

*Home Page*

*Title Page*

*Contents*



*Page 1 of 5*

*Go Back*

*Full Screen*

*Close*

*Quit*

# IT608 Course Project

## “Profile-Based Speech Coding using Clustering on Elementary Sounds”

**Chetan Vaity (02329901)**  
**Shweta Agrawal (02329013)**  
**Amreek Singh (02329025)**

KReSIT,  
Indian Institute of Technology,  
Mumbai.

Home Page

Title Page

Contents



Page 2 of 5

Go Back

Full Screen

Close

Quit

## Motivation

- In a person's speech, similar *elementary sounds* are repeated
- Identify and store these sounds once, and use them to code speech samples
- What will the quality be like?

*Home Page*

*Title Page*

*Contents*



*Page 3 of 5*

*Go Back*

*Full Screen*

*Close*

*Quit*

## Results

- The 20ms sound slices cluster well
- The decoded sound is discontinuous but intelligible.
- Space savings: compression factor of 200 (600)



## References

- George Tzanetakis and Perry Cook, Multifeature audio segmentation for browsing and annotation, Proceedings 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, New York, Oct. 17-20, 1999
- Regine Andre-Obrecht, A new statistical approach for automatic sound segmentation of continuous speech signals, IEEE Transactions on Acoustics, Speech and Signal Processing, Vol 36, No. 1, January 1988
- Suresh Balakrishna, Speech Recognition using Mel Cepstrum features, Mississippi State University, 1998
- CLUTO, A clustering toolkit  
(<http://www-users.cs.umn.edu/karypis/cluto/>)
- cepstral, MFCC feature extraction algorithms  
(<http://tiger.la.asu.edu/software.htm>)
- ISIP Automatic Speech Recognition (<http://isip.msstate.edu/projects/speech/>)
- John Watkinson, The Art of Digital Audio, 3rd Edition

*Home Page*

*Title Page*

*Contents*



*Page 5 of 5*

*Go Back*

*Full Screen*

*Close*

*Quit*