Voice Recognition in Hardware
### As an Input Device

Tao Ben Schardl
Nicholas Harrington

A Functional Overview

- Small sections of incoming audio ($\approx 20ms$) are analyzed and converted into vectors that parameterize the incoming audio.
- Each word is composed of a list of these vectors. Endpoints for possible words are determined by the overall energy of incoming audio.
- Words trained into the hardware are compared against incoming audio signals using a dynamic time warping alogorithm (DTW).
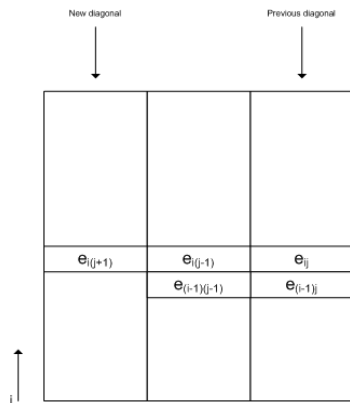
Vector Generation
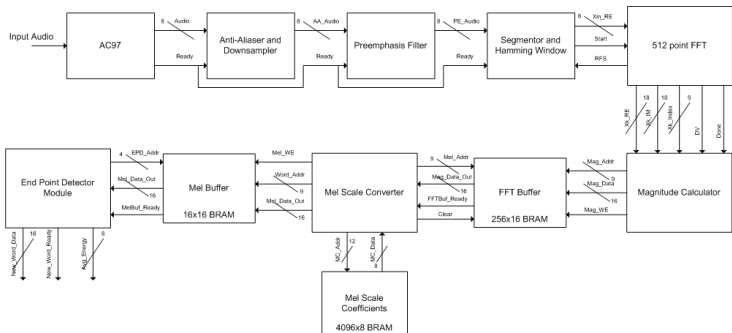
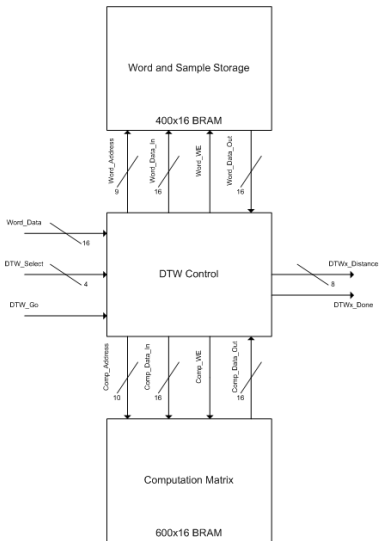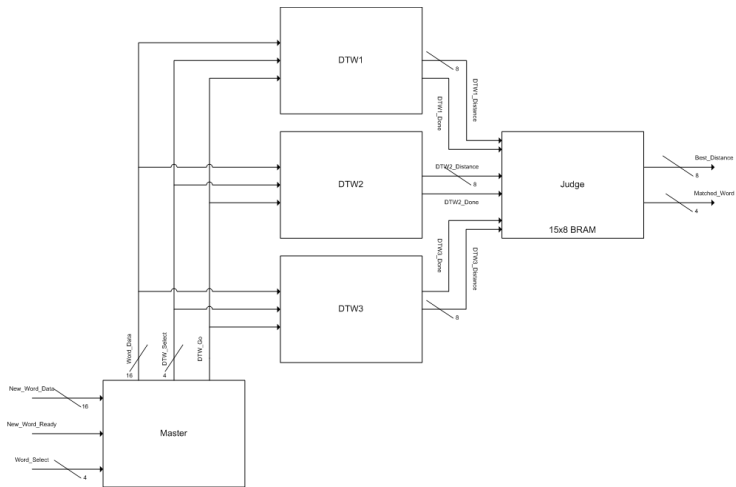# Vector Generation
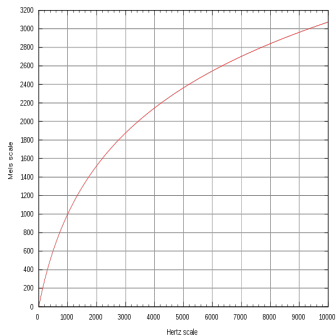
Borrowed from wikipedia

Algorithm

$$e_{i(j+1)} = d_{i(j+1)} + \min(e_{ij}, e_{i(j-1)}, e_{(i-1)(j-1)})$$
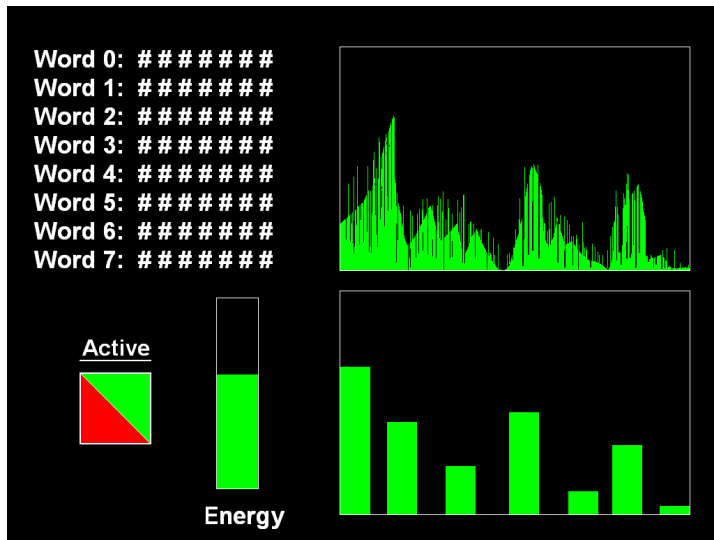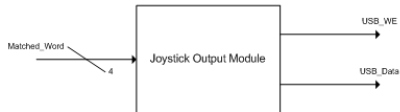$$d_{ij} = d(w_i, s_{j-i})$$

DTW

DTW Control

Mock Layout

Output Module
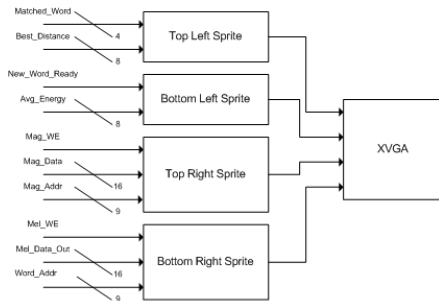
November 2008

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
9	10	11 Veteran's Day	12	13	14	15
16	17	18	19	20	21 Working FFT	22
23 Working DTW	24 Mel Scale Values	25	26 Finalized Dist. Func	27 Thanksgiving	28	29 Output Modules
30						

December 2008

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
	1 P	2 A	3 N	4 I	5 C	6 Working System
7	8	9	10	11	12	13

http://homeschooling.about.com

Thank You For Your Time