14**	**2131.40**
Standard Deviation(SD)			**156.49**	**123.51**	**616.49**

Table 5. Estimated Format Frequencies of Marathi Words by LPC using PRAAT (PT) Tool.

English Written Marathi Words	Marathi Words in phonetic form	Duration in Sec.	F1	F2	F3
Aachari	आचारी	0.55	468.37	947.29	2760.29
Aajoba	आजोबा	0.48	397.92	700.61	2769.87
Zadala	झाडाला	0.63	566.78	1135.21	3291.71
Vastu	वस्तु	0.47	421.96	1148.27	3559.05
Antaralat	अंतराळात	0.65	472.41	745.98	2308.18
Aadhar	आधार	0.36	474.64	888.78	3254.87
Balachi	बाळाची	0.75	503.02	1119.99	2821.13
Banachi	बाणाची	0.67	520.60	1351.08	2998.12
Bharat	भारत	0.39	441.13	1156.05	3125.76
Chandichya	चांदीच्या	0.60	465.74	921.34	2809.42
Criketacha	क्रिकेटचा	0.62	410.10	1136.36	3144.37
Davpech	डावपेच	0.47	464.96	999.27	3065.15
Mean			**467.30**	**1020.85**	**2992.33**
Standard Deviation(SD)			**45.58**	**180.54**	**310.93**

Table 6. Estimated Format Frequencies of Marathi Numerals by LPC using MATLAB (MT) Tool.

English Numbers	Marathi Numbers	Marathi Numbers in Phonetic form	Duration in Sec.	F1	F2	F3
8	८	आठ	0.30	213.49	289.41	644.67
39	३९	एकोणचाळीस	0.69	102.96	228.77	812.80
44	४४	चौरेचाळीस	0.73	159.28	221.64	683.77
135	१३५	एकशे पस्तीस	1.30	162.59	248.51	875.83
475	४७५	चारशे पंचाहत्तर	1.49	131.37	237.46	768.37
838	८३८	आठशे अडोतीस	1.09	123.55	226.59	854.25
1785	१७८५	एक हजार सातशे पंच्याऐशी	2.08	150.12	282.87	632.12
2699	२६९९	दोन हजार सहाशे नव्याण्णव	2.22	167.47	306.35	875.34
4555	४५५५	चार हजार पाचशे पंचावन्न	2.27	176.92	286.99	577.39
5000	५०००	पाच हजार	0.93	217.38	288.63	570.08
6689	६६८९	सहा हजार सहाशे एकोणनव्वद	2.20	264.70	271.09	872.86
9867	९८६७	नऊ हजार आठशे सदुसष्ट	2.39	141.30	273.20	626.15
Mean				**167.59**	**263.46**	**732.80**
Standard Deviation(SD)				**45.35**	**29.30**	**122.52**

Table 7. Estimated Format Frequencies of Marathi Vowels by LPC using MATLAB (MT) Tool.

English Written Marathi Vowels	Marathi Vowels in Phonetic form	Duration in Sec.	F1	F2	F3
a	अ	0.54	209.59	263.51	564.94
aa	आ	0.71	311.10	413.73	671.39
i	इ	0.73	100.88	240.20	831.95
ee	ई	0.74	104.90	116.14	847.33
u	उ	0.58	149.31	220.40	831.22
oo	ऊ	0.59	110.16	221.08	851.47
ae	ए	0.66	158.93	249.86	545.71
aae	ऐ	0.66	158.86	250.64	578.24
o	ओ	0.75	166.87	273.64	847.62
ou	औ	0.68	165.83	285.89	554.35
am	अं	0.54	219.66	327.77	559.48
aha	अः	0.50	273.08	366.25	561.33
Mean			**177.43**	**269.09**	**687.09**
Standard Deviation(SD)			**65.51**	**76.13**	**140.44**

Table 8. Estimated Format Frequencies of Marathi Words by LPC using MATLAB (MT) Tool

English Written Marathi Words	Marathi Words in phonetic form	Duration in Sec.	F1	F2	F3
Aachari	आचारी	0.55	93.62	295.22	862.11
Aajoba	आजोबा	0.48	93.52	233.07	851.31
Zadala	झाडाला	0.63	103.01	275.13	843.23
Vastu	वस्तु	0.47	118.92	272.58	753.80
Antaralat	अंतराळात	0.65	158.28	249.57	830.39
Aadhar	आधार	0.36	254.93	269.90	773.24
Balachi	बाळाची	0.75	107.61	275.94	858.01
Banachi	बाणाची	0.67	266.22	267.52	857.79
Bharat	भारत	0.39	102.01	269.17	854.11
Chandichya	चांदीच्या	0.60	91.99	315.96	781.73
Criketacha	क्रिकेटचा	0.62	199.95	259.52	856.32
Davpech	डावपेच	0.47	150.7	165.64	852.93
Mean			**145.06**	**262.44**	**831.25**
Standard Deviation(SD)			**63.10**	**36.83**	**38.56**

VII. Conclusion

This research work has reported the implementation of Natural Sounding Speech synthesizer for Marathi. The important feature of concatenative speech synthesizer is restricted to only one speaker and one voice for producing natural and intelligible speech signals. The throughout experiment was carried out by DSP tool available in MATLAB software. The formant frequency results which are determined by formant detection techniques through LPC and Cepstral analysis using MATLAB and PRAAT tool. The synthesized speech signals are extracted from the formant detection technique that separated the peaks are denoted as F1, F2 and F3 that results are reported in table 3-8. This results are given good and high quality performance, with the help it produce the natural synthetic speech and generated the waveform of corresponding input text.

Acknowledgements

The authors are indebted and thankful to Dr. S. C. Mehrotra, Professor, (Ramanujun Geospatial Chair), Department of Computer Science & IT, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad (MS) for valuable suggestion and discussion. Author G. D. Ramteke would like to thankful for providing the financial support of Raisoni Doctoral Fellowship, North Maharashtra University, Jalgaon.

References

[1]. A. Watanabe, "Formant estimation method using inverse-filter control, " IEEE Trans. Speech and Audio Processing, vol. 9, 2001, pp. 317-326.
[2]. Roy C. Snell and Fausto Milinazzo, "Formant location from LPC analysis data", IEEE Transaction on Speech and Audio Processing, Vol.1, No-2, 1993.
[3]. R.K. Bansal, and J.B. Harrison, "Spoken English for India, a Manual of Speech and Phonetics"', Orient Longman, 1972.
[4]. Aimilios Chalamandaris, Sotiris Karabetsos, "A Unit Selection Text-to-Speech Synthesis System Optimized for Use with Screen Readers", IEEE Transactions on Consumer Electronics, Vol. 56, No. 3, 2010, pp.1890-1897.
[5]. Akemi Iida, Nick Campbell," A database design for a concatenative speech synthesis system for the disabled", ISCA, ITRW on Speech Synthesis, 2001.
[6]. A.Chauhan, Vineet Chauhan, Gagandeep Singh, Chhavi Choudhary, Priyanshu Arya, "Design and Development of a Text-To-Speech Synthesizer System", IJECT Vol. 2, Issue 3, Sept. 2011.
[7]. Chomtip Pronpanomchai, Nichakant Soontharanont, Charnchai Langla and Narunat Wongsawat, "A Dictionary-Based Approach for Thai Text to Speech (TTTS)", The Third International Conference on Measuring Technology and Mechatronics Automation, pp. 40 – 43, 2011.
[8]. D. J. Ravi and Sudarshan Patilkulkarni, "A Novel Approach to Develop Speech Database for Kannada Text-to-Speech System", Internal Journal on Recent Trends in Engineering & Techonology, Vol. 05, No. 01, 2011, pp. 119-122.
[9]. Muhammad Masud Rashid, Md. Akter Hussian, M. Shahidur Rahman, "Text Normalization and Diphone Preparation for Bangla Speech Synthesis", Joural of Multimedia, Vol. 5, No. 6, 2010, pp. 551-559.
[10]. Mustafa Zeki, Othman O. Khalifa, A. W. Naji, "Development of An Arabic Text-To-Speech System", International Conference on Computer and Communication Engineering (ICCCE 2010), 2010, pp. 1-5.
[11]. L. Welling & H. Ney, "Formant estimation for speech recognition," IEEE Trans. Speech and Audio Processing, vol.6, 1998, pp. 36-48.
[12]. Patti Adank, Roeland van Hout, and Roel Smits, "A comparison between human vowel normalization strategies and acoustic vowel transformation techniques", 2001 Euro speech.
[13]. D.Gargouri, Med Ali Kammoun and Ahmed ben Hamida, "Formants Estimation Techniques for Speech Analysis", International Conference on Machine Intelligence, Tozeur–Tunisia, Nov.5-7, 2005, pp. 96-100.
[14]. D.Gargouri, Med Ali Kammoun and Ahmed ben Hamida,"A Comparative Study of Formant Frequencies Estimation Techniques", Proceedings of the 5th WSEAS International Conference on Signal Processing, Istanbul, Turkey, 2006, pp15-19.
[15]. Parminder Singha, Gurpreet Singh Lehal, "Text-To-Speech Synthesis System for Punjabi Language", 2011, pp. 140 – 143.
[16]. G. D. Ramteke, S. S. Nimbhore, R. J. Ramteke, "De-noising Speech of Marathi Numerals Using Spectral Subtraction" , Advances in Computational Research, ISSN: 0975-3273 & E-ISSN: 0975-9085, Volume 4, Issue 1, 2012, pp.-122-124.
[17]. S. S. Nimbhore, G. D. Ramteke, R. J. Ramteke, "Pitch Estimation of Devnagari Vowels using Cepstral and Autocorrelation Techniques for Original Speech Signal", International Journal of Computer Applications Journal(0975-8887), Volume 55 - Number 17, Oct-2012, pp.38-43.
[18]. Knowledge is treasure that will follow that will follow its owner everywhere, available at: http://www.balmitra.com.
[19]. http:// www.cfilt.iitb.ac.in/wordnet/
[20]. http://www.learn101.org/marathi_phrases.php.