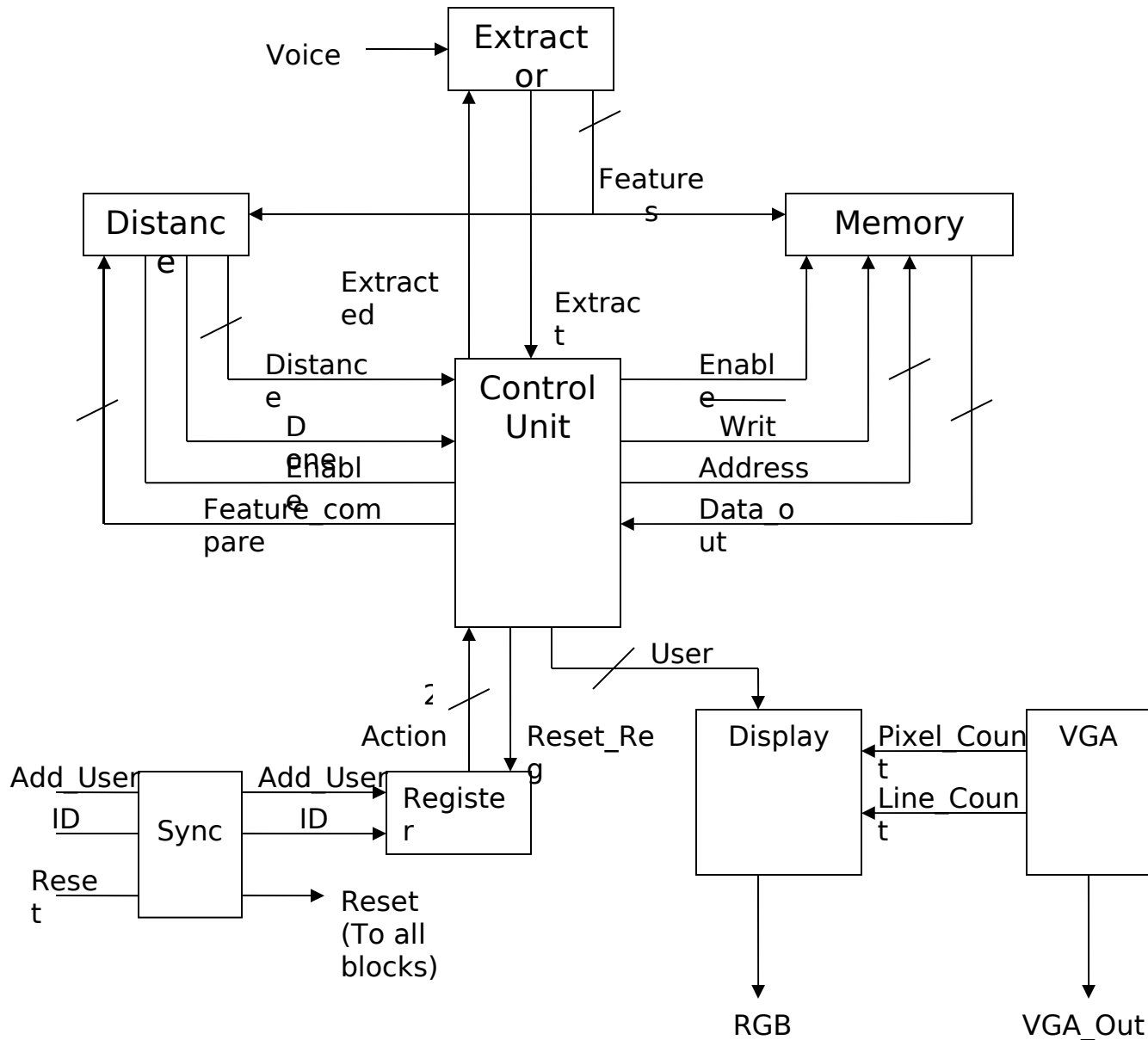# Speech Recognition System

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.

Block Diagram

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

Voice

↓

AC'97

↓

Hamming Window

↓

DFT

↓

Mel Filters

↓

Log

↓

DCT

↓

Spectral Coefficients
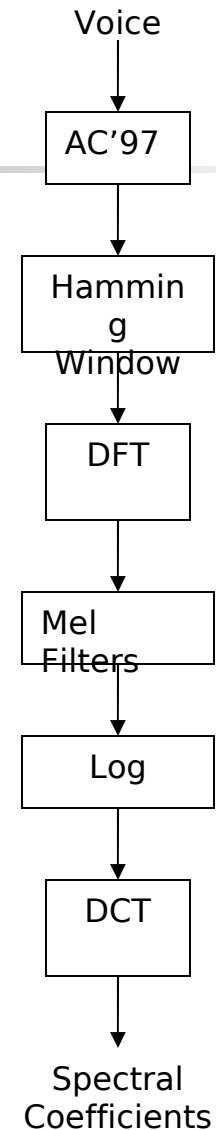
Distance Block

- 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 = Σ smallest dist in each TI row and column of the distance matrix

Distance Metric Calculation Example

-					4										
-					3										
-					4										
X					7										
X					9										
X					7										
X					8										
I					8										
I					8										
S					7										
S	6	7	5	7	7	6	5	1	2	8	8	9	8	9	6
S					9										
-					3										
-					2										
-					4										
	-	-	-	-	-	-	-	S	S	I	I	I	X	X	-

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?