A Novel Fuzzy Approach to Speech Recognition

Ramin Halavati, Saeed B. Shouraki, Pujan Ziaie
Sharif University of Technology
Tehran, Iran

Presented by: Pujan Ziaie (pujan_nacon@Yahoo.com)

Summery

- Introduction: Speech Recognition
- Proposed Model
- Recognition Approach
- Training process
- Results
Speech Recognition

- Several Methods
  - HMM (Hidden Markov Models), TDNN (Time Delay NN), ... 

- Common Problems:
  - Effect of Noise
  - Recognition Speed

- Fuzzy approach:
  - To Ignore details such as noise.
  - similarity with human recognition process.
Human Voice Recognition

- Imprecise processing
- Deciding upon a rough measurement of amplitude
- No counting on speech frames (relative lengths)
- Sensitive to lower frequencies
Proposed Model

- **Base Data:**
  - Speech Spectrogram
  - Phonemes Specification (developed by using GA)

- **Data manipulation:**
  - Stretching Using MEL Filter Banks. (Human’s ear is more sensitive to low frequencies and less to high ones.)
  - Fuzzification to reduce amount of data. (Human do not use that much precise data.)
  - Calculating the belongness to each phoneme
Proposed Model

- Spectrogram:
Proposed Model

- After MEL-Stretching
Proposed Model

Data Reduction (Fuzzification)

In the first step, the original signal frames are divided into 25 vertical ranges and then, the values inside each range are sorted so that the more powerful ones are moved to top.

In the second step, the top 10% values of each range are chosen and averaged and the result is replaced with the all the value of that range, making all values in each vertical range similar.
Proposed Model

- Fuzzification (Contd.)
Proposed Model

- Phoneme definition necessities:
  - Colors
  - Lengths (5 MFs)
Proposed Model

- Sample Phoneme Definition:

Range 25: Black or Blue
Range 24: Black or Blue

Range 4: Red or Yellow
Range 3: Blue or Magenta
Range 2: Black or Blue or Magenta
Range 1: Black or Blue or Magenta
Recognition Method

- The existence of appropriate phoneme definitions is assumed

- Recognition
  - Compare the given sample with all phoneme definitions
  - Choose the one with highest compatibility value
Recognition Method

☐ Single Phoneme Comparison:
  ■ Comparing the color pattern of the phoneme with all frames of the given sample.
  ■ Finding the matching sequences.
  ■ Comparing the length of a matching sequence with the required length.
# Recognition Method

## Sample, Step One:

<table>
<thead>
<tr>
<th>Input</th>
<th>Pattern</th>
<th>Compatibility</th>
<th>After applying MAX</th>
</tr>
</thead>
</table>
| -     | Range 25: Black or Blue  
Range 24: Black or Blue | Range 25: 100% or 10%  
Range 24: 100% or 10% | Range 25: 100%  
Range 24: 100% |
| -     | Range 4: Green or Yellow  
Range 3: Blue or Magenta  
Range 2: Black or Blue or Magenta  
Range 1: Black or Blue or Magenta | Range 4: 0% or 20%  
Range 3: 10% or 100%  
Range 2: 10% or 90% or 0%  
Range 1: 10% or 90% or 0% | Range 4: 20%  
Range 3: 100%  
Range 2: 90%  
Range 1: 90% |

Final Result after applying MIN: 20%
Recognition Method

Sample, Step Two:

1. Output of Step 1:
   
   \[
   \begin{array}{cccccccccccc}
   85 & 79 & 75 & 65 & 55 & 45 & 55 & 98 & 78 & 78 & 77 & 76 & 54 & 82 & 83 & 88 & 99 & 98 & 78 & 77
   \end{array}
   \]

2. Assuming the 75% as a threshold, the lengths are:

   \[3\quad 5\quad 7\]

3. Selecting the max Length:

   \[
   \begin{array}{cccccccc}
   82 & 83 & 88 & 99 & 98 & 78 & 77
   \end{array}
   \]

4. Computing Best Match Value:

   \[
   \left(\frac{82 + 83 + 88 + 99 + 98 + 78 + 77}{7}\right) = 86
   \]

5. Assuming requested Average Length for the Pattern:

   Compatibility = 86 * \text{IsAverage}(7)
Training

- To get the proper phoneme’s specification (colors and length)

- Using GA for data improvement
Training Method

- Genetic Algorithm
  - Each Genome:
    - Color Definitions
    - Length Definitions
    - Phoneme Descriptions
  - Cross Over:
    - Combination of two genomes phoneme Description part
  - Mutation:
    - Randomly change a color or length definition.
    - Randomly change a phoneme description part
Training Approach: flowchart

Start

Create 100 Random Genomes and add them to the gene pool.

Is Best Genome's Fitness acceptable?

Yes

Terminate.

No

Sort Genomes Based on their Fitnesses.

Throw out the last 50% Genomes.

Randomly choose some genomes and add their cross-overs to the gene pool.

Add a mutated copy of all available genomes to the gene pool.
Experimental Results

- Comparison with HMM

<table>
<thead>
<tr>
<th></th>
<th>Fuzzy Approach</th>
<th>HMM Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; correct answers:</td>
<td>85%</td>
<td>62.28</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt; correct answers (out of 62):</td>
<td>95%</td>
<td>79.60</td>
</tr>
<tr>
<td>6&lt;sup&gt;th&lt;/sup&gt; correct answers (out of 62):</td>
<td>98%</td>
<td>86.98</td>
</tr>
</tbody>
</table>

- One of the top three guesses has been correct.
Future Works

- To encounter color transitions in the model.
- To enhance horizontal segmentations.
- To test noise immunities.
- To alter model to represent and recognize words.
Acknowledgment

☐ Special thanks to professor Hirota (TIT) for his useful advices and also giving me the opportunity to participate in the conference
Thank you

Any questions?