A Syntax–Directed Level Building Algorithm for Large Vocabulary Handwritten Word Recognition


1Laboratoire d’Imagerie, de Vision et d’Intelligence Artificielle
École de Technologie Supérieure
Montréal, QC, Canada

2Centre for Pattern Recognition and Machine Intelligence
Montréal, QC, Canada

3Pontifícia Universidade Católica do Paraná
Curitiba, PR, Brazil

Presentation Outline

- Problem Statement
- Objective
- System Overview
- Syntax–Directed LBA (SDLBA)
- Experimental Results
- Concluding Remarks

Problem Statement

- Accuracy and speed decrease with the increase of the lexicon size
  - More similar words
  - More entries to match against the sequence of observations
- Limitations
  - Computational resources (memory and CPU)

Proposal

- Considering
  - Limitations of the baseline system to deal with large lexicons
  - The Zip Code is unavailable
- Goal
  - Overcome problems related to speed, memory usage and accuracy
  - Build a Large Vocabulary HWR System
- How ??
  - Optimizing the search space
System Overview

Characteristics

- Reading French City Names (single and compound words)
- Words segmented into graphemes
- Cursive and handprinted letters
- Lexicon driven and character HMMs

Rennes cedex
Laval Cedex 9

System Overview

Pre-Processing

Original Image
Baseline and character slant correction
Segmentation into graphemes
Feature Sequences

Character and Word Models

PARIS
P A R I S

Decision
**System Overview**

**Recognition Engine**

- Flat Lexicon X Lexical Tree
  - Words share a sequence of $N$ initial characters
  - Reduction in the search space
- SDLBA: Lexical Tree with LBA
  - Decode the lexical tree
  - Consider both uppercase and lowercase models
  - Time asynchronous search
  - Contextual Information between levels

**Advantages of the LBA**

- Facilitates pruning
- Facilitates inclusion of contextual Information

\[ \delta_t(l, j) = \max_{0 \leq i < j} \left[ \delta_t(l, i) + a_{ij}^c(l) + CI \right] \]

- Facilitates inclusion of other contextual dependent models
- Problem: Sub-optimal solution

**Flat Lexicon x Lexical Tree**

**SDLBA and Lexical Tree**
System Overview
Differences between Viterbi and SDLBA

Experimental Results
SDLBA x Baseline System

Concluding Remarks

Analysis of the Results
- Improvement of 4.5 – 6.6 times in the recognition speed*
- Reduction of 0.0 – 1.25% in the recognition rate for TOP10*
- The problem of recognition speed has been solved partially

We believe that it is still possible to speed up the recognition and improve the accuracy
- Constraining the search inside the nodes
- Pruning the lexical tree based on the partial likelihoods of the intermediate nodes: beam search
- Incorporation of contextual information into the probabilistic framework: word metrics into the LBA

*10–30,000 entries respectively

Experimental Results
SDLBA x Baseline System

<table>
<thead>
<tr>
<th>Lexicon</th>
<th>% Chars</th>
<th>% Frames</th>
<th>X Speed</th>
<th>% TOP1</th>
<th>% TOP5</th>
<th>% TOP10</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>5.28</td>
<td>3.04</td>
<td>4.51</td>
<td>0.17</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>100</td>
<td>17.6</td>
<td>12.03</td>
<td>4.62</td>
<td>0.45</td>
<td>0.17</td>
<td>0</td>
</tr>
<tr>
<td>1,000</td>
<td>30.32</td>
<td>27.21</td>
<td>5.31</td>
<td>0.89</td>
<td>0.51</td>
<td>0.51</td>
</tr>
<tr>
<td>5,000</td>
<td>40.12</td>
<td>45.34</td>
<td>5.97</td>
<td>1.66</td>
<td>0.96</td>
<td>0.71</td>
</tr>
<tr>
<td>10,000</td>
<td>44.55</td>
<td>55.93</td>
<td>6.36</td>
<td>1.64</td>
<td>1.18</td>
<td>0.98</td>
</tr>
<tr>
<td>20,000</td>
<td>49.09</td>
<td>68.87</td>
<td>6.59</td>
<td>1.81</td>
<td>1.17</td>
<td>1.03</td>
</tr>
<tr>
<td>30,000</td>
<td>51.77</td>
<td>77.76</td>
<td>6.64</td>
<td>2.39</td>
<td>1.37</td>
<td>1.25</td>
</tr>
</tbody>
</table>

Concluding Remarks

Analysis of the Results
- Improvement of 4.5 – 6.6 times in the recognition speed*
- Reduction of 0.0 – 1.25% in the recognition rate for TOP10*
- The problem of recognition speed has been solved partially

We believe that it is still possible to speed up the recognition and improve the accuracy
- Constraining the search inside the nodes
- Pruning the lexical tree based on the partial likelihoods of the intermediate nodes: beam search
- Incorporation of contextual information into the probabilistic framework: word metrics into the LBA

*10–30,000 entries respectively
Future Work

Time–Length Constraints and Node Pruning

\[
\begin{array}{c}
\text{Level } l = 0 \\
\text{Level } l = 1 \\
\text{Level } l = \ldots \\
\text{Level } l = L-1
\end{array}
\]

\[
\begin{array}{c}
\text{FLFT}(l) \\
\text{FLFT}(l) \\
\text{FLFT}(l) \\
\text{FLFT}(l)
\end{array}
\]

\[
\begin{array}{c}
P_n > P_{th} \quad \ldots \\
\text{Level 3} \\
\text{Level 4} \\
\text{Level L-1}
\end{array}
\]

\[
\begin{array}{c}
C \quad O \quad U \quad R \quad T \\
V \quad \times \quad \times \quad \times \\
G \quad \ldots \\
T \quad \times \\
L \quad \ldots
\end{array}
\]