VAACHANA - A KANNADA SPEECH SYNTHESIZER

PROJECT REPORT

Submitted by

KARTHIK T. **KRISHNARAJ BARVATHAYA B.**

NANDEESHA G.M. UTHAM H.

In partial fulfillment of the requirements for the degree of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

(Visveswaraiah Technological University)

Under the Guidance of

Mr. D.S. RAJESH. B.E, M.Tech

Assistant Professor

Department of Computer Science and Engineering

K.V.G. College of Engineering

SULLIA, D.K. 574327, KARNATAKA

2003-2O04

K.V.G COLLEGE OF ENGINEERING
SULLIA, D.K. 574 327 - KARNATAKA

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

*Certified that the project work entitled "**VAACHANA - A KANNADA SPEECH SYNTHESIZER**" is a bonafide work carried out by*

KARTHIK T.	**Reg. No. 4KV00CS031**
KRISHNARAJ BARVATHAYA B.	**Reg. No. 4KV00CS035**
NANDEESHA G.M.	**Reg. No. 4KV00CS042**
UTHAM H.	**Reg. No. 4KV00CS093**

*in partial fulfillment for the award of degree of "**BACHELOR OF ENGINEERING**" in **COMPUTER SCIENCE AND ENGINEERING*** of the ***VISVESWARAIAH TECHNOLOGICAL UNIVERSITY, BELGAUM*** *during the year* 2003-2004. *It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the Report deposited in the departmental library. The project report has been approved as it satisfies the academic requirements in respect of Project work prescribed for the Bachelor of Engineering Degree.*

Mr. D.S. RAJESH.	**Prof. SRIKANTA MURTHY K.**	**Prof. S.G. GOPALA KRISHNA**
B.E, M.Tech	**B.E. M.Tech, (Ph.D)MISTE, AMIE**	**B.E., M.Tech,(Ph.D) MIE, MISTE**
Assistant Professor	**Head of the Department**	**Principal**
Project Guide		

Name of the Student :
University Seat No :

Name of the Examiners **Signature with Date**

1.

2.

ABSTRACT

ABSTRACT

Most of the information in world of computer is accessible to a few who can read or understand a particular language. But it could be very much helpful for the common man if the computer talks to him in his language.

Text-To-Speech (TTS) is a technology that converts a written text into human understandable voice. A TTS synthesizer is a computer-based system that can be able to read any text aloud that is given through standard input devices. VAACHANA is particularly a Kannada text to speech converting application, which serves the localization of computer application.

ACKNOWLEDGEMENT

ACKNOWLEDGEMENT

Motivation and guidance is the key towards the success. We would like to thank all the source of motivation and who guided us with due respect and gratitude.

We express our deep sense of gratitude to our project guide Mr. D. S. RAJESH, Assistant Professor, Department of Computer Science and Engineering for his timely help and advice.

We are extremely grateful to our project coordinator Mr. NARENDRA U. P, Lecturer, Department of Computer Science and Engineering who encouraged us to complete this project in time.

We extend our heart full thanks to our beloved HOD Prof. SRIKANTA MURTHY K, Computer Science and Engineering who has been the constant driving behind the completion of this project.

We are extremely grateful to our beloved Principal Prof. S.G. GOPALA KRISHNA, who encouraged us.

We are indebted to Dr. KURUNJI VENKATRAMANA GOWDA, President AOLE for having provided an environment with all facilities that helped us in completing our project.

Finally we would like to extend our thanks to all teaching and non-teaching staff-members of Computer Science and Engineering for their co-operation in completing this project.

CONTENTS

CONTENTS

CHAPTER 1
INTRODUCTION

CHAPTER 1

INTRODUCTION

Today there is a wide spread talk about improvement of the human interface to the computer. Because no longer people want to sit and read data from the monitor. Since there is a painstaking effort to be taken, this involves strain to their eyes. In this aspect Speech Synthesis is becoming one of the most important steps towards improving the human interface to the computer.

The art of making PC's talk has always entranced the human community. After all, voice is one of the best alternatives for hours of eyestrain involved in going through any document. Also Voice is a better interface when it comes to illiterate people rather than Graphic User Interface in English. So research is being done through out the world for improving the Human Interface to the computer and one of the best options found out till date is the ability of a computer to speak to humans. Here comes the role of the Text To Speech (TTS) engines. Text-To-Speech is a process through which input text is analyzed, processed and “understood”, and then the text is rendered as digital audio and then “spoken”. It is a small piece of software, which will speak out the text inputted to it, as if reading from a newspaper. There have been many developments found around the world in the development of TTS Engines in various languages like English, French, German etc and even in Hindi. This has not been tried out till date, (according to our knowledge) in Kannada. So here is the first step towards making computers speak to Kannadigas around the world.

Types of TTS Systems

Most Text To Speech engines can be categorized by the method that they use to translate phonemes into audible sound. Some TTS Systems are listed below:-

Prerecorded

In this kind of TTS Systems we maintain a database of prerecorded words. The main advantage of this method is good quality of voice. But limited vocabulary and need of large storage space makes it less efficient.

Formant

Here voice is generated by the simulation of the behavior of human vocal cord. Unlimited vocabulary, need of low storage space and ability to produce multiple featured voices makes it highly efficient, but robotic voice, which is sometimes not appreciated by the users.

Concatenated

This is the technique, which is implemented in Vaachana. In this kind of TTS systems, text is phonetically represented by the combination of its syllables. These syllables are concatenated at run time and they produce phonetic representation of text. Key features of this technique are unlimited vocabulary and good voice. But it can't produce multiple featured voices, needs large storage space.

Various methodologies of implementation, prospects and challenges of implementation of a Kannada TTS engine with regard to speech synthesizer and its high level applications are presented here. The Implementation of this TTS is done using the concatenation method. Integral parts of a Text To Speech engine are phoneme identifier, voice mapping and speech synthesizer.

Different applications of TTS in our day-to-day life

Telephony

Automation of telephone transactions (e.g., banking operations), automatic call centres for information services (e.g., access to weather reports), etc.

Automotive

Information released by in-car equipments such as the radio, the air conditioning system, the navigation system, the mobile phone (e.g., voice dialing), embedded telematic systems, etc.

Multimedia

Reading of electronic documents (web pages, emails, bills) or scanned pages (output of an Optical Character Recognition system).

Medical

Disabled people assistance: personal computer handling, demotic, mail reading.

Industrial

Voice-based management of control tools, by drawing operator's attention on important events divided among several screens.

CHAPTER 2
SYSTEM ANALYSIS

CHAPTER 2

SYSTEM ANALYSIS

2.1 IDENTIFICATION OF THE NEED

Language technologies can provide solutions in the form natural interfaces so that digital content can reach to the masses and facilitate the exchange of information across different people speaking different languages. There are already many speech synthesizers existing for English. But there is no standard speech synthesizer for Kannada. Government of Karnataka has made several efforts to make village people to take advantages of advances in the field of IT. But still there exists DIGITAL DIVIDE between the people who can use computer and who cannot, because of illiteracy. “VAACHANA – A Kannada Speech Synthesizer” has been developed for the aid of such people, which can read out the text, which is in Kannada.

2.2 PRELIMINARY INVESTIGATION

In current scenario there are only few softwares developed for Kannada language in the field of speech. If such software is developed it will be very helpful to many areas such as Localization (l10n). In this aspect TTS become one of the most important step in computer field. VAACHANA primarily concerns the extraction of characters based on the corresponding Unicode’s and generate sound.

2.3 FEASIBILITY STUDY

2.3.1 TECHNICAL

VAACHANA has been tested repeatedly for various Kannada words. Conclusion drawn from the tests is that, it is easily able to pronounce the simple Kannada letters. But it finds difficult to pronounce some complex Kannada words like ಜೌತ , ಯಜ್ಞ .

The talking speed of VAACHANA is little bit lagging when compared to the natural speed of the human beings.

2.3.2 ECONOMICAL

The system is implemented in Windows XP Platform. Since JDK1.4.2, is used as developing tool, it can be run in any platform like LINUX, Macintosh etc.

2.3.3 OPERATIONAL

A rich Graphical User Interface is provided in VAACHANA. So the novice user need not require extra knowledge to use the system. However a new user may find difficult in typing the Kannada letters.
To overcome this problem a suitable help file is provided in this package.

CHAPTER 3

SOFTWARE REQUIREMENT SPECIFICATION

CHAPTER 3

SOFTWARE REQUIREMENT SPECIFICATION

In order to read aloud the typed text the Kannada speech synthesizer should contain the following two components, viz.

- Text processing component
- Speech generation component.

The objective of text processing component is to process the given input text and produce appropriate sequence of phonemic units.
These phonemic units are realized by speech generation components. Speech generation component then creates the waveform for the identified Kannada characters.

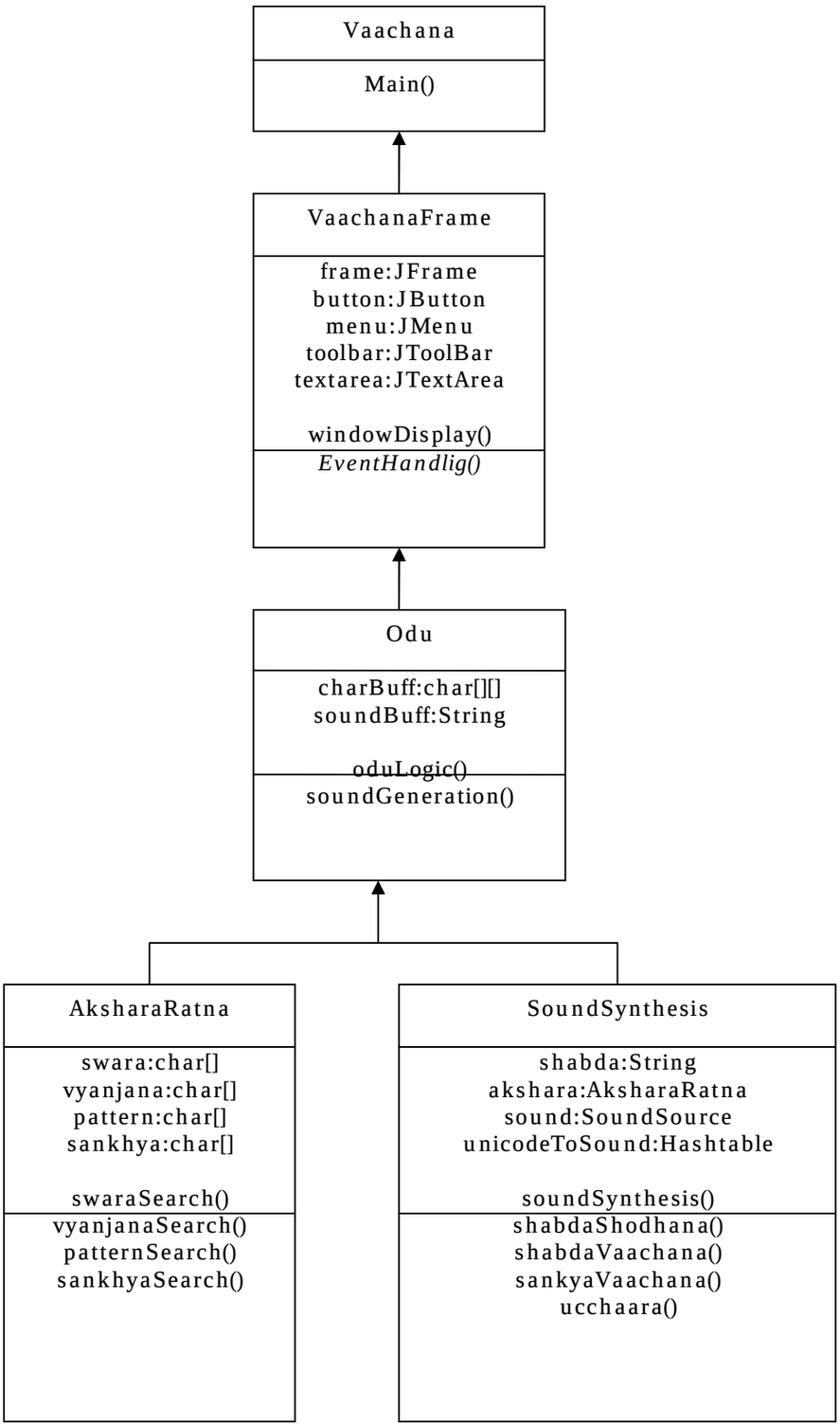
Software Requirements

- Operating system: Windows XP/ Fedora core-I
- Software: jdk 1.4.

Hardware Requirements

- Any x86 class processor
- 32 MB RAM
- 1 GB Hard Disk Space.
- Speaker connected to the computer.

UML CLASS DIAGRAM

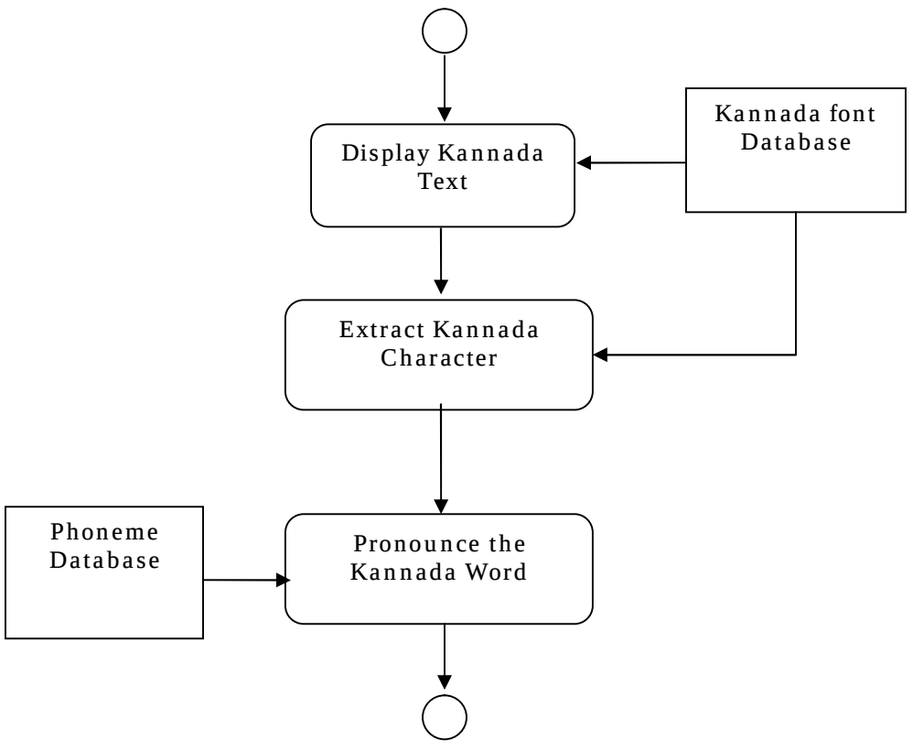


CHAPTER 4
DESIGN

CHAPTER 4

DESIGN

4.1 DATA FLOW DIAGRAM



4.2 MODULES OF VAACHANA

4.2.1 Word Processor module

This is similar to a normal Word Processor, which include Kannada fonts (e.g. Tunga) and is used to enter text phonetically. This will get the text entry from the user who wants the data to be spoken aloud.

Kannada language is implemented using Unicode standard. Unicode is the latest solution for Internationalization (i18n) and Localization (l10n) of computer application. JAVA is one of the programming languages that support Unicode completely.

There are many encoding standards used in text documentation because Unicode supports large number of scripts. If Unicode is used documents can be opened in different systems having different Operating system and architecture.

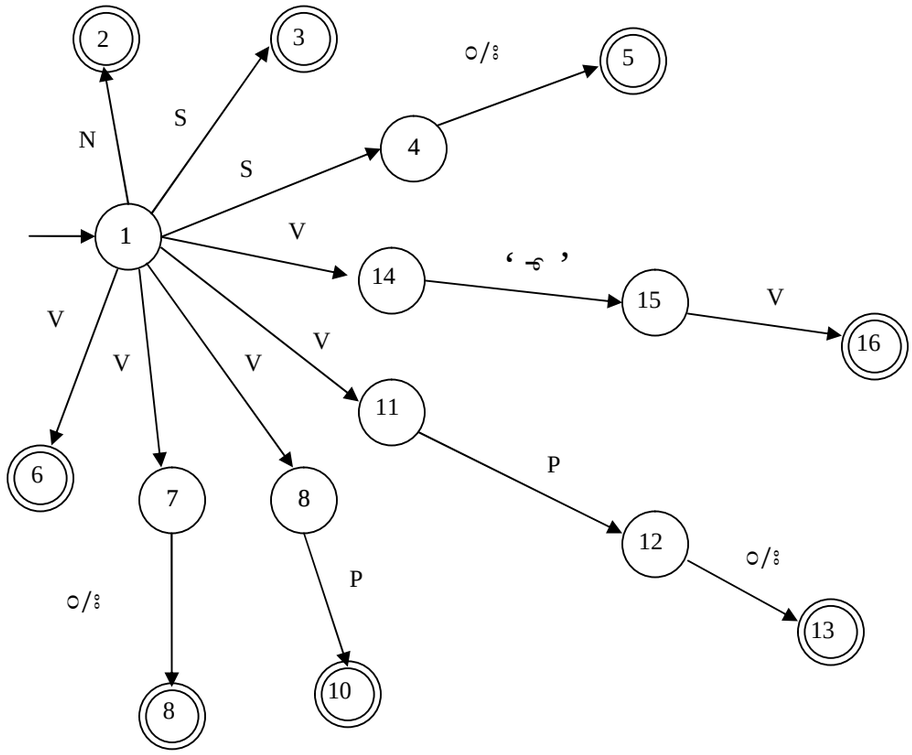
4.2.2 Phoneme identifying module

Phonemes are basic sound units of a language. There are various phonemes in a language, which form the basics of talking that language. Kannada language has around 13 vowels (Swara), around 34 consonants (Vyanjana) and two Anuswara and Visarga. The basic units of the writing system in Indian languages are characters which are an orthographic representation of speech sounds. A character in Indian language scripts is close to a syllable and can be typically of the following form: C, V, CV, VC, CCV and CVC, where C is a consonant and V is a vowel.

Kannada script is phonetic in nature. There is almost one to one correspondence between what is written and what is spoken. When a text sentence is input into word processor, it will divide the various words into their respective phonemes by parsing the various characters and their neighboring characters.

For example, ಪ್ರಕೃತಿ ಸೌಂದರ್ಯ can be parsed as the combination of ಪ್ರ, ಕ್ರ, ತಿ, , ಸೌ, ಂ, ದ, ರ್ಯ . Here letters delimited by comma is extracted from the original word.

4.2.3 Kannada Character identification model



V – Vyanjana S – Swara

P – Pattern (್, ಾ, ೀ,.....) N – Number

“ಕ್ರೌರ್ಯ”, is first split into “ಕ್ರೌ”, “ರ್ಯ”

ಕ್ರೌ = ಕ + ್ + ರ + ೌ

ರ್ಯ= ರ + ್ + ಯ

4.2.4 Sound generating module:

This part consists of two sub segments.

i) Data Base Mapper.

ii) Sound Player.

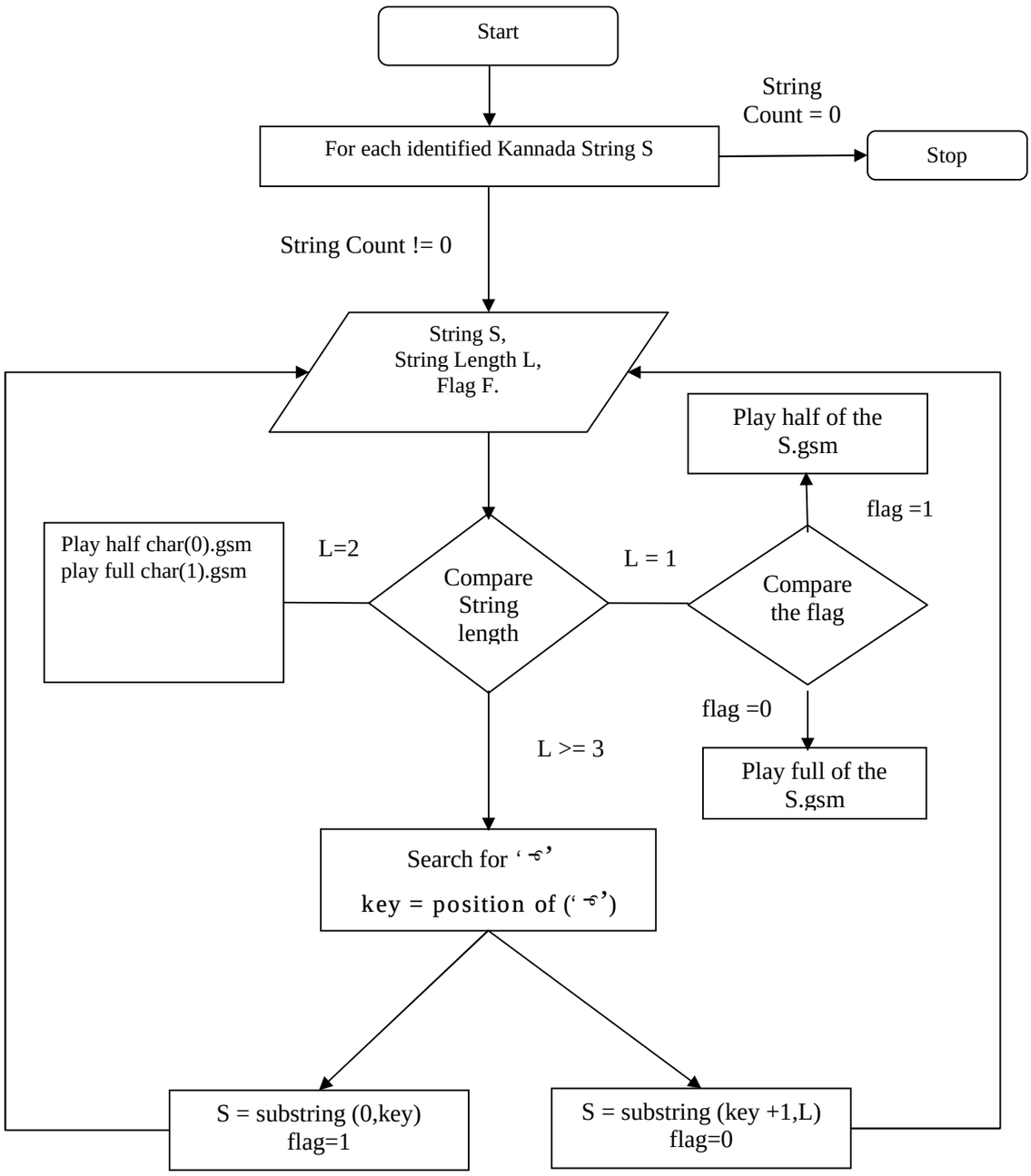
The Data Base Mapping part of the TTS is the One, which maps the various phonemes to their corresponding sounds present in the sound database. Thus after this phase, the output will be the sound file

of the corresponding phonemes. Then the Sound Player will play these sound files, according to the speed required for each of the letters.

Database Details

The database has the following structure. All sound files stored in the database are gsm compressed .gsm files (The GSM standard by The Communications and Operating Systems Research Group (KBS) at the Technische Universitaet Berlin) recorded at 16KHz as 16bit signed linear samples.

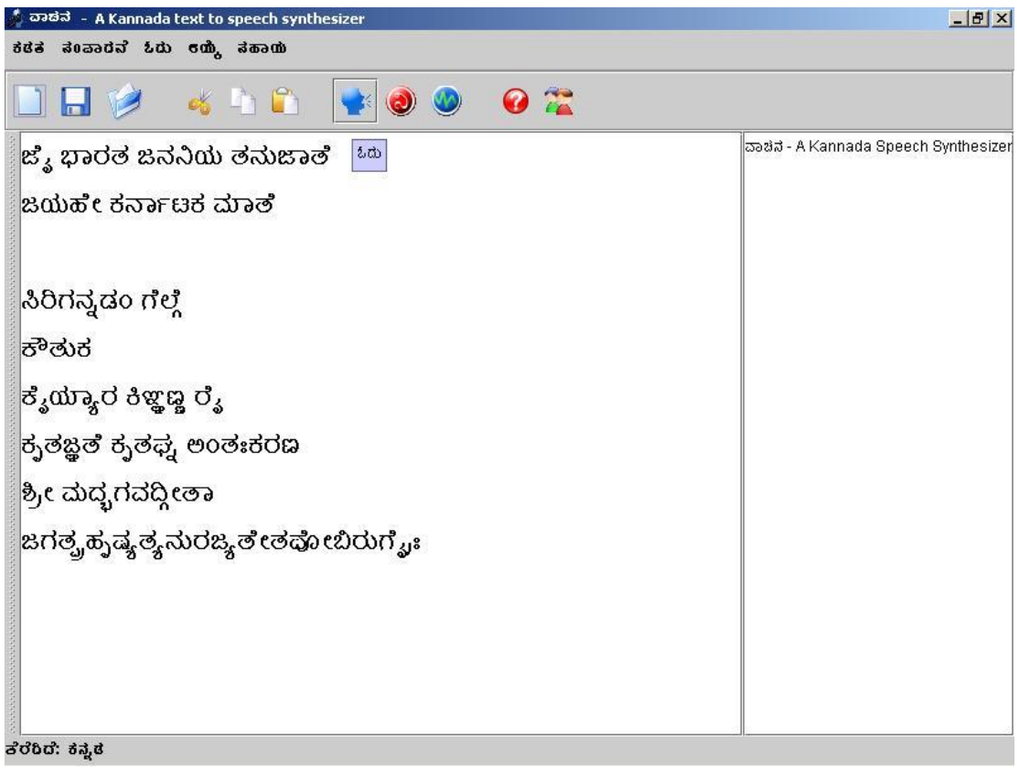
4.3 FLOW CHART FOR PLAYING THE KANNADA WORDS

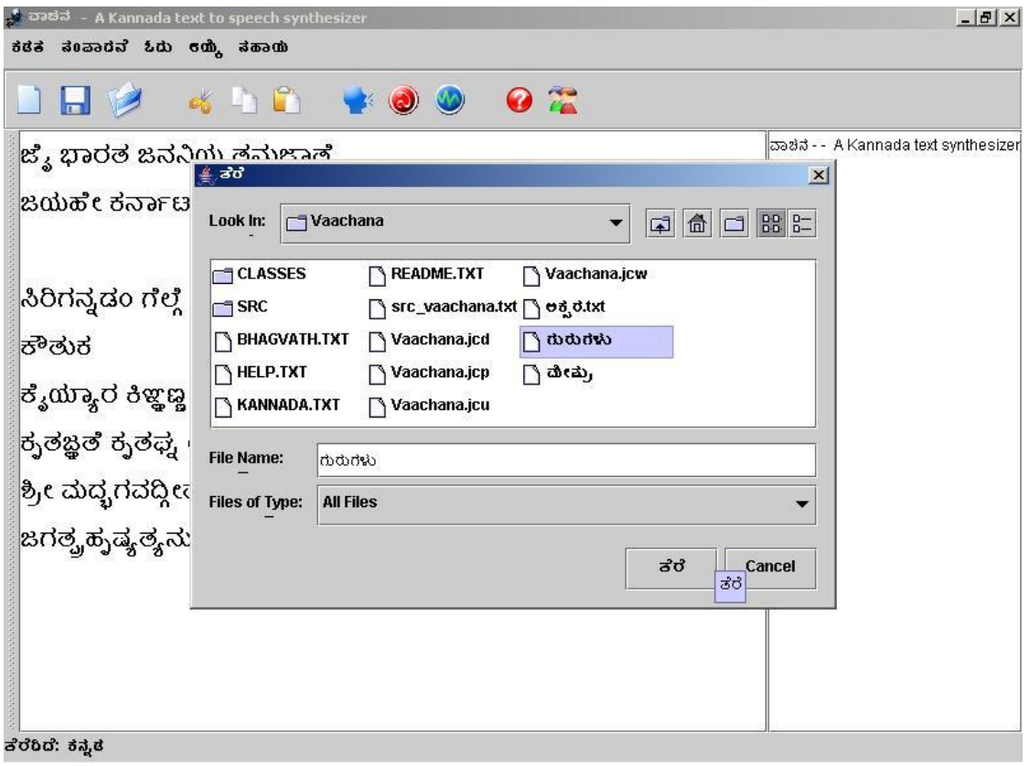


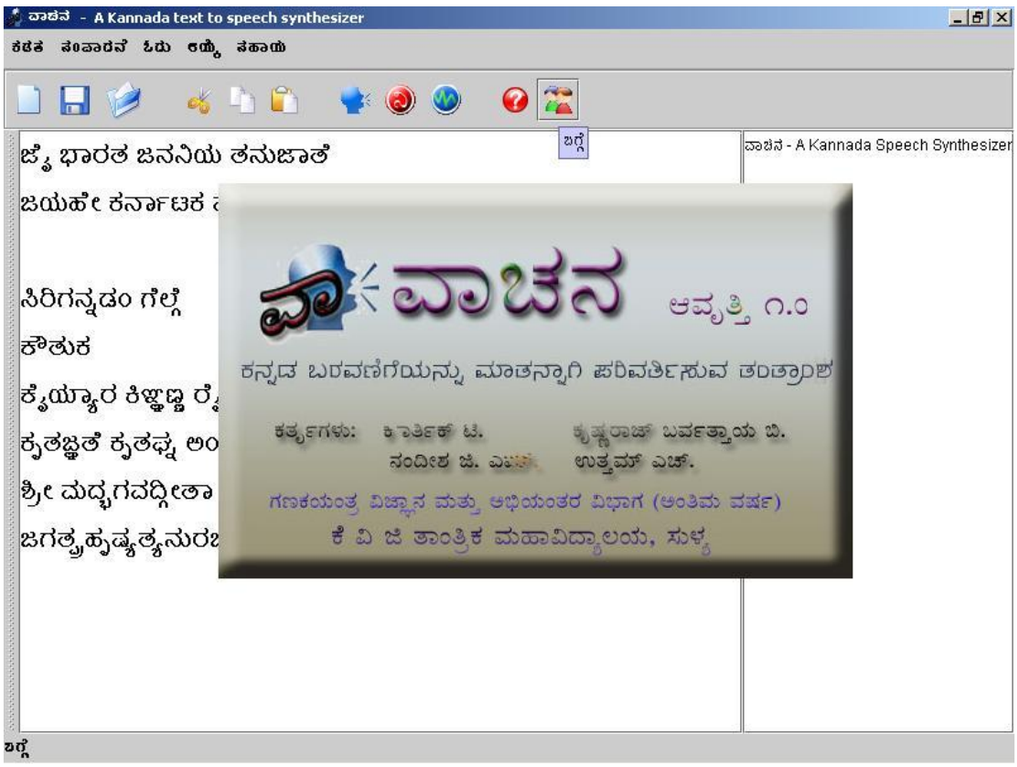
CHAPTER 5
SCREEN LAYOUT

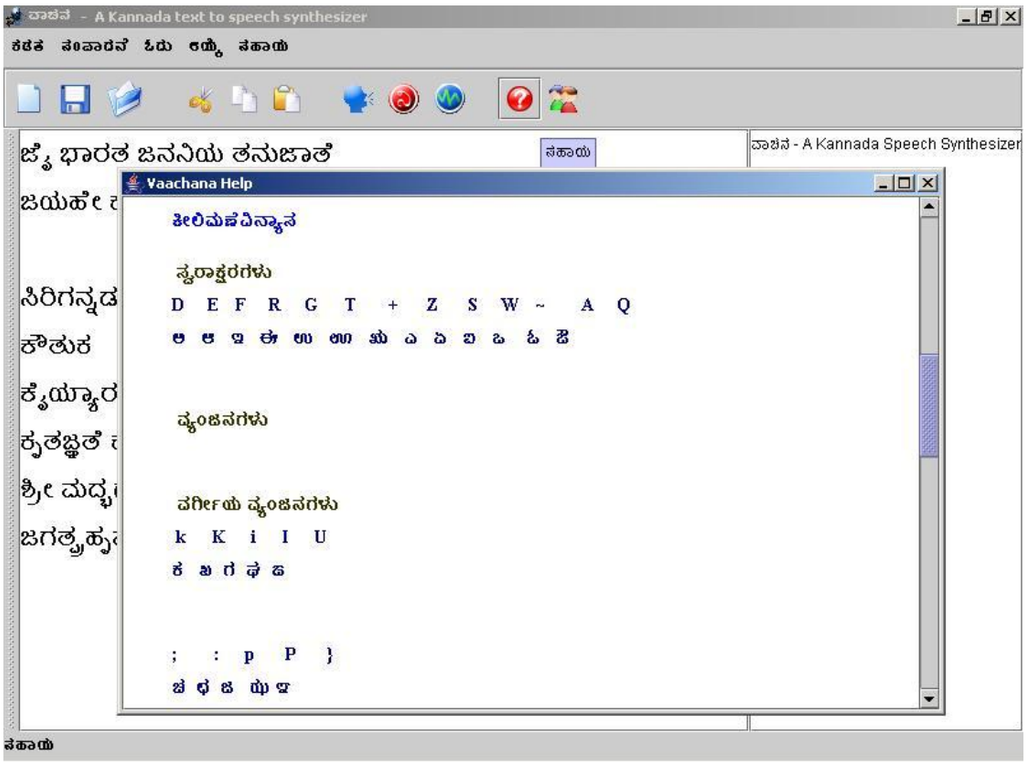
CHAPTER 5

SCREEN LAYOUT




ವಾಚನ - A Kannada text to speech synthesizer
ಕಡತ ಪಂಪಾರವೆ ಓದು ರಯ್ಕೆ ಪಹಾಯ
ವಾಚನ - - A Kannada text synthesizer
ತೆರೆ
Look In:
Vaachana
CLASSES
README.TXT
Vaachana.jcw
SRC
src_vaachana.txt
ಅಕ್ಷರ.txt
BHAGVATH.TXT
Vaachana.jcd
ಗುರುಗಳು
HELP.TXT
Vaachana.jcp
ಮೀಮ್ಸ
KANNADA.TXT
Vaachana.jcu
File Name:
ಗುರುಗಳು
Files of Type:
All Files
ತೆರೆ
Cancel
ತೆರೆ
ತೆರೆರಿದೆ: ಕನ್ನಡ


ವಾಚನ - A Kannada text to speech synthesizer
ಕಡತ ಸಂಪಾದನೆ ಓದು ಕಡ್ಡಿ ಸಹಾಯ
ಜೈ ಭಾರತ ಜನನಿಯ ತನುಜಾತೆ
ಜಯಹೇ ಕರ್ನಾಟಕ
ಸಿರಿಗನ್ನಡಂ ಗೆಲ್ಗೆ
ಕೌತುಕ
ಕೈಯ್ಯಾರ ಕಿಳ್ಞಣ್ಣ ರೈ
ಕೃತಜ್ಞತೆ ಕೃತಘ್ನ ಅಂ
ಶ್ರೀ ಮದ್ಭಗವದ್ಗೀತಾ
ಜಗತ್ಪ್ರಹೃಷ್ಯತ್ಯನುರಜ
ಬಗ್ಗೆ
ವಾಚನ - A Kannada Speech Synthesizer
ವಾಚನ
ಆವೃತ್ತಿ ೧.೦
ಕನ್ನಡ ಬರವಣಿಗೆಯನ್ನು ಮಾತನ್ನಾಗಿ ಪರಿವರ್ತಿಸುವ ತಂತ್ರಾಂಶ
ಕರ್ತೃಗಳು: ಕಾರ್ತಿಕ್ ಟಿ. ಕೃಷ್ಣರಾಜ್ ಬರ್ಮತ್ತಾಯ ಬಿ.
ನಂದೀಶ ಜಿ. ಎಂ. ಉತ್ತಮ್ ಎಚ್.
ಗಣಕಯಂತ್ರ ವಿಜ್ಞಾನ ಮತ್ತು ಅಭಿಯಂತರ ವಿಭಾಗ (ಅಂತಿಮ ವರ್ಷ)
ಕೆ ವಿ ಜಿ ತಾಂತ್ರಿಕ ಮಹಾವಿದ್ಯಾಲಯ, ಸುಳ್ಯ
ಬಗ್ಗೆ


ವಾಚನ - A Kannada text to speech synthesizer
ಕಡತ ಸಂಪಾದನೆ ಓದು ಕಯ್ಕೆ ಸಹಾಯ
ಜೈ ಭಾರತ ಜನನಿಯ ತನುಜಾತೆ
ಸಹಾಯ
ವಾಚನ - A Kannada Speech Synthesizer
ಜಯಹೇ ಕ
ಸಿರಿಗನ್ನಡ
ಕೌತುಕ
ಕೈಯ್ಯಾರ
ಕೃತಜ್ಞತೆ ಸ
ಶ್ರೀ ಮದ್ಭ
ಜಗತ್ಪ್ರಭ
Vaachana Help
ಕೀಲಿಮಣೆವಿನ್ಯಾಸ
ಸ್ವರಾಕ್ಷರಗಳು
D E F R G T + Z S W ~ A Q
ಅ ಆ ಇ ಈ ಉ ಊ ಋ ಎ ಏ ಐ ಒ ಓ ಔ
ವ್ಯಂಜನಗಳು
ವರ್ಗೀಯ ವ್ಯಂಜನಗಳು
k K i I U
ಕ ಖ ಗ ಘ ಙ
; : p P }
ಚ ಛ ಜ ಝ ಞ
ಸಹಾಯ

CHAPTER 6

USER MANUAL

CHAPTER 6

USER MANUAL

The package VAACHANA is designed to be very user friendly. Different tool bars, menus are provided for easy usage. Because it is basically designed for Kannada language all menus and tool tip texts are given in Kannada language. A special “Help” file is provided with the package will assist the novice user to type in Kannada. With a little practice one can become master in this Kannada keyboard layout.

To use this package i.e. to read out the Kannada text either one has to type the text in the text area or else he could open the pre-stored text from the memory. After the text is ready to read, he has to select “ಓದು” (Speak) option either from the toolbar or from the menu option.

Along with this, a slider bar is provided to change the frequency of voice to be read. Conventional text editor options like open, save, cut, copy and paste are also been provided for effective usage of the package.

CHAPTER 7
FUTURE SCOPE OF THE PROJECT

CHAPTER 7

FUTURE SCOPE OF THE PROJECT

The various future expansion possibilities for VAACHANA are,

1. Improvement of the smoothness of the sound.
2. Inclusion of prosody and the naturalization of the voices like human expressions.
3. Reading of special cases like date and number.
4. Inclusion of different kinds of voices and graphical faces.
5. Import and Export of documents.
6. Controlling the reading speed.

CHAPTER 8
CONCLUSION

CHAPTER 8

CONCLUSION

As per the goal of this project an attempt is made to show how the computer speaks out the Kannada text. Here the provision is provided to the user to input the Kannada text and he can listen to his text.

Present system just pronounces the Kannada character; however the “naturalness” of the synthetic speech needs to be improved for implementing the expressions of the human beings.

By developing such systems, relationship between human and computer becomes much closer. Thus it helps in overcoming the problem of DIGITAL DIVIDE.

BIBILIOGRAPHY

BIBILIOGRAPHY

1. Steven Holzner, Java 2 AWT, Swing, XML and Java Beans Programming Black Book, Dreamtech Press, First Edition, 2002.
2. E. Balagurusamy, Programming with Java, A Primer, Tata McGraw-Hill Publishing Co. Ltd, 2nd Edition, 2000.
3. Robert Eckstein, Marc Loy and Dave Wood, Java Swing, O'Reilly, Shroff Publishers and Distributors Pvt. Ltd, February 2001
4. DHVANI home page- http://dhvani.net.
5. CDAC home page -http://www.ncb.ernet.in/matrubhasha.