UNIVERSITY OF BARCELONA

Centre de Visió per Computador

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

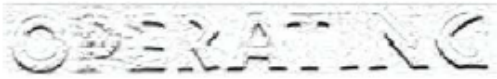
Text detection – Gradient features

Features for generalizing text structure????

ICDAR 2003 conference

T

$\frac{\partial T}{\partial x}$



$\frac{\partial T}{\partial y}$

$|\nabla T|$

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

$$\nabla T = \left(\frac{\partial T}{\partial x}, \frac{\partial T}{\partial y}\right)$$

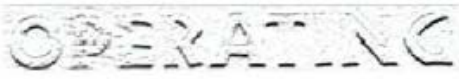
Text detection – Celcius transform

34	62	98
34	62	98
34	34	98

⇨ 1 1 0 / 1 0 / 1 1 0 ⇨ $(11010110)_2$ ⇨ CT = 214

T

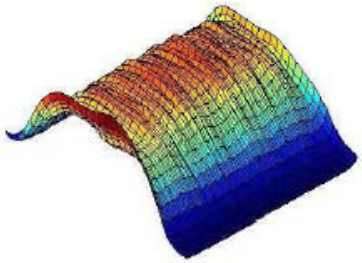
$\frac{\partial T}{\partial x}$



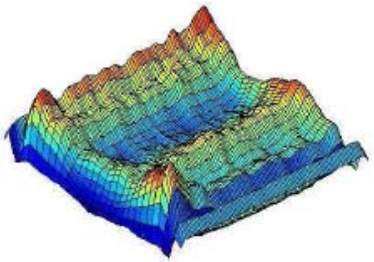
$\frac{\partial T}{\partial y}$

$|\nabla T|$

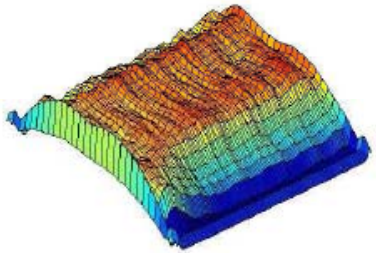
CT



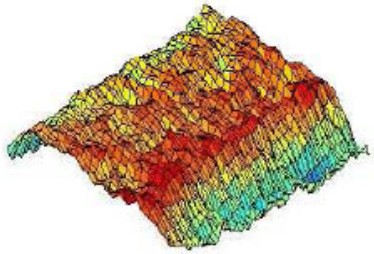
$|\frac{\partial T}{\partial x}|$



$|\frac{\partial T}{\partial y}|$



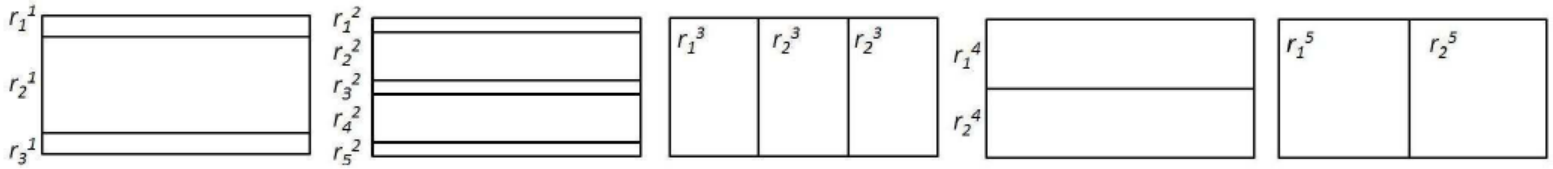
$|\nabla T|$



CT

ICDAR 2003 conference

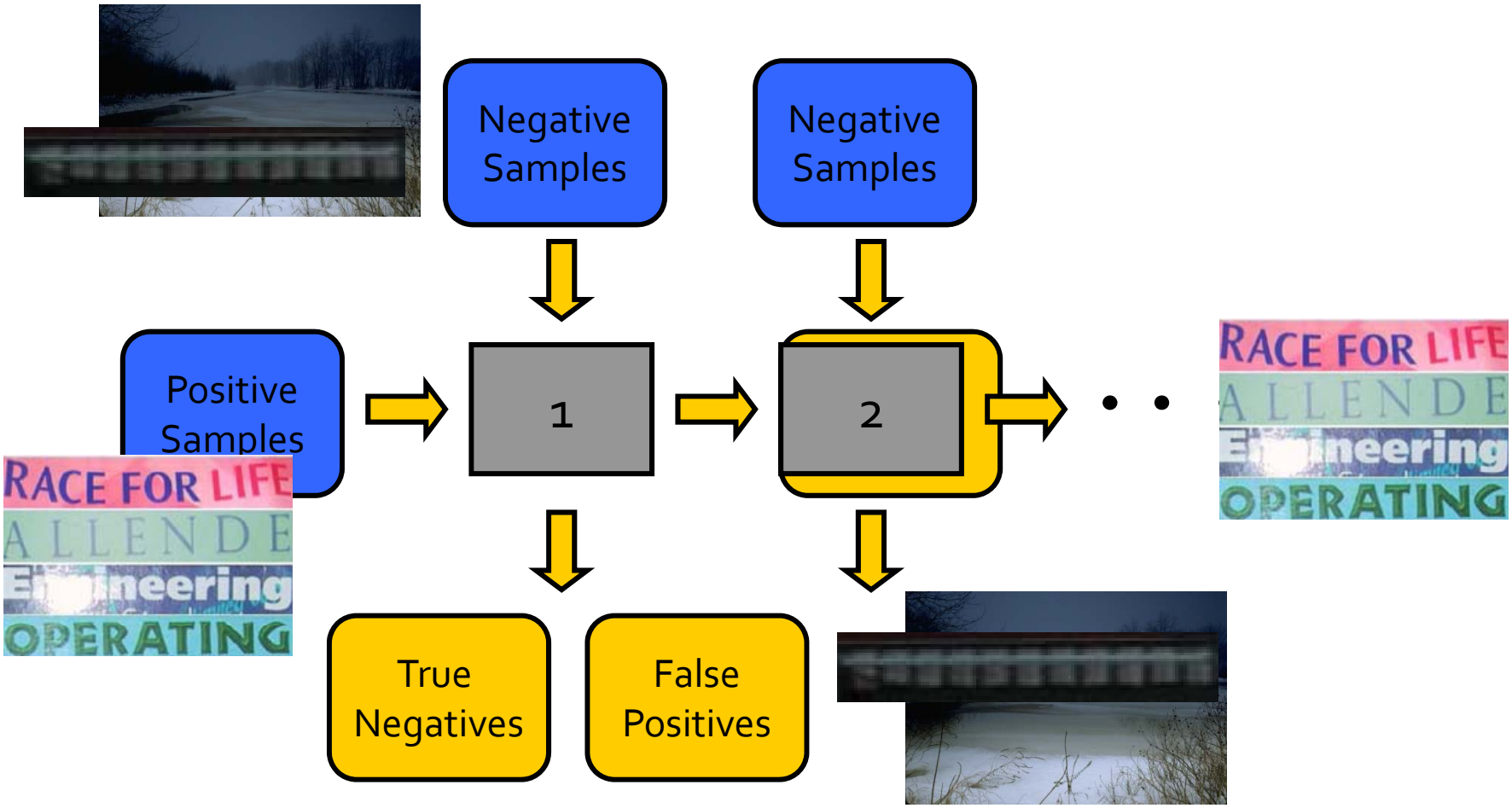
Text detection – Set of features



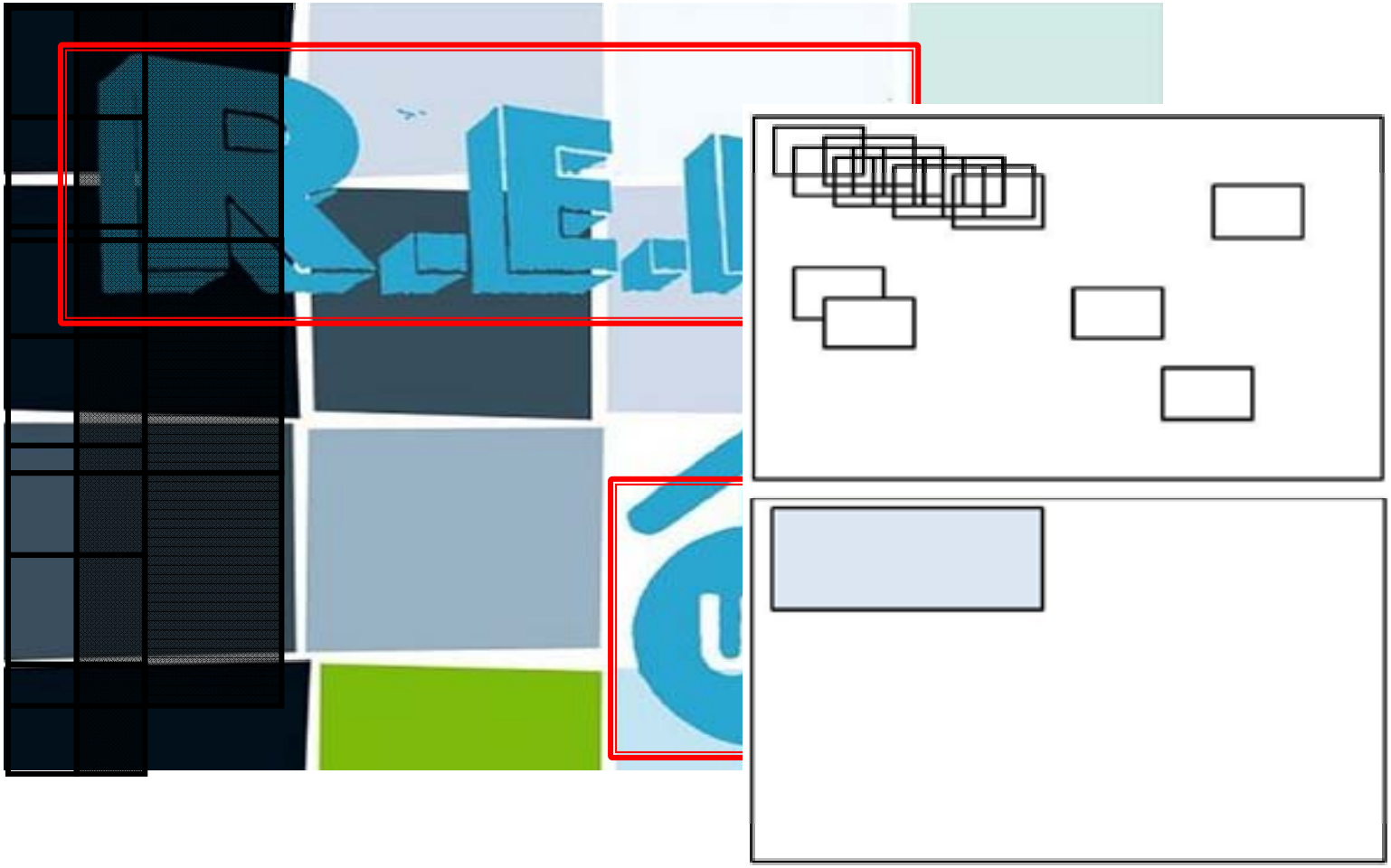
R_1	$\sum_k r^1_{2,k} - \sum_k r^1_{1,k}$ $\sum_k r^1_{2,k} - \sum_k r^1_{3,k}$ $\sum_k r^1_{2,k} - (\sum_k r^1_{1,k} + \sum_k r^1_{3,k})$	$r^1_1 : 1/6h \times 2h$ $r^1_2 : 2/3h \times 2h$ $r^1_3 : 1/6h \times 2h$
R_2	$\sum_k r^2_{2,k} - (\sum_k r^2_{1,k} + \sum_k r^2_{3,k})$ $\sum_k r^2_{4,k} - (\sum_k r^2_{3,k} + \sum_k r^2_{5,k})$ $\sum_k r^2_{2,k} + \sum_k r^2_{4,k} - (\sum_k r^2_{1,k} + \sum_k r^2_{3,k} + \sum_k r^2_{5,k})$	$r^2_1 : 1/9h \times 2h$ $r^2_2 : 1/3h \times 2h$ $r^2_3 : 1/9h \times 2h$ $r^2_4 : 1/3h \times 2h$ $r^2_5 : 1/9h \times 2h$
R_3	$\sum_k r^3_{2,k} - \sum_k r^3_{1,k}$ $\sum_k r^3_{2,k} - \sum_k r^3_{3,k}$ $\sum_k r^3_{2,k} - (\sum_k r^3_{1,k} + \sum_k r^3_{3,k})$	$r^3_1 : h \times 2/3h$ $r^3_2 : h \times 2/3h$ $r^3_3 : h \times 2/3h$
R_4	$\sum_k r^4_{2,k} - \sum_k r^4_{1,k}$	$r^4_1 : h/2 \times 2h$ $r^4_2 : h/2 \times 2h$
R_5	$\sum_k r^5_{2,k} - \sum_k r^5_{1,k}$	$r^5_1 : h \times h$ $r^5_2 : h \times h$

11 x 4 = 44 features per text region

Classifier Cascade - learning



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**

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

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

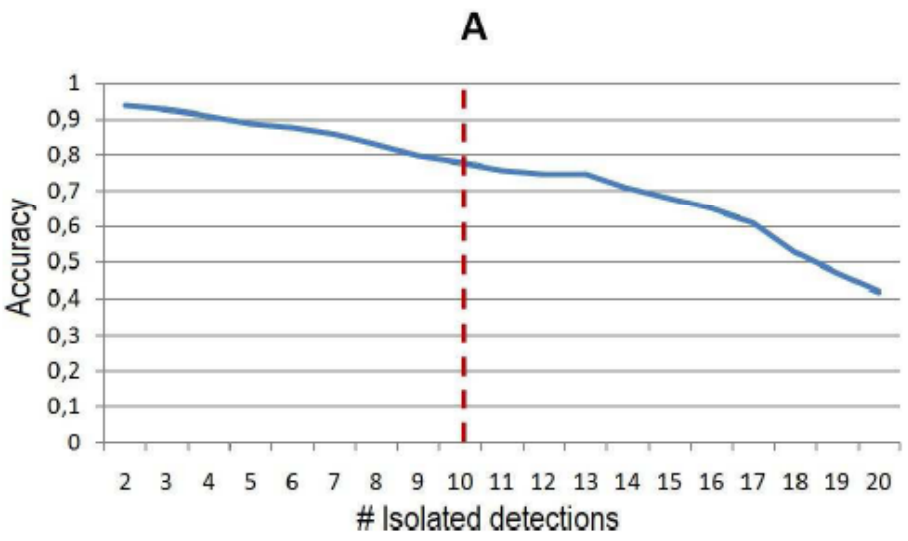
$$Performance = \frac{\text{Detected text area} \cap \text{ground truth text area}}{\text{Detected text area} \cup \text{ground truth text area}}$$

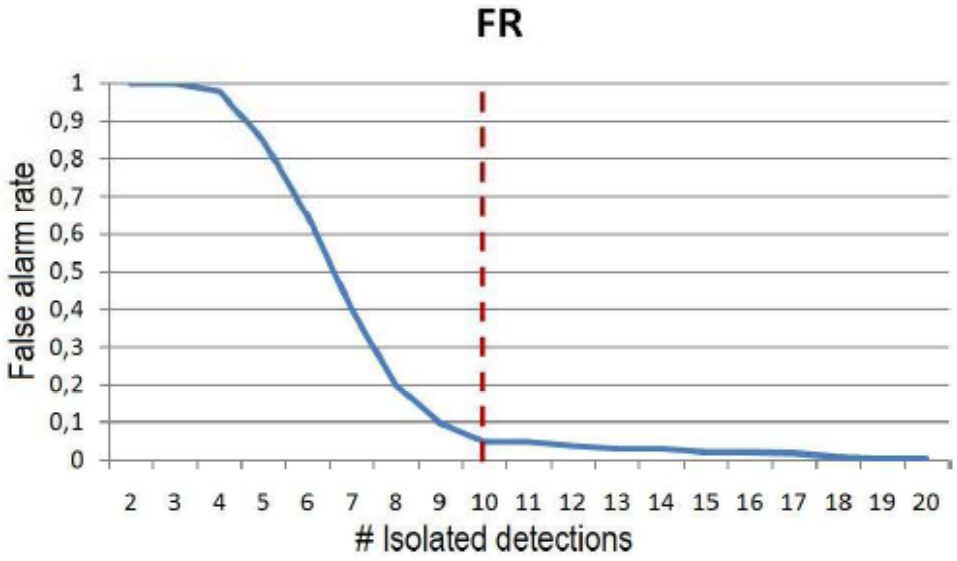
Text detection – Results

- 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



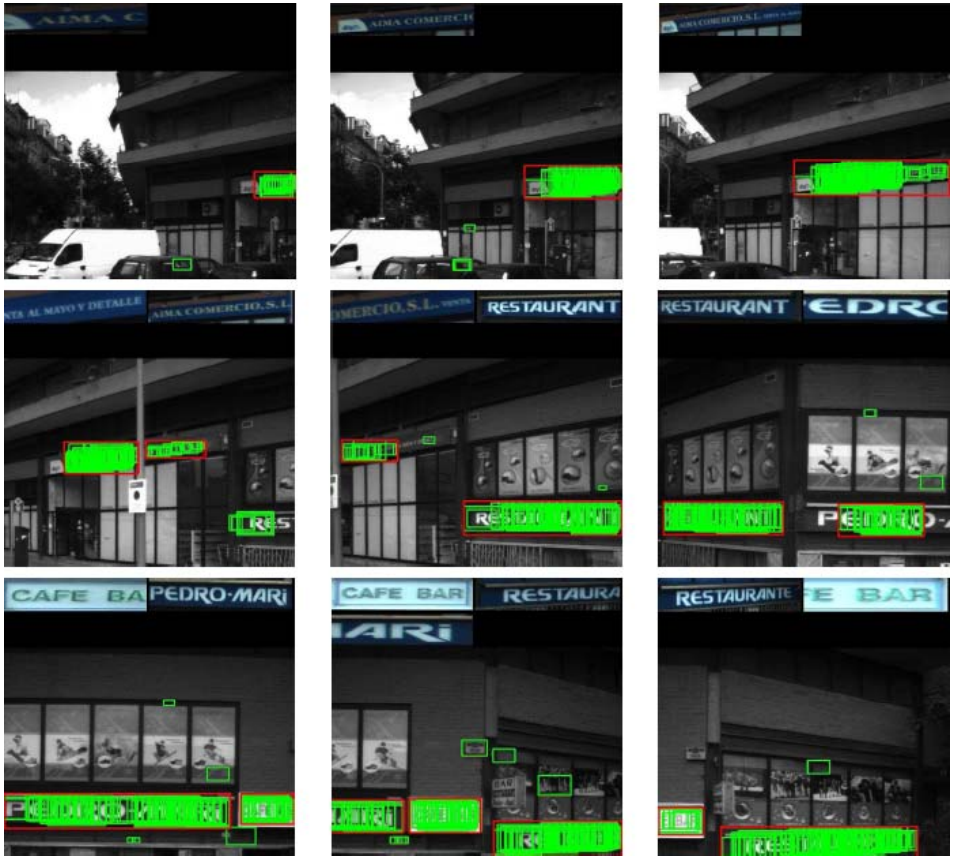


Text detection – Results

- 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: *Performance = 0.63*

Text translation application in mobile platforms

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/)**
- Future work:
  - Analyze FP and FN and include proper complex features
    - Final levels of the cascade
    - Weighted by complexity

UNIVERSITY OF BARCELONA

Centre de Visió per Computador

Thank you!!

Sergio Escalera,
Xavier Baró,
Jordi Vitrià,
Petia Radeva

2009