

Major Project

Nepali Speech Recognition

Presented By:

Chetan Prajapati (060/BCT/509)

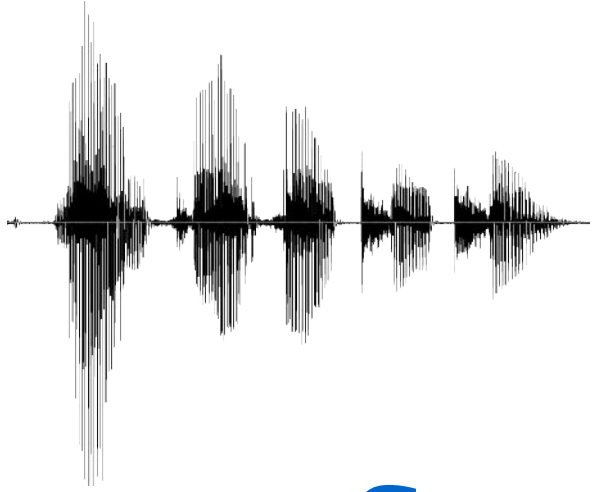
Jiwan Nyoupane (060/BCT/512)

Jwalanta Deep Shrestha (060/BCT/513)

Shishir Jha (060/BCT/541)

February, 2008

Introduction



Speech -> Digital Text

नेपाल

Research Based Project

Aimed at finding best possible method for NSR

Nepali Input System

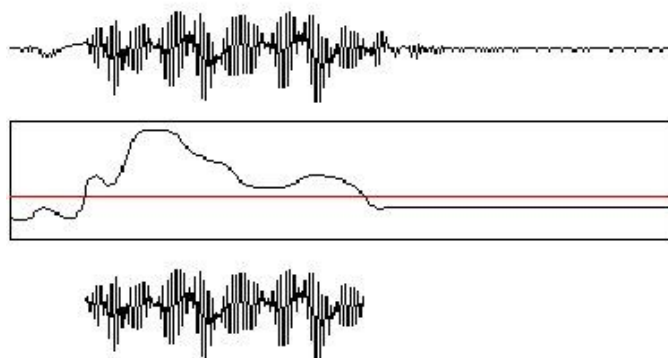
Inconvenient

Nepali Speech Recognition

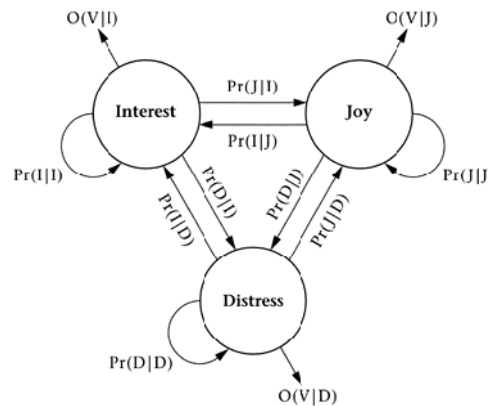
Could be an alternative

Existing Methods

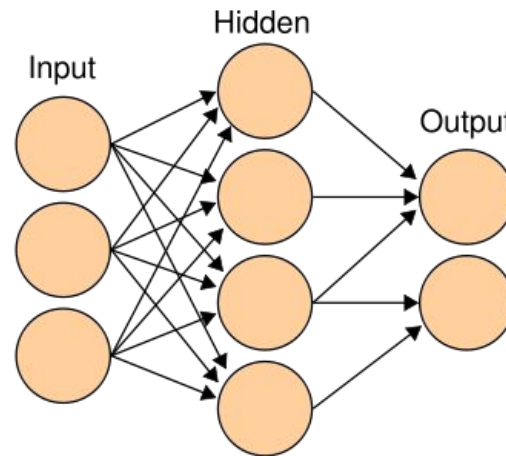
Template Matching



Hidden Markov Model (HMM)



Neural Network



Template Matching + HMM

Widely used

Implementation: Conventional Methods

Experiment:

MFCC & Direct FFT

Results

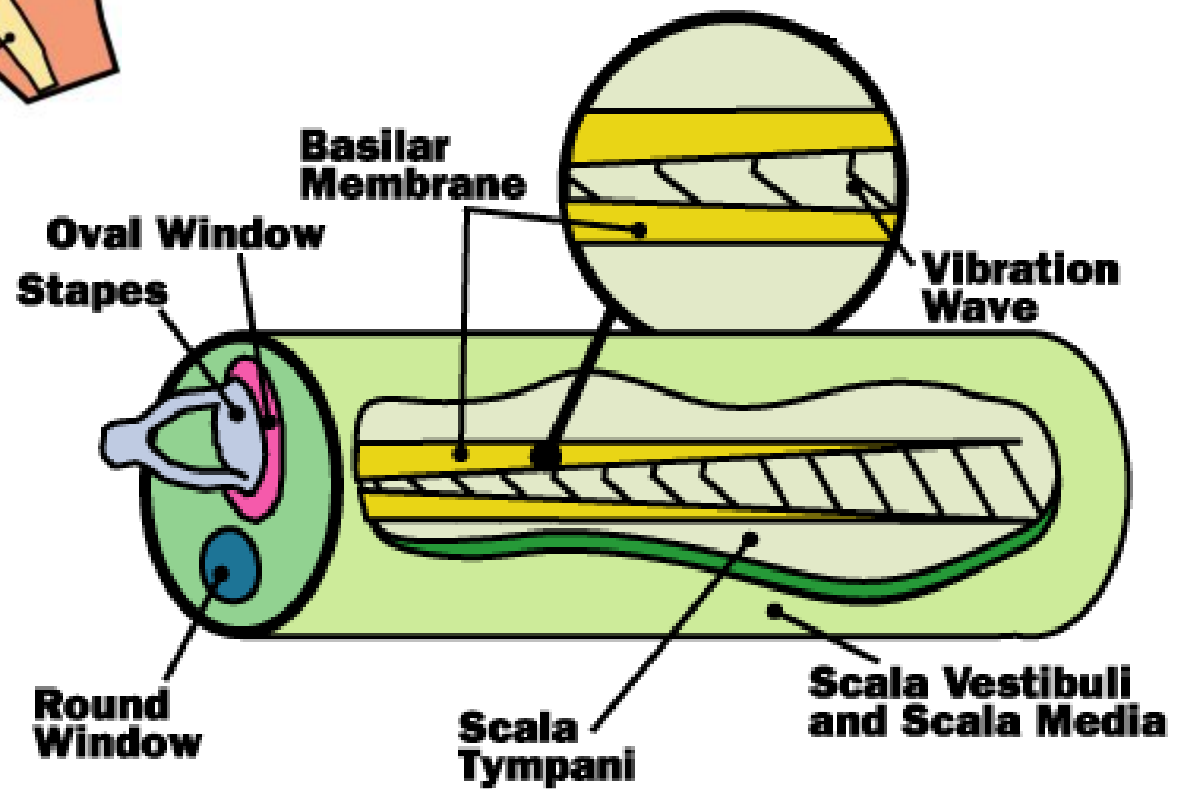
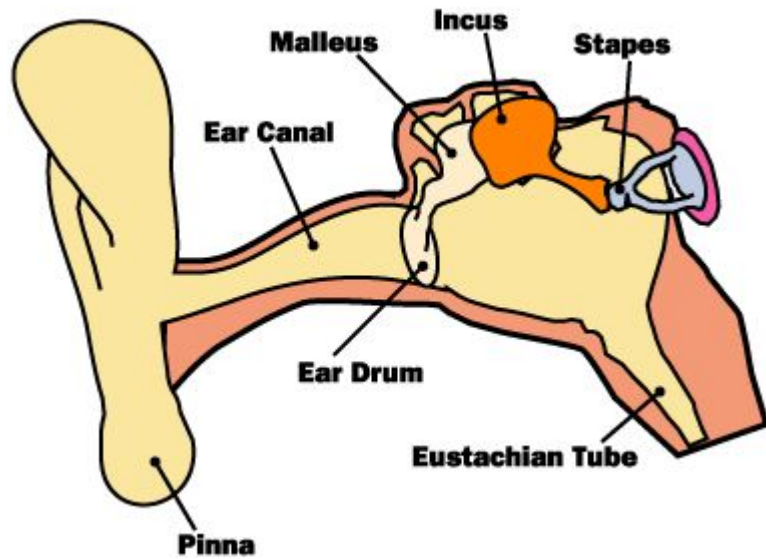
S.No.	Methods	Window Size (Samples/frame)	No. of Windows	Accuracy
1	MFCC	64	12	0.656327
2	MFCC	48	12	0.628571
3	MFCC	32	14	0.632653
4	FFT	48	12	0.673010
5	FFT	64	12	0.657595

Conclusion

Low accuracy

**Good Enough..
..but not acceptable**

Human Hearing

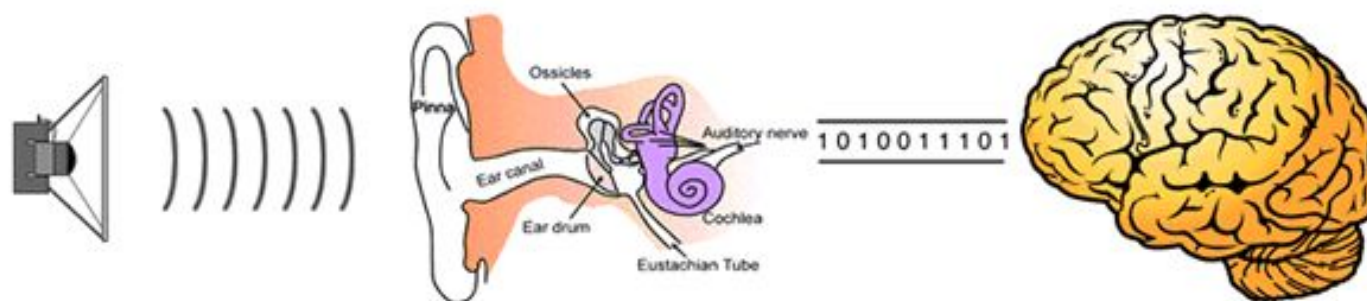


Human ear

Complex Fourier Analyzer

Ear Model

Simulation of Ear and Brain



Fibers in Cochlea

Objects

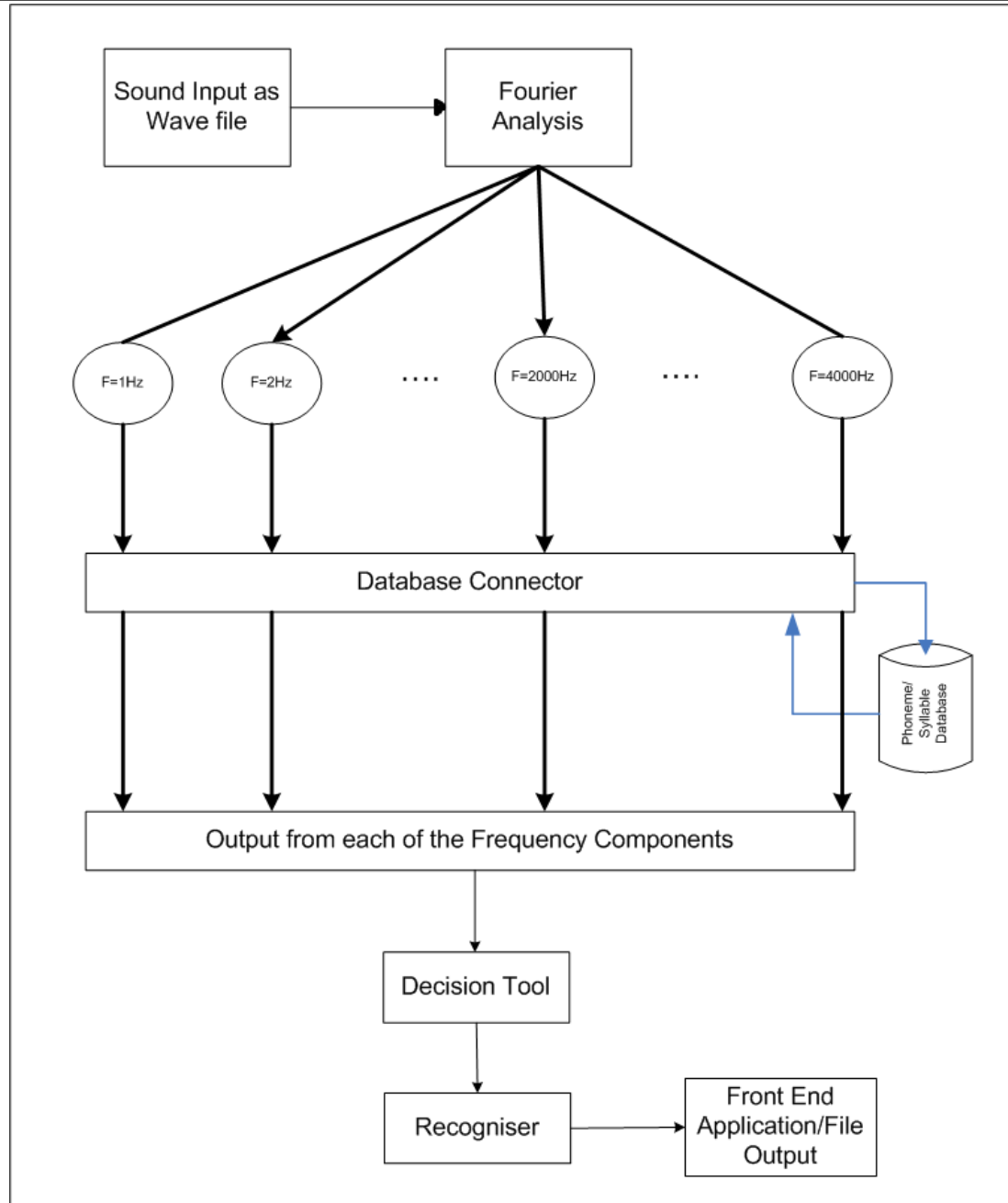
Fundamental Frequencies

Corresponds to objects

Memory

Each fiber has its own memory

Design



Hearing Mode

Listens to voice

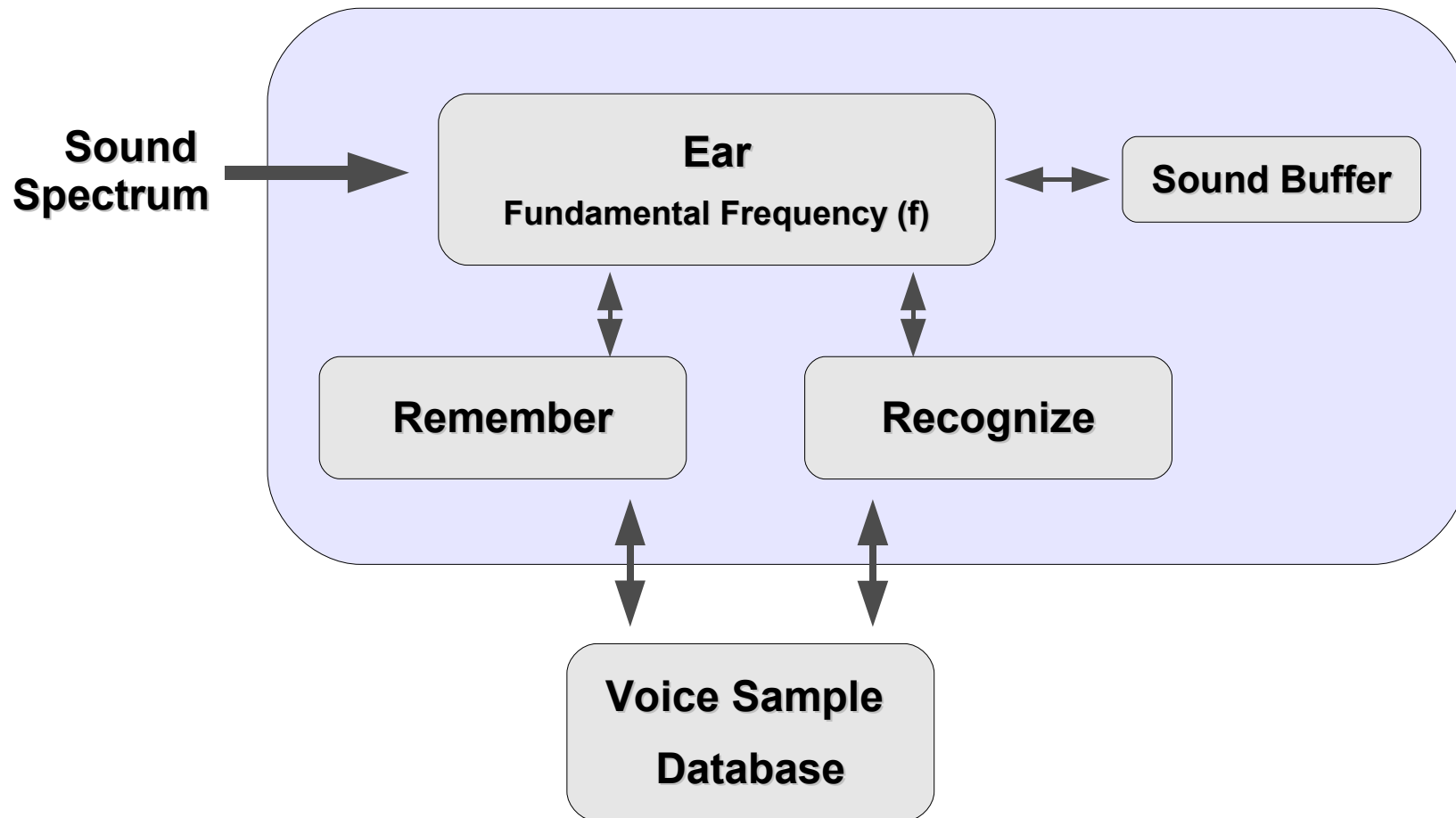
Training Mode

Stores voice samples to memory (db)

Recognition Mode

Recognizes voice sample

Ear Model



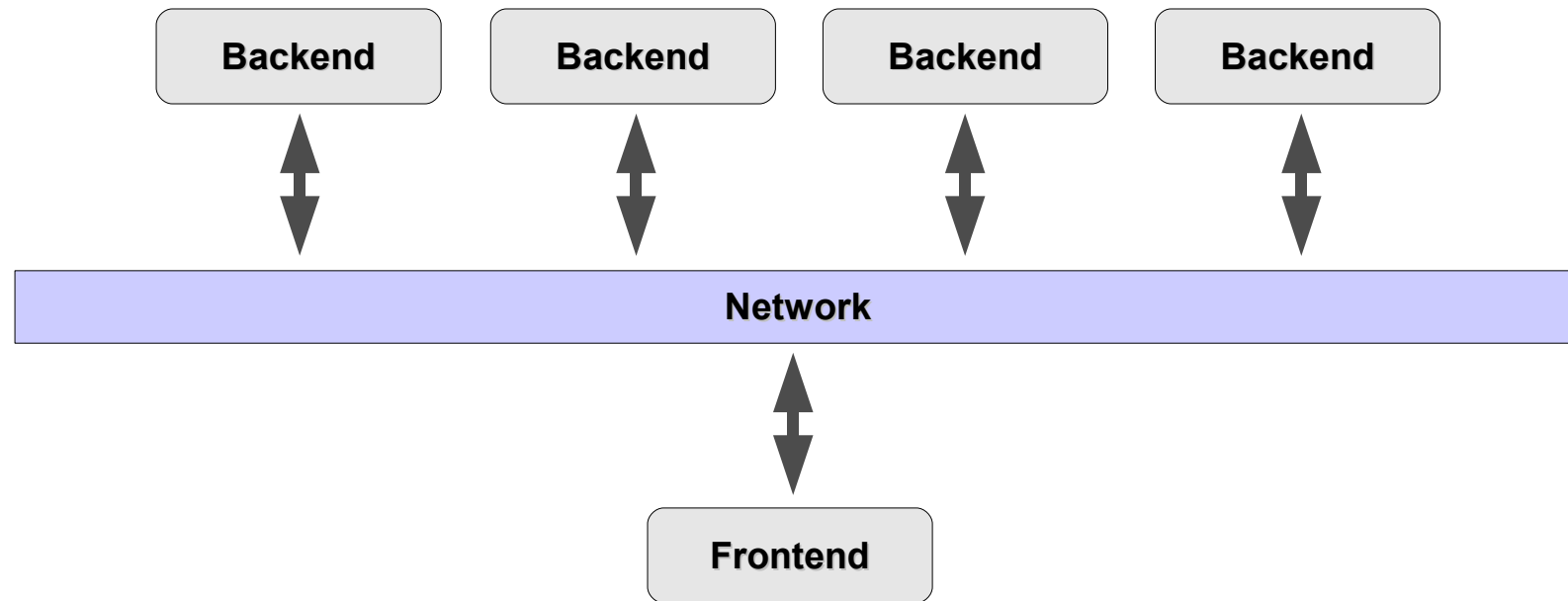
Backend

Training, Recognition

Frontend

Sound Input, Interface, Text Display

Multiple Backends



Spelling Checking

Bayes Theorem

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}.$$

Implementation

Backend

C++

Frontend

Python, wxWidget, PortAudio

Database

SQLite

Output & Accuracy

Seems Promising

Is giving better accuracy

Limitation

Narrow Domain

Currently Offline

Future Enhancements

Better Recognition

Neural Network Based Training

Performance

Optimized backend, distributed model

Better GUI

References

- 1.Ram Prasad Gyawali, et al. नेपाली व्याकरण तथा रचना**
- 2.Udhyakumar.N, Swaminathan.R and Ramakrishnan.S.K. Multilingual Speech Recognition for Information Retrieval in Indian context**
- 3.M. Kumar , N. Rajput , A. Verma. A large-vocabulary continuous speech recognition system for Hindi**
- 4.Roberto Gemello, Franco Mana, Dario Albesano. Hybrid HMM/Neural Network based Speech Recognition in Loquendo ASR**
- 5.Dr. Joseph Picone, Fundamentals of Speech Recognition**
- 6.Joseph Keshet, Shai Shalev-Shwartz, Yoram Singer, Dan Chazan. (Hebrew University) Phoneme Alignment Based on Discriminative Learning**
- 7.Prof. R. K. Joshi. The Phonemic approach for Sanskrit text**
- 8.Yun-Hsuan Sung. Speech Recognition and Synthesis**

© Original Artist
Reproduction rights obtainable from
www.CartoonStock.com



'YOU CAN HAVE THIS SPEECH RECOGNITION
SOFTWARE BACK - I SAID 'TURN OFF'
AND MY WIFE SLAPPED ME.'



Thank You