Text Detection in Urban Scenes

Sergio Escalera, Xavier Baró, Jordi Vitrià, Petia Radeva

2009
Text detection

- Applications:
  - Security: plate control
  - Social systems: Navigation and translation
  - Etc.

- Difficulties:
  - Changes in appearance!
Layout

- Text detection
- Text features
- Learning text
- Results
- Conclusions
Text detection – Gradient features

Features for generalizing text structure???

| \nabla T | = \sqrt{\frac{\partial T}{\partial x}^2 - \frac{\partial T}{\partial y}^2} 

\nabla T = \left( \frac{\partial T}{\partial x}, \frac{\partial T}{\partial y} \right)
Text detection – Celcius transform

\[
\begin{array}{ccc}
34 & 62 & 98 \\
34 & 62 & 98 \\
34 & 34 & 98 \\
\end{array}
\Rightarrow
\begin{array}{ccc}
1 & 1 & 0 \\
1 & 0 & \Rightarrow (11010110)_2 \\
1 & 1 & 0 \\
\end{array}
\Rightarrow CT = 214

\[
T
\]

\[
\frac{\partial T}{\partial x}
\]

\[
\frac{\partial T}{\partial y}
\]

\[
|\nabla T|
\]

ICDAR 2003 conference
### Text detection – Set of features

<table>
<thead>
<tr>
<th>Region</th>
<th>Feature Expression</th>
<th>$r_1$</th>
<th>$r_2$</th>
<th>$r_3$</th>
<th>$r_4$</th>
<th>$r_5$</th>
<th>$r_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_1$</td>
<td>$\sum_k r_{2,k}^1 - \sum_k r_{1,k}^1$</td>
<td>$r_1^1$</td>
<td>$r_2^1$</td>
<td>$r_3^1$</td>
<td>$r_4^1$</td>
<td>$r_5^1$</td>
<td>$r_6^1$</td>
</tr>
<tr>
<td></td>
<td>$\sum_k r_{2,k}^1 - (\sum_k r_{1,k}^1 + \sum_k r_{3,k}^1)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R_2$</td>
<td>$\sum_k r_{2,k}^2 - (\sum_k r_{1,k}^2 + \sum_k r_{3,k}^2)$</td>
<td>$r_1^2$</td>
<td>$r_2^2$</td>
<td>$r_3^2$</td>
<td>$r_4^2$</td>
<td>$r_5^2$</td>
<td>$r_6^2$</td>
</tr>
<tr>
<td></td>
<td>$\sum_k r_{2,k}^2 - (\sum_k r_{1,k}^2 + \sum_k r_{3,k}^2 + \sum_k r_{5,k}^2)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R_3$</td>
<td>$\sum_k r_{2,k}^3 - \sum_k r_{1,k}^3$</td>
<td>$r_1^3$</td>
<td>$r_2^3$</td>
<td>$r_3^3$</td>
<td>$r_4^3$</td>
<td>$r_5^3$</td>
<td>$r_6^3$</td>
</tr>
<tr>
<td></td>
<td>$\sum_k r_{2,k}^3 - (\sum_k r_{1,k}^3 + \sum_k r_{3,k}^3)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R_4$</td>
<td>$\sum_k r_{2,k}^4 - \sum_k r_{1,k}^4$</td>
<td>$r_1^4$</td>
<td>$r_2^4$</td>
<td>$r_3^4$</td>
<td>$r_4^4$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R_5$</td>
<td>$\sum_k r_{2,k}^5 - \sum_k r_{1,k}^5$</td>
<td>$r_1^5$</td>
<td>$r_2^5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$11 \times 4 = 44$ features per text region
Classifier Cascade - learning

- Negative Samples
- Positive Samples
- True Negatives
- False Positives

1 2 ••
Classifier Cascade - testing
### Text detection – Results

- **Data**
  - Mobile mapping system
  - Cover data set

- **Methods**
  - Gradient and CT features
  - Cascade of classifiers
  - Gentle Adaboost 50 decision stumps

- **Measurements**

\[
A = \frac{\# \text{Hits}}{\# \text{Text regions in the analyzed frames}}
\]

\[
FR = \min \left( \frac{\# \text{False positive detections}}{\# \text{Analyzed frames}}, 1 \right)
\]

\[
Performance = \frac{\text{Detected text area} \cap \text{ground truth text area}}{\text{Detected text area} \cup \text{ground truth text area}}
\]
Text detection – Results

- Mobile Mapping – Highway sequences
- 2000 frames (20 km)
- A=0.89
- FR=0.02
Text detection – Results

- Mobile Mapping – Highway sequences
- 2000 frames (20 km)
- $A=0.89$
- $FR=0.02$
- Mobile Mapping – Urban sequences
- 2000 frames (8 km)
- A=0.78
- FR=0.29
Text detection – Results

- Mobile Mapping – Urban sequences
Text detection – Cover data set

- Mobile Mapping – Cover data set

\[
Performance = \frac{\text{Detected text area} \cap \text{ground truth text area}}{\text{Detected text area} \cup \text{ground truth text area}}
\]

1000 cover images tested: \( \text{Performance} = 0.63 \)
Text translation application in mobile platforms

ACCURATE TEXT DETECTION IN CLUTTER SCENES
Conclusions

- Simple text features that allow high generalization of text structure
- Fast and accurate text detection over real data using a cascade of classifiers
- High performance over real data from a Mobile Mapping system and a new Cover data set ([http://bcnpcl.uab.cat/](http://bcnpcl.uab.cat/))

Future work:
- Analyze FP and FN and include proper complex features
  - Final levels of the cascade
  - Weighted by complexity
Thank you!!