Speech Recognition System

Jaime Díaz
Raiza Muñiz
System Overview

- Closed-Set Speaker ID system
- Two active states
  - Speaker training
  - Speaker ID
- Project partition
  - Jaime – DSP (feature extraction & comparison).
  - Raiza – control, memory & video output.
Extractor Block

- Processes ~ 3.5 sec audio
- Outputs 16 Spec. Coeff.

- Issue: number of samples
  - Need to process small chunks
  - Pipelining to reduce gates
  - Customization ⇔ less portable
Distance Block

- Compares Spec. Coef. (SC)
  - Input vs Stored (Speech)
  - Outputs a distance metric

- Comparison: Dynamic Time Warping
  - Calc. Euclidean distance bet the SC of input vs stored for each time interval.
  - Dist = \( \sum \) smallest dist in each TI row and column of the distance matrix.
Control Block

- Tells all other blocks what to do.
- Drives the direct user I/O interface
  - ADD or ID user inputs.
  - Video output
- Drives Memory Read/Write cycles
- Supplies Distance Block stored SC vectors.
Other Blocks

- Memory – store/read user SC as needed
- Register – tell Control requested action
- Video interface – feedback to the user
Thank You!

Questions?