Automatic Mandarin Prosody Boundary Detecting Based on Tone Nucleus Features and DNN Model

Ju Lin, Yanlu Xie, Wei Zhang, Jinsong Zhang

Speech Acquisition and Intelligent Technology (SAIT) Lab, Beijing Language and Culture University

1. Overview

Problem: Automatic detection of Mandarin prosody boundary.
Method: Detecting prosody boundary based on tone nucleus features and Deep Neural Network (DNN) model.
> This method firstly calculated the boundary-related parameters by applying the tone nucleus features.
> Then, the parameters were modeled by DNN.
Conclusion: Considering interaction between adjacent tones, the method of using tone nucleus model to extract boundary-related parameters is effective.

2. Tone Nucleus Model

2.1 Tone nucleus

![Diagram of F0 contour of adjacent "Tone 4 Tone 1", with a prosodic phrase boundary between them]

1) Tone nucleus denotes the F0 target and serves as the primary acoustic cue for tone perception.
2) Articulatory transitions are the F0 variations occurring as the transitions to or from the final targets.

2.2 Tone nucleus and prosody boundary

![Diagram of F0 contour showing transition between "Tone 4" and "Tone 1"

3. DNN Model

3.1 Corpus: ASCCD

<table>
<thead>
<tr>
<th>Total</th>
<th>B0</th>
<th>B1</th>
<th>B2</th>
<th>B3</th>
<th>B4</th>
</tr>
</thead>
<tbody>
<tr>
<td>87628</td>
<td>61518</td>
<td>16334</td>
<td>8442</td>
<td>7449</td>
<td>4072</td>
</tr>
<tr>
<td>100%</td>
<td>62.9%</td>
<td>16.7%</td>
<td>8.6%</td>
<td>7.8%</td>
<td>4.2%</td>
</tr>
</tbody>
</table>

5.1 Corpus: ASCCD

5.2 Experiment results

<table>
<thead>
<tr>
<th>System</th>
<th>B0</th>
<th>B1</th>
<th>B2</th>
<th>B3</th>
<th>B4</th>
</tr>
</thead>
<tbody>
<tr>
<td>SY-CART</td>
<td>90.9</td>
<td>48.6</td>
<td>50.9</td>
<td>80.8</td>
<td>61.7</td>
</tr>
<tr>
<td>SY-DNN</td>
<td>96.1</td>
<td>21.2</td>
<td>41.4</td>
<td>75.3</td>
<td>74.6</td>
</tr>
<tr>
<td>TN-DNN</td>
<td>96.1</td>
<td>22.6</td>
<td>47.6</td>
<td>83.6</td>
<td>75.1</td>
</tr>
</tbody>
</table>

6. Conclusion

This paper proposed a method of automatic prosody boundary detecting based on acoustic features of tone nucleus and DNN model, the experimental results demonstrate that the method is effective.