Lip Reading to Text

Waleed Deaney

Supervisor: Prof. I.M Venter

Co-Supervisor: Mr. M Ghaziasgar

Mentor :Mr. K Abrahams

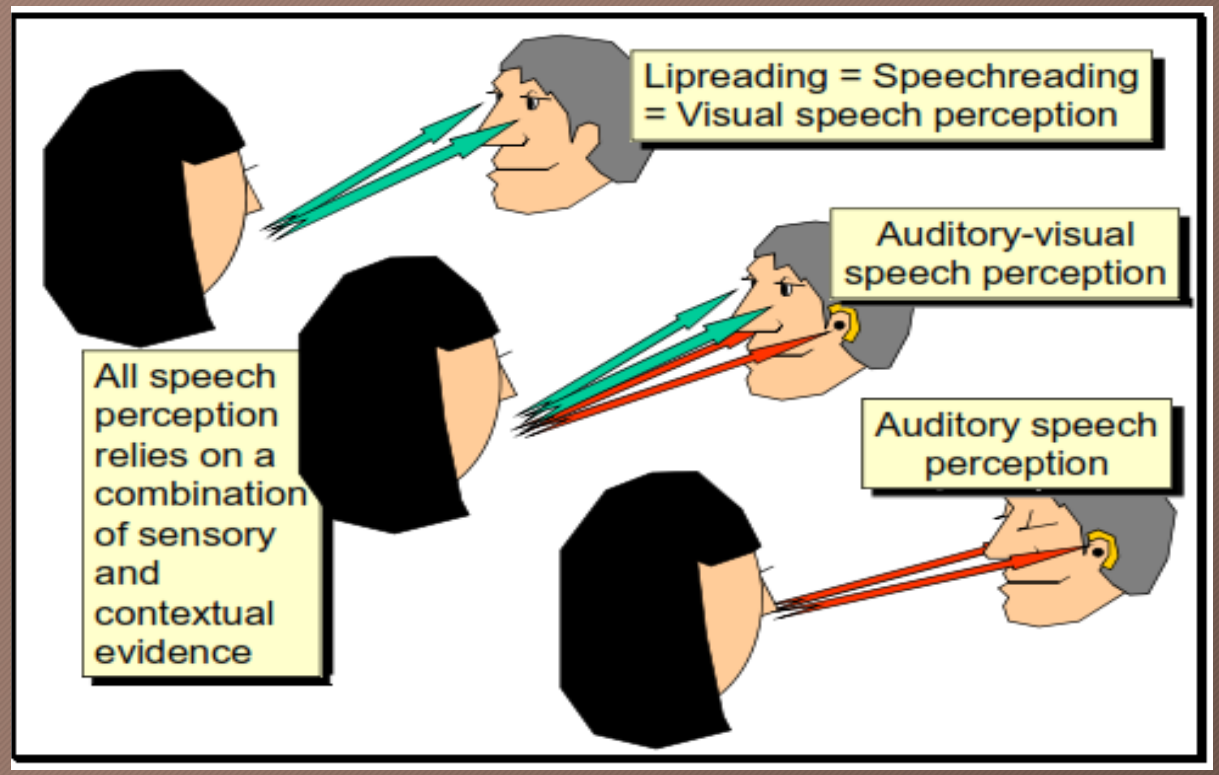
Co-mentor: Mr. N de la Cruz

Overview

- Background
- User Requirements
- Requirements Analysis Design
- Project Plan

Background

- Speech is a multi-modal form of interaction



Background (cont.)

- Lip Reading or Speech Reading is a visual way of “listening” to someone
- Speech recognition software:
  - Auditory
  - Limitations:
    - Noisy environments
    - Multiple speakers

User Requirements

- Visual Speech Recognition System
- Results should be displayed as TEXT
- Easy to use
- Start and Stop at any given time
- System should use Web Cam
  - Input Video

Requirements Analysis Design Scope

- User should clearly pronounce sounds or letters
- One user in frame at a time
- User should be facing camera directly
- Regard the mouth as the region of interest

Requirements Analysis Design(cont.)
Lip Reader to Text System

Face and Mouth Detection

Pre-Processing

Training and Testing

Project Plan

GOAL	Due Date
Research • Learn how to use OpenCV	End of Term 1
• Accurately locate mouth and extract features	End of Term 2
Implementation • Train the system to recognize a sounds or letters • Optimize image for better recognition	End of Term 3
Test and Evaluate • Add more training and testing data	End of Term4

References

- Bradski, G. and Kaehler, A. (2008). Learning OpenCV.
- McGurk, H. and MacDonald, J. (1976). Hearing lips and seeing voices.
- Mehrotra, H., Agrawal, G., and Srivastava, M. (2009). Automatic lip contour tracking and visual character recognition for computerized lip reading.

Questions?