

# Typographical Features for Scene Text Recognition

Jerod Weinman



Department of Computer Science  
Grinnell College

# Scene Text: What Challenges?

---

- Small sample ( < 50 chars)
- Out-of-plane distortion
- Uneven lighting
- Unusual fonts
- Word segmentation
- Character segmentation



Via Via



DELANO'S



KELLOGG



REVOLVE



MARTIAL ARTS



CRAFTS



Résumés



Résumés



Résumés



Résumés



Résumés



Résumés



Résumés



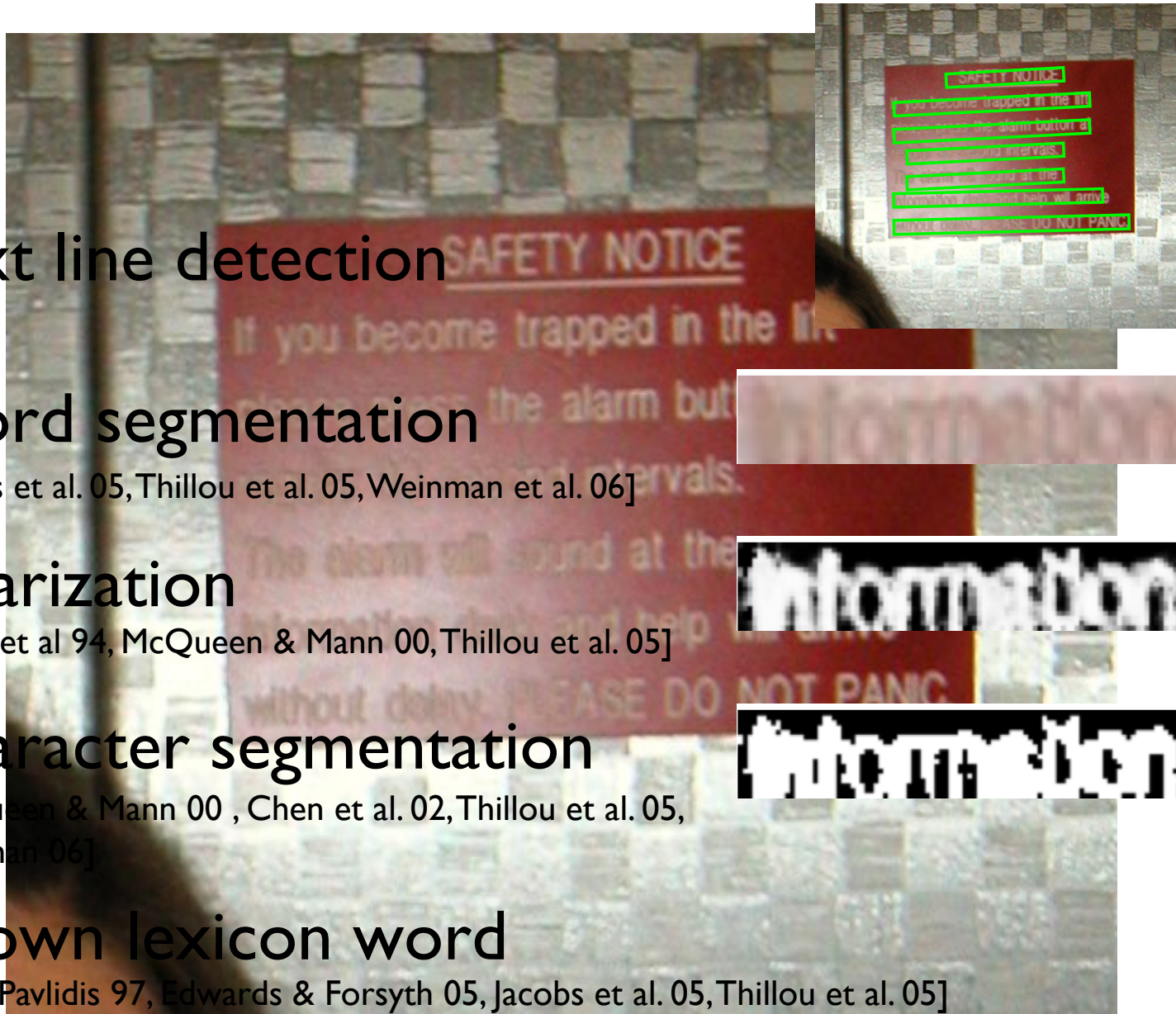
Résumés



Résumés

# What Assumptions?

- Text line detection
- Word segmentation  
[Jacobs et al. 05, Thillou et al. 05, Weinman et al. 06]
- Binarization  
[Ohya et al 94, McQueen & Mann 00, Thillou et al. 05]
- Character segmentation  
[McQueen & Mann 00, Chen et al. 02, Thillou et al. 05, Weinman 06]
- Known lexicon word  
[Shi & Pavlidis 97, Edwards & Forsyth 05, Jacobs et al. 05, Thillou et al. 05]



# Overview

---

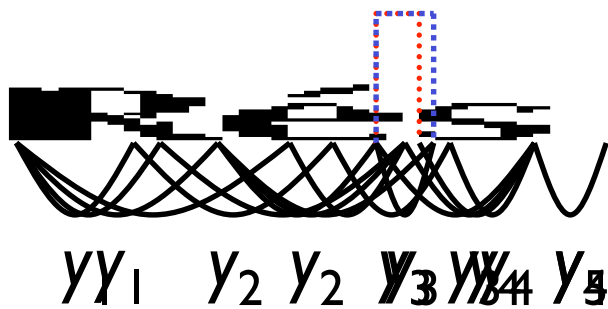
- Weinman et al. in *ICPR 2008*
  - Joint character segmentation and recognition
  - Open vocabulary; lexical bias
- Semi-Markov model
- Typographical regularities
- Experimental Results
- Conclusions

# Semi-Markov Model

$$p(\mathbf{y}|\mathbf{x};\vec{\theta}) = \frac{1}{Z(\mathbf{x})} \exp\{U(\mathbf{y}, \mathbf{x}; \vec{\theta})\}$$

[Lafferty et al. 01]

[Sarawagi & Cohen 05]



- Appearance

$$U^A(y_i, \mathbf{x})$$

- Bigrams

$$U^B(y_i, y_{i+1})$$

- Lexicon

$$U^L$$

- Parse

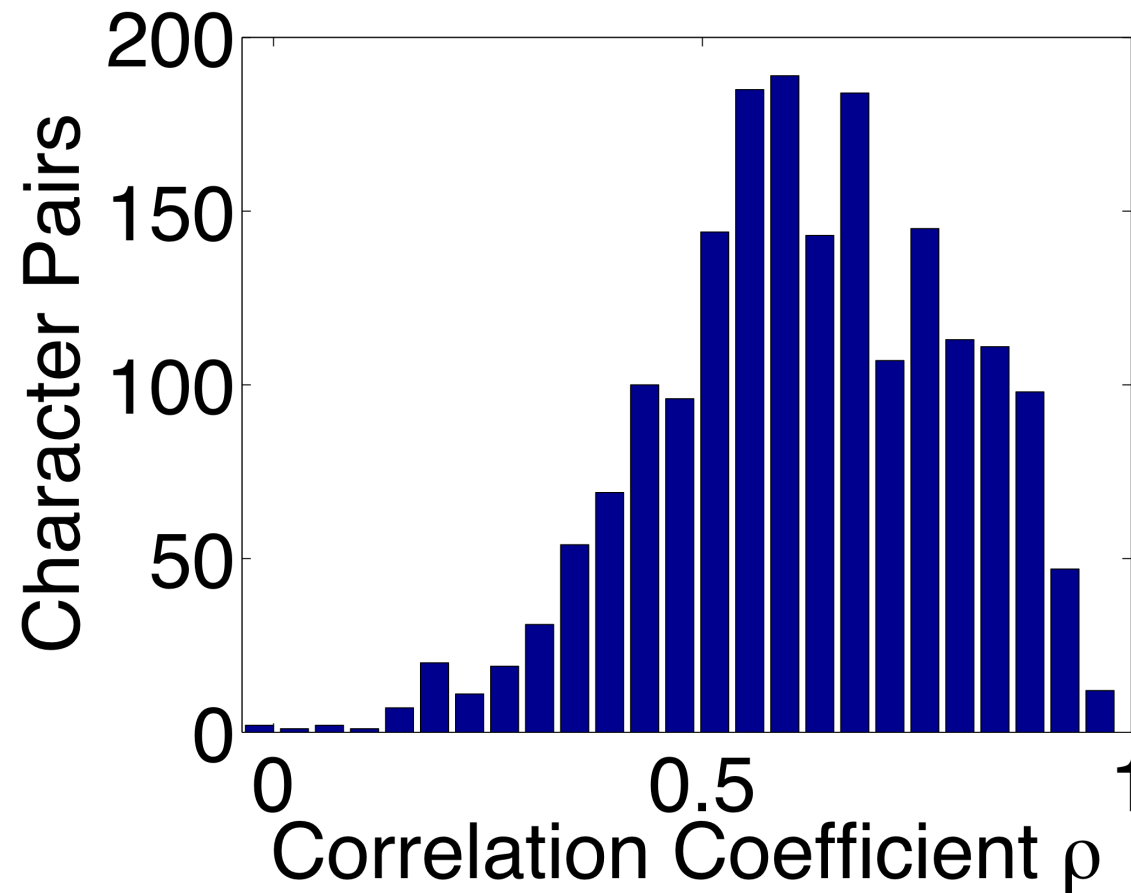
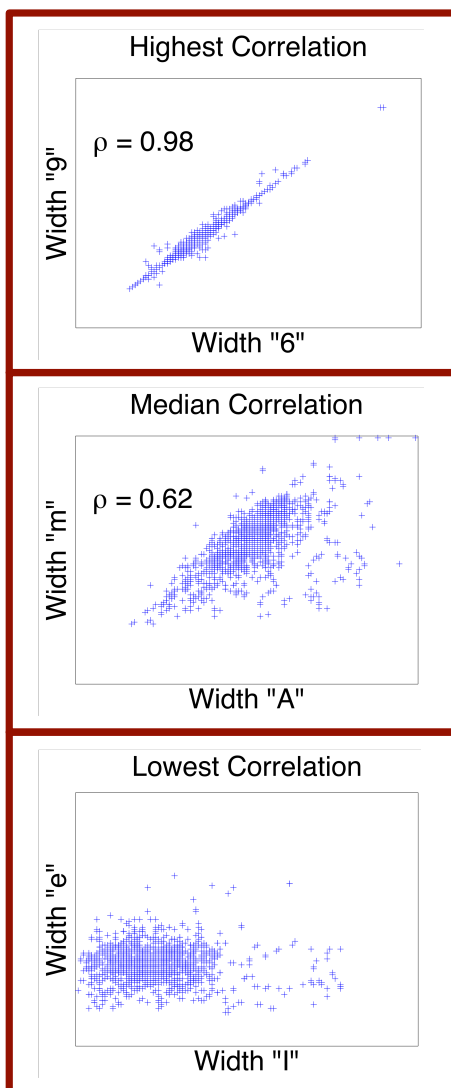
$$U^P(y_i, y_{i+1}, \mathbf{x})$$

# Typography Observations

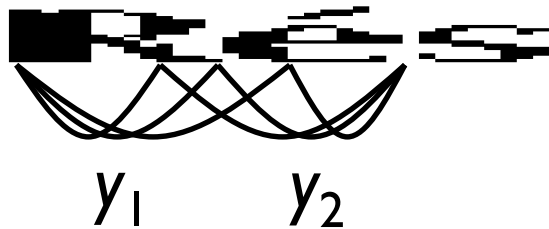
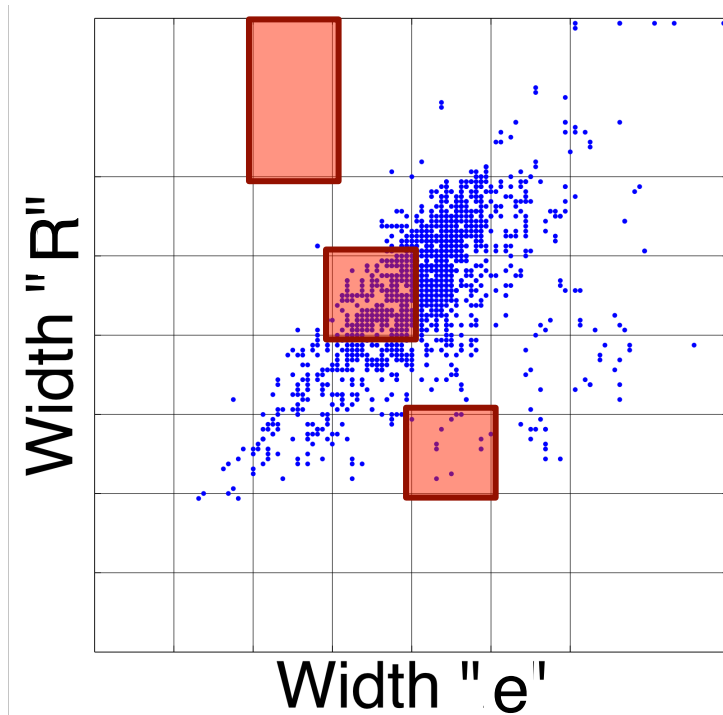
N. SHARPE

Free checking.

Free checking.



# Character-Conditional Width Bigram



- Appearance

$$U^A(y_i, \mathbf{x})$$

- Bigrams

$$U^B(y_i, y_{i+1})$$

- Lexicon

$$U^L$$

- Parse

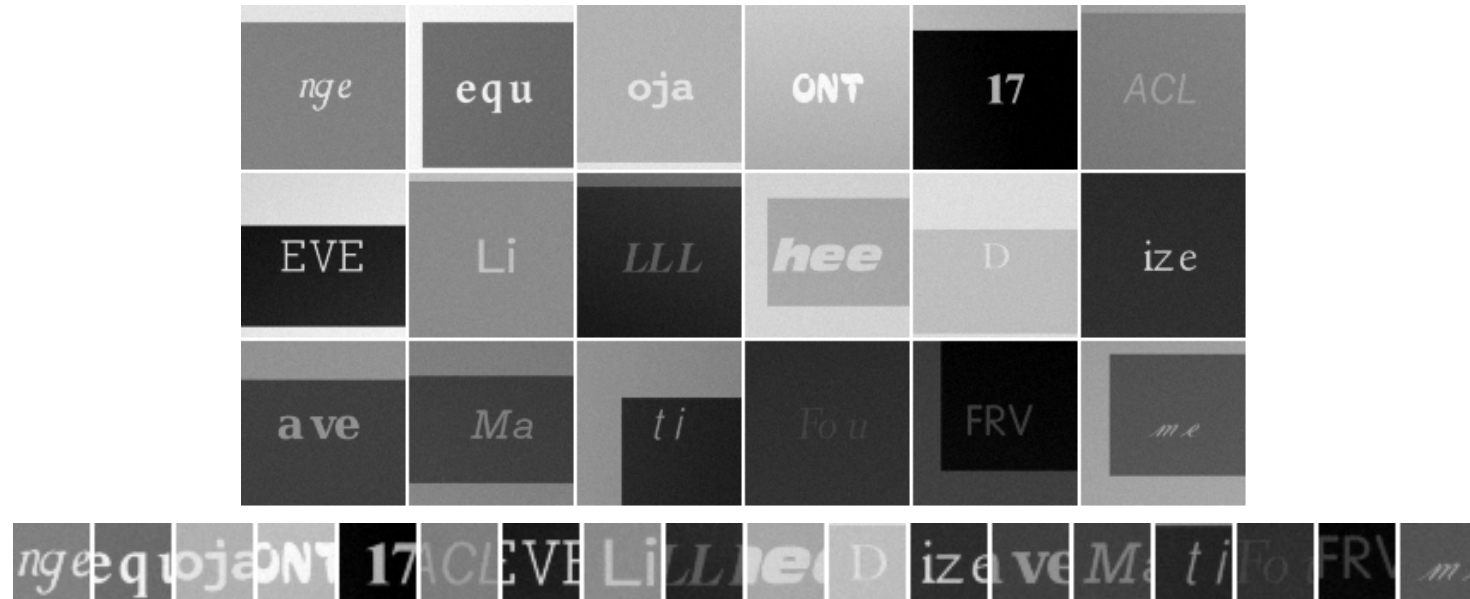
$$U^P(y_i, y_{i+1}, \mathbf{x})$$

- Width Bigram

$$U^W(y_i, y_{i+1}, w_i, w_{i+1})$$

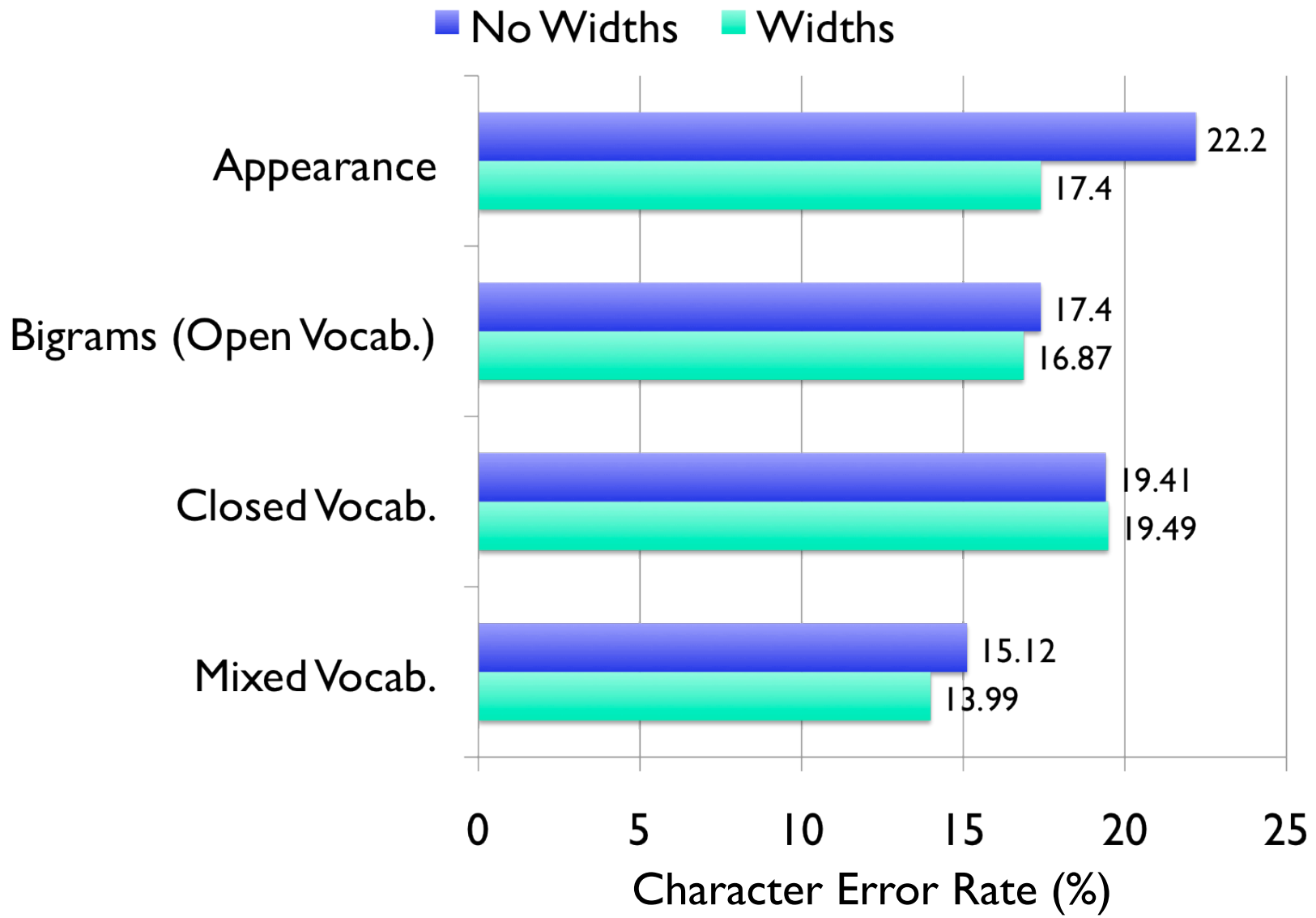
# Experimental Data

- Synthetic characters (1600 x 3 fonts)



- English text (82 books, Project Gutenberg)
- Lexicon (50th freq. percentile, SCOWL)
- 1140 char., 12 pixel x-height from 85 signs

# Results



# Recognition Comparison v. OmniPage

DOUGLASS

DOUGLASS

uucL, ass

Free checking,

Free checking

Free ehe in

LIBRARY

LIBRARY

LIBRFIRY

AMHERST

AMHERST

A), 1HERbT

TAVERN

TAVERN

TAVF

Traditional Asian  
Healing Arts

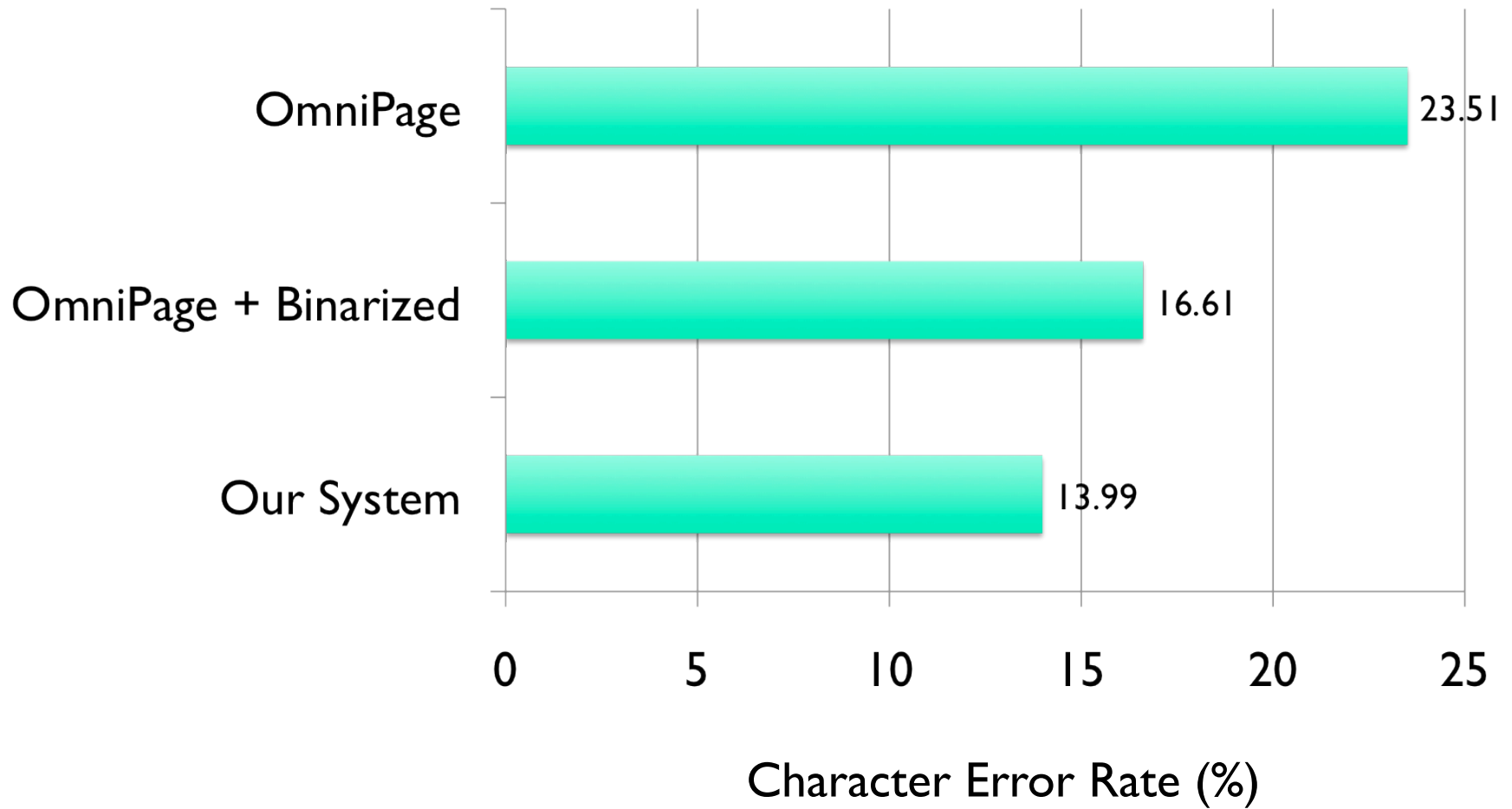
Tradiliorpal Asi'?

Traditional Asian  
Healing Arts

HealiIN Arts

# Comparison

---



# Low Resolution

---

Resumes

Resumes

WM/ I WS

krauts

11.-4her

† tttA wms

# Summary

---

- No binarization
- No prior word segmentation
- Joint character/word segmentation and recognition
- Open vocabulary with a lexical bias
- Learned typographical regularities

**Thank you!**



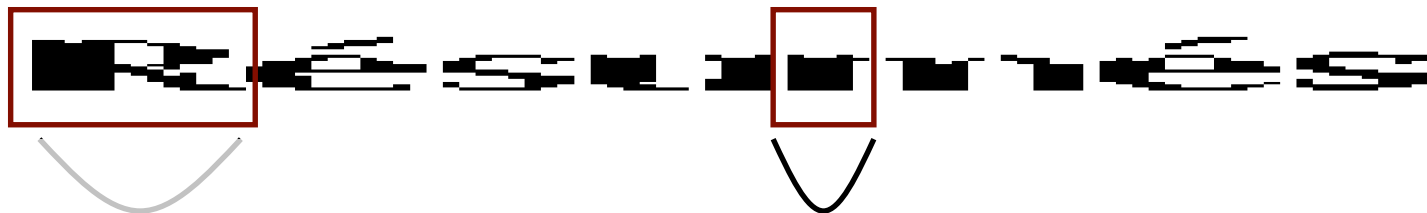
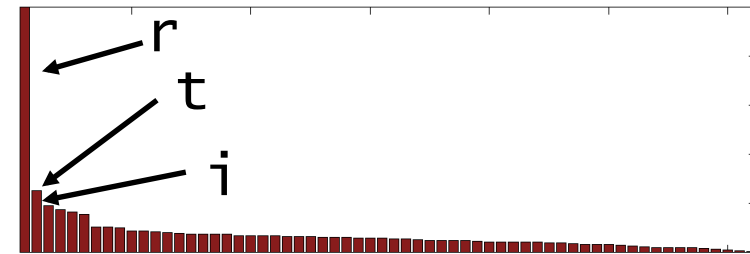
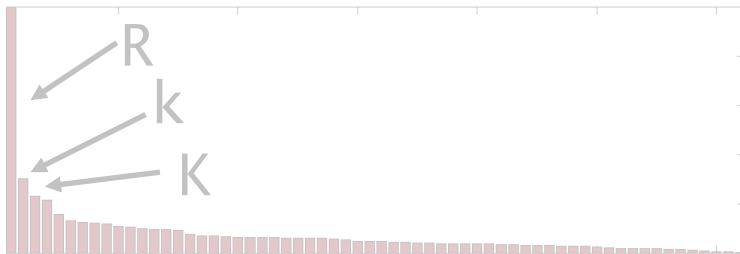
# Results

---

<b>System</b>	<b>Char Err.</b>
OmniPage	23.51%
OmniPage + Binarized	16.61%
Our System	13.99%

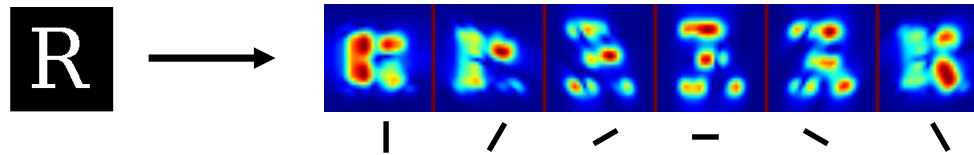
# Interpretation Scoring

---



# Recognition Information

- Appearance



- Bigrams

$$P(\text{TH} | \text{English}) = \frac{39}{1000}$$

$$P(\text{QU} | \text{English}) = \frac{1.4}{1000}$$

$$P(\text{IN} | \text{English}) = \frac{21}{1000}$$

$$P(\text{QA} | \text{English}) = \frac{.0001}{1000}$$

- Lexicon



{  
a  
Aaberg  
Aachen  
:  
:  
:  
:  
zymurgy  
Zyuganov  
Zzz

# Interpretation Graph

---

[Weinman et al. 2008]

[Jacobs et al. 2005]

[Lucas et al. 2003]

