Unconstrained Signature and Logo Detection and Matching for Off-line Document Image Retrieval

Guangyu Zhu, David Doermann

Laboratory for Language and Media Processing
Institute for Advanced Computer Studies
University of Maryland, College Park

Agenda

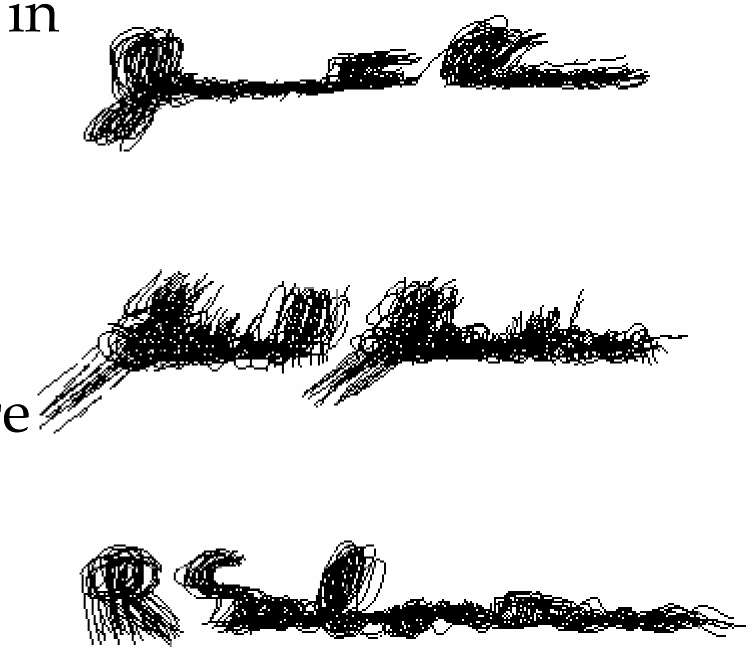
- Signature Detection Using Multi-scale Structural Saliency (to appear in CVPR 2007)
- Signature Matching for Document Image Retrieval
- Unconstrained Logo Detection in Document Images

Motivations

- Signatures and signed initials provide a new dimension for document image retrieval
- Two important aspects of this problem
  - ❑ Signature detection
  - ❑ Signature recognition
- Solution to this problem will greatly benefit off-line signature verification and identification in a range of application domains
- Signature detection is largely an open problem in literature

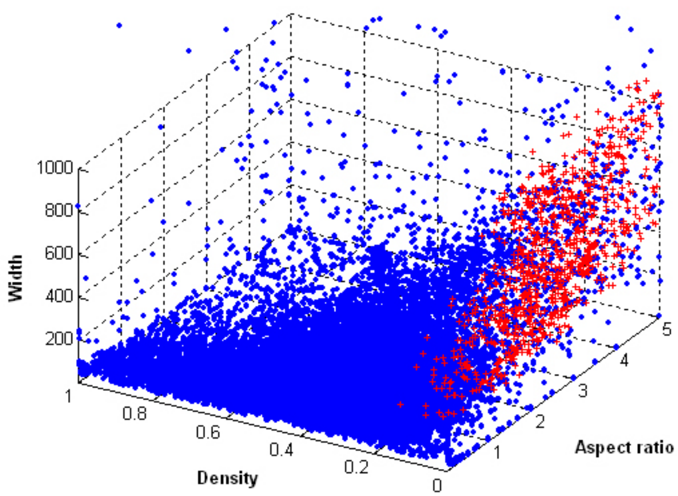
Challenges

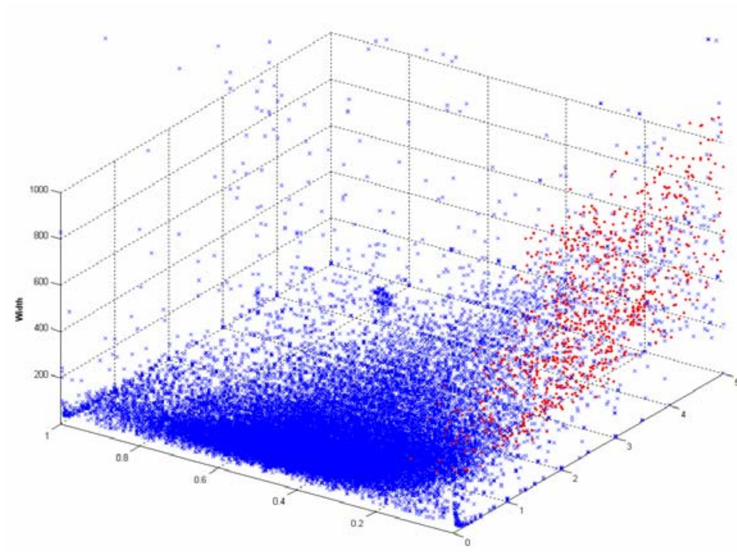
- Detecting free-form objects in cluttered backgrounds is a challenging problem in computer vision
- 2D nature of off-line signatures
  - Difficult to recover tempo order of unconstrained off-line handwriting [1]
- Large intra-class variations of signature
  - Intersession variability
  - Larger variations than other forms of handwriting
- Computation complexity



Intersession variability shown by Sabourin *et al.* [6]

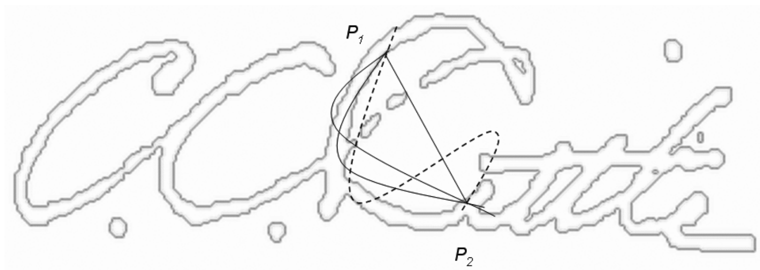
Intra-class Variations of Signatures





Overview of our approach

- We treat a signature as a global symbol. Rather than focusing on local features that typically have large variations, our approach aims to capture the structural saliency of a signature by searching over multiple scales
- We consider identifying salient structure and grouping its parts in two separate steps
- Two keys questions we addressed are:
    - How to effectively model off-line signature production under reasonable assumptions without its temporal information
    - What to effectively measure the structural saliency of signatures under such production model



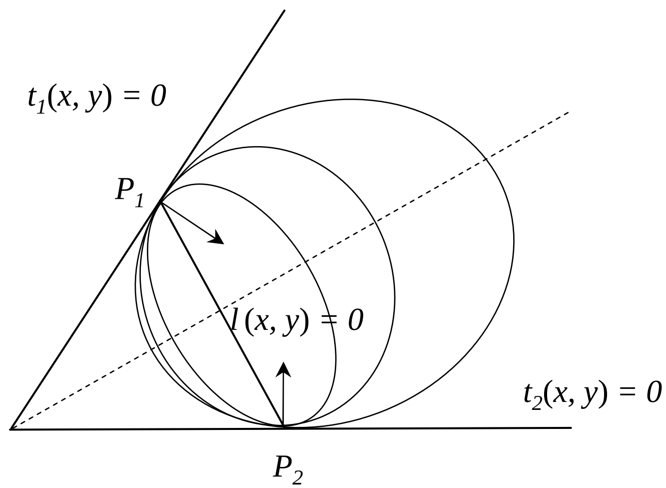
UMIACS
University of Maryland Institute for Advanced Computer Studies

Signature production model

- We assume that
  - The wrist moves in a cycloid fashion with reference to a sequence of shifting virtual baselines.
  - Local baseline changes as the wrist moves its position with respect to the document.
  - Within a local curve segment, we consider that the baseline remains unchanged.
  - The locus of the pen maintains a proportional distance from the local center point (*focus*) to the local baseline (*directrix*).
- This is equivalent to viewing signatures as piece-wise concatenations of small elliptic segments.
- The model imposes one additional constraint that limits the group of feasible second-order curves to smoother ellipses.

Measure of saliency for signatures

- How to measure the global saliency of a signature in the form of dynamic curvature without recovering its temporal order.



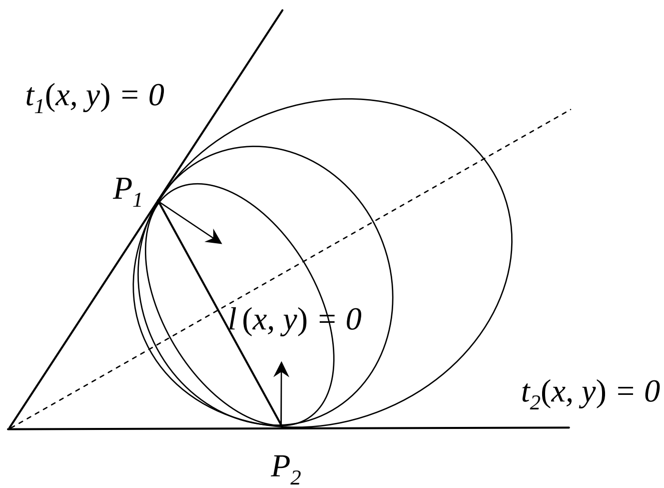
- Knowing two points P_1 and P_2 and their gradient directions, we know a family of second-order curves that pass both points

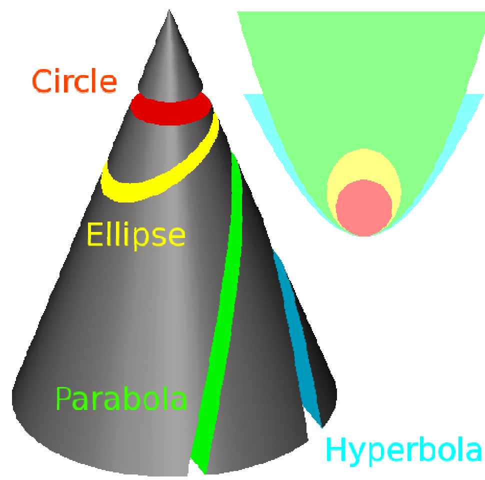
$$f(x,y) \equiv l^2(x,y) - \lambda t_1(x,y) t_2(x,y) = 0$$
$$= ax^2 + 2hxy + by^2 + 2gx + 2fy + c = 0 \qquad (1)$$

UMIACS
University of Maryland Institute for Advanced Computer Studies

Measure of saliency for signatures

- In the Cartesian coordinate system, the graph of a quadratic equation in two variables is always a conic section



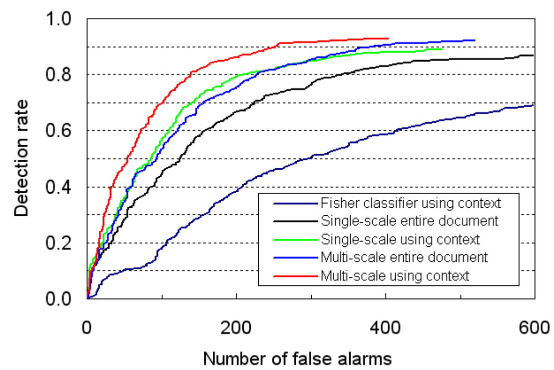


- For two points on a signature, i.e. for a set of $\{(x_1, y_1), (x_2, y_2), (p_1, q_1), (p_2, q_2)\}$, the range of λ value that corresponds to ellipses

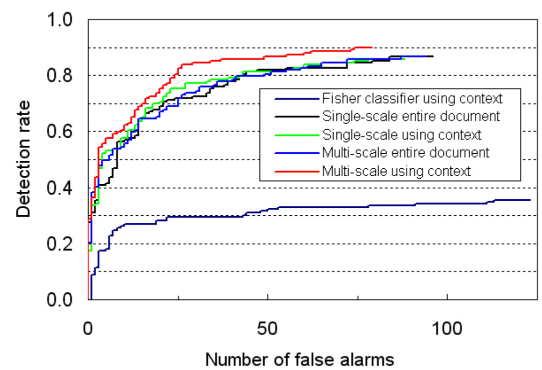
$$0 < \lambda < \frac{4[p_1(x_2 - x_1) + q_1(y_2 - y_1)][p_2(x_1 - x_2) + q_2(y_1 - y_2)]}{(p_1 q_2 - p_2 q_1)^2}$$

Evaluation

- We used two large collections of real world documents — Tobacco-800 and University of Maryland Arabic datasets.
- Using document context, our multi-scale signature detector achieves 92.8% and 86.6% detection rates for the Tobacco-800 and Maryland Arabic datasets, at 0.3 false-positives per image.



(a)



(b)

ROC curves for (a) Tobacco-800 dataset and (b) Maryland Arabic dataset.

Evaluation

Examples of detected signatures from Tobacco-800 and their saliency maps.

Evaluation

Examples of detected signatures from Maryland Arabic dataset and their saliency maps.

Evaluation

(a)

(b)

Examples of (a) falsely alarms (b) missed signatures

References

1. D.S. Doermann and A. Rosenfeld, Recovery of temporal information from static images of handwriting, *Int'l J. Computer Vision*, vol. 15, pp. 143-164, 1995.
2. R. Plamondon and S. N. Srihari, On-Line and off-Line handwriting recognition: a comprehensive survey, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 63-84, Jan. 2000.
3. A. Sha'ashua and S. Ullman. Structural saliency: the detection of globally salient structures using a locally connected network, In *Proc. Int'l Conf. Computer Vision*, pages 321–327, 1988.
4. T. Alter and R. Basri. Extracting salient curves from images: an analysis of the saliency network. *Int'l J. Computer Vision*, 27(1):51–69, 1998.
5. E.A. Maxwell, *The Methods of Plane Projective Geometry Based on the Use of General Homogeneous Coordinates*. Cambridge, England: Cambridge Univ. Press, 1946.
6. R. Sabourin, G. Genest, F. J. Preteux, Off-Line signature verification by local granulometric size distributions, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 9, pp. 976-988, Sept 1997.
7. J. M. Hollerbach, An oscillation theory of handwriting, AI Technical Report 534, March 1981.
8. J. Canny, A computational approach to edge detection, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679-697, 1986.
9. J. K. Guo and M.Y. Ma, Separating handwritten material from machine printed text using hidden Markov models, in *Proc. Int'l Conf. Document Analysis and Recognition*, pp. 439-443, 2001.
10. Y. Zheng, H. Li, D. Doermann, Machine printed text and handwriting identification in noisy document images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.26, no. 3, pp. 337-353, March 2004.

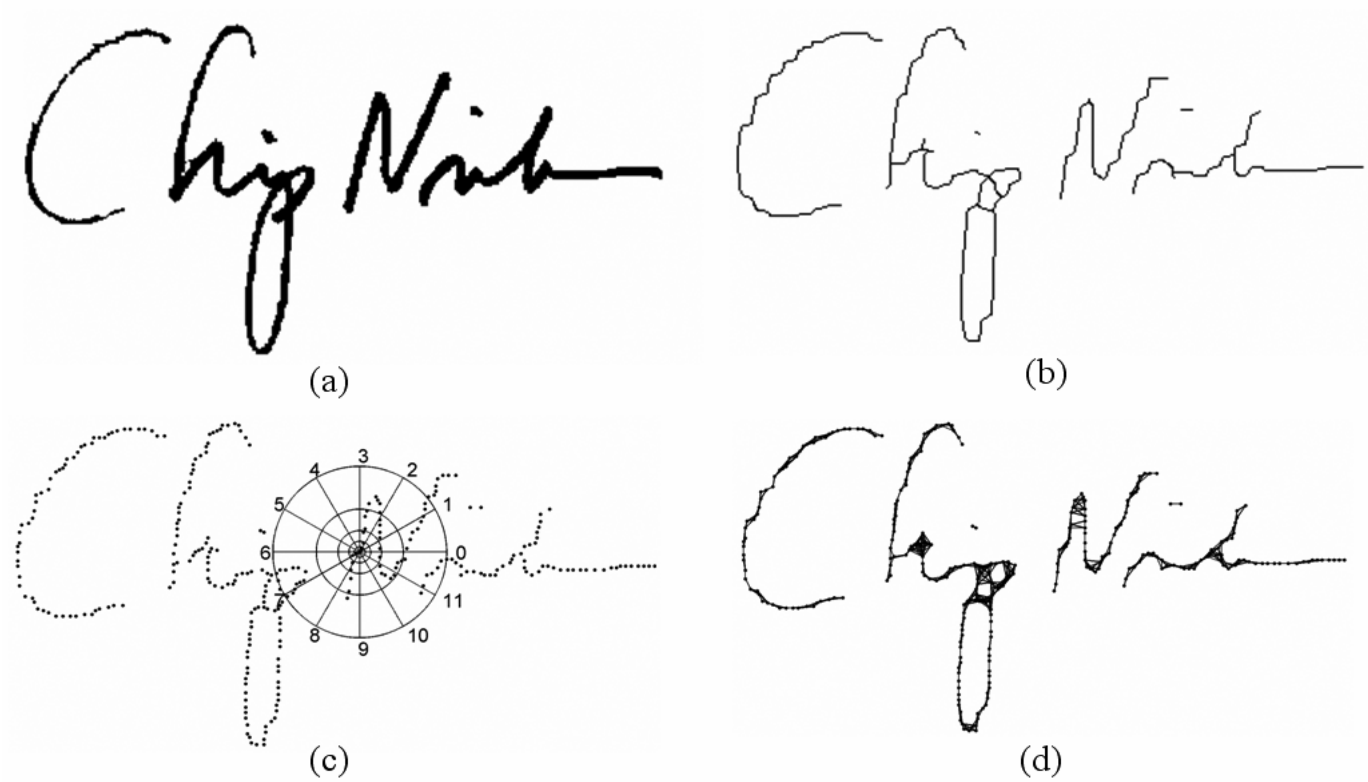
Outline

- Signature Detection Using Multi-scale Structural Saliency
- Signature Matching for Document Image Retrieval
- Unconstrained Logo Detection in Document Images

Motivations

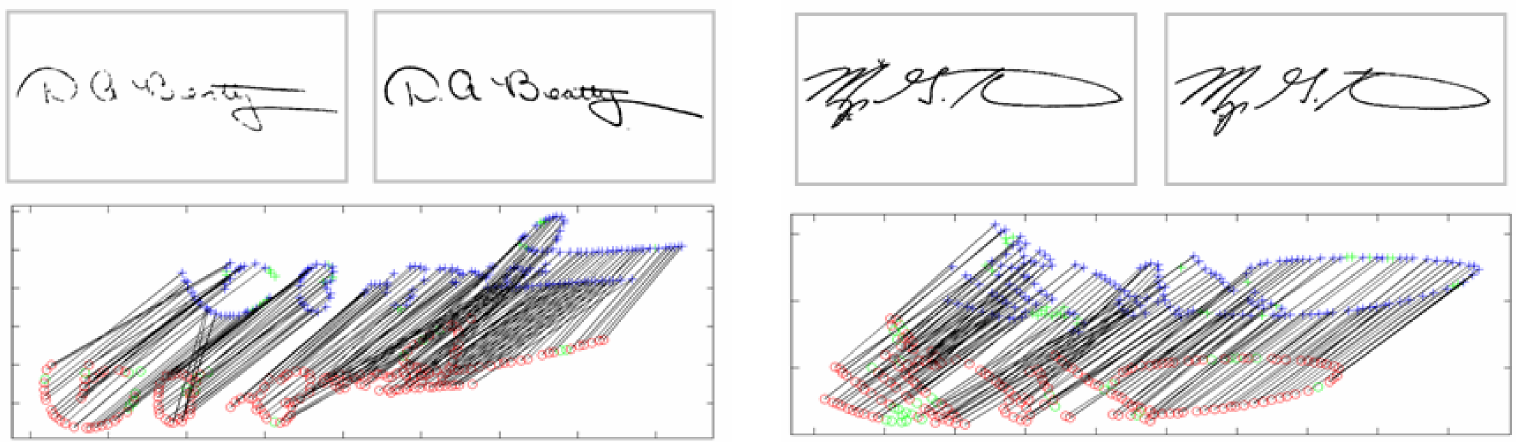
- We treat a signature as a shape
- Employ shape matching techniques for signature recognition
  - Shape representations
  - Shape matching algorithms
  - Measure of dissimilarities for shapes

Shape representation



Shape matching using points sampled from skeletons. (a) Original signature. (b) Extracted skeleton [7]. (c) Shape context descriptor [4]. (d) Local neighborhood graph [5].

Shape matching



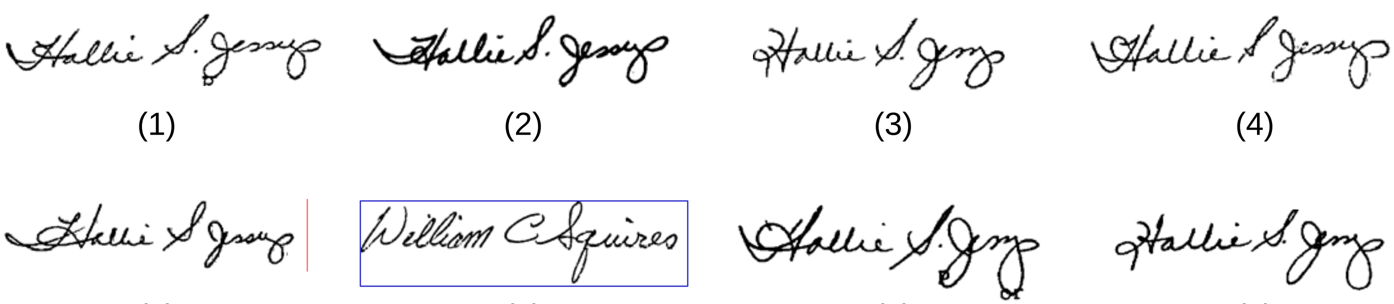
(a) (b)

Visualization of shape matching results using the graph-based non-rigid shape matching algorithm. For both signatures, we use 200 point sampled along their skeletons. After 5 iterations, 181 and 170 points are matched in (a) and (b), respectively.

Shape matching evaluation

A query with eight relevant signature instances

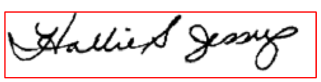
Top eight retrieved in the ranked list



(1) (2) (3) (4)

(5) (6) (7) (8)

Relevant instance outside the top eight in the ranked list



(10)

A signature query example. Among the total of eight relevant signature instances, seven appear in the top eight of the 460-element ranked list, giving an average precision of 94.2%, and an R-Precision of 87.5%. The irrelevant signature that is ranked among the top eight is highlighted with a dashed box.

UMIACS University of Maryland Institute for Advanced Computer Studies

Signature matching results

Table 1: Signature retrieval result using different similarity measures.

Similarity measures	Mean average precision	Mean R-Precision
D_{sc}	66.9%	62.8%
D_{af}	61.3%	57.0%
D_{be}	59.8%	55.6%
D_{re}	52.5%	48.3%
$D_{sc} + D_{be}$	78.7%	74.3%
$D_{sc} + D_{af} + D_{sc} + D_{re}$	84.5%	80.8%

Table 2: Signature retrieval result using multiple instances of signatures from the same person in each query.

Number of instances	Mean average precision	Mean R-Precision
One	84.5%	80.8%
Two	88.6%	85.2%
Three	91.3%	88.1%

References

1. S. Loncaric, "A survey of shape analysis techniques," Pattern Recognition, vol. 31, no. 8, pp. 983-1001, 1998.
2. R.C. Velkamp and M. Hagedoorn, "State of the art in shape matching," Utrecht University, Netherlands, Tech. Rep. UU-CS-1999-27, 1999.
3. H. Chui and A. Rangarajan, "A new point matching algorithm for non-rigid registration," Computer Vision and Image Understanding, vol. 89, nos. 2-3, pp. 114-141, 2003.
4. S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," IEEE Trans. Pattern Anal. Machine Intell., vol. 24, no. 4, pp. 509-522, 2002.
5. Y. Zheng and D. Doermann, "Robust point matching for non-rigid shapes: a relaxation labeling based approach," IEEE Trans. Pattern Anal. Machine Intell., vol. 28, no. 4, pp. 643-649, April 2006.
6. C.R. Dyer and A. Rosenfeld, "Thinning algorithms for gray-scale pictures," IEEE Trans. Pattern Anal. Machine Intell., vol. 1, no. 1, pp. 88-89, 1979.
7. T.Y. Zhang and C.Y. Suen, "A fast parallel algorithm for thinning digital patterns," Communications of the ACM, vol. 27, no. 3, pp. 236-239, 1984.

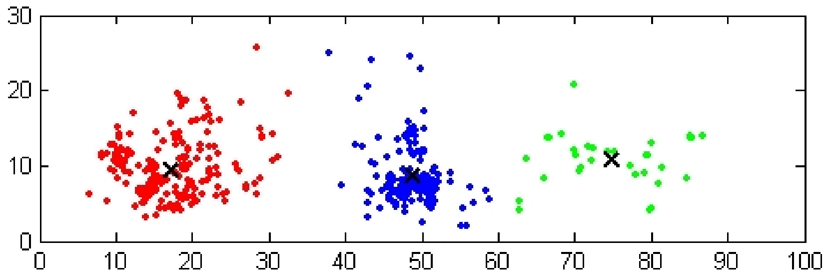
Outline

- Signature Detection Using Multi-scale Structural Saliency
- Signature Matching for Document Image Retrieval
- Unconstrained Logo Detection in Document Images (to appear in ICDAR 2007)

Overview

- Propose a joint formulation for logo detection and extraction using a boosting strategy across multiple image scales
- At a coarse scale, a trained Fisher classifier performs an initial classification using features from document context and connected components.
- Each logo candidate region is further classified at successively finer image scales by a cascade of simple classifiers

Feature selection and extraction



Positions of logos in the Tobacco-800 dataset relative to the entire document.

We define context distance as

$$D_c(P) = \min_{i \in \{1,2,\cdots,k\}} (|p_x - c_x^i| + \lambda |p_y - c_y^i|)$$

Table 3: Features used for classification.

Context distance	Aspect ratio
Spatial density	Area

UMIACS
University of Maryland Institute for Advanced Computer Studies

Evaluation

- We use accuracy and precision as evaluation metrics

$$\text{Accuracy} = \frac{\text{\# of correctly detected logos}}{\text{\# of logos in groundtruth}}$$

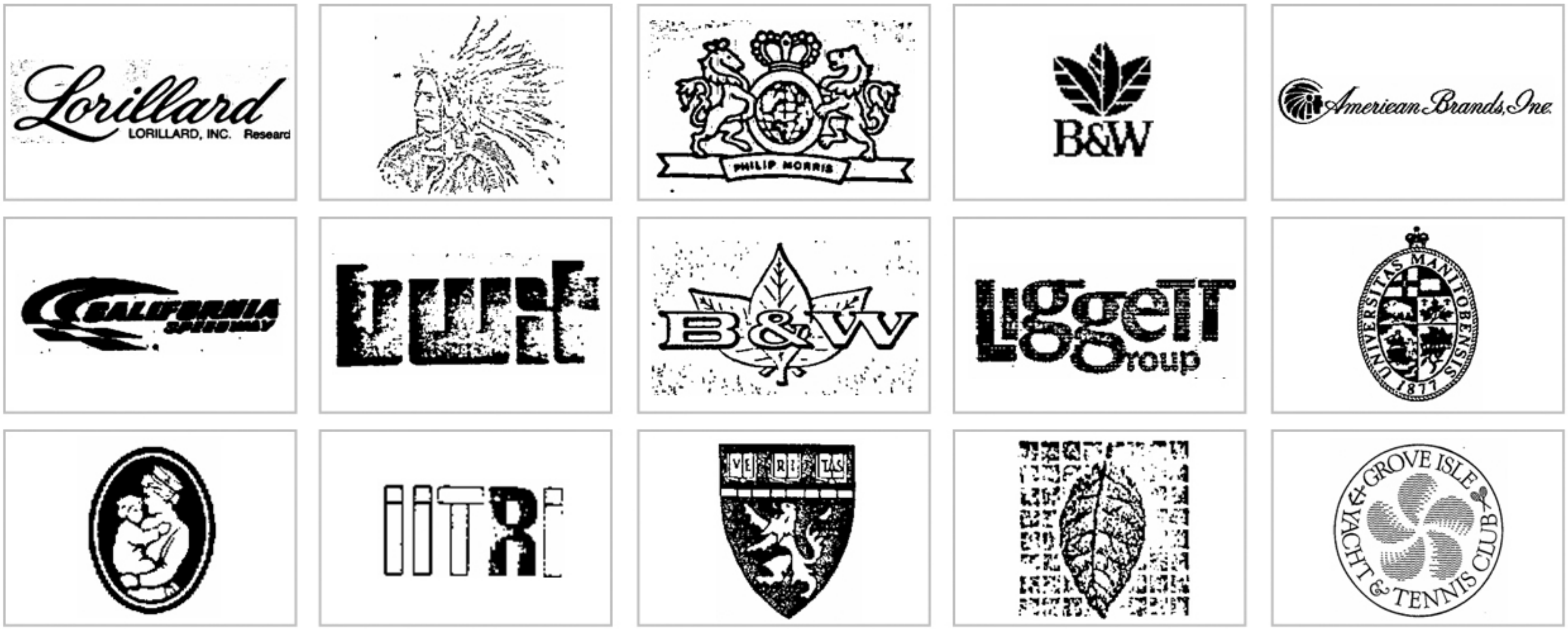
$$\text{Precision} = \frac{\text{\# of correctly detected logos}}{\text{\# of detected logos}}$$

- We consider a logo *correctly detected* if and only if the detected region contains more than 75% overlapping pixels with the groundtruth AND its area is less than 125% of the area of the groundtruth.

Table 4: Positions of logos in the Tobacco-800 dataset relative to the entire document.

	Accuracy	Precision
Improved spatial density [8]	39.3%	32.1%
Fisher classifier only, *i.e.* $\lvert\mathcal{S}\rvert = 1$	59.2%	41.7%
Multi-scale approach with $\lvert\mathcal{S}\rvert = 2$	57.0%	68.1%
Multi-scale approach with $\lvert\mathcal{S}\rvert = 3$	84.2%	73.5%

Evaluation



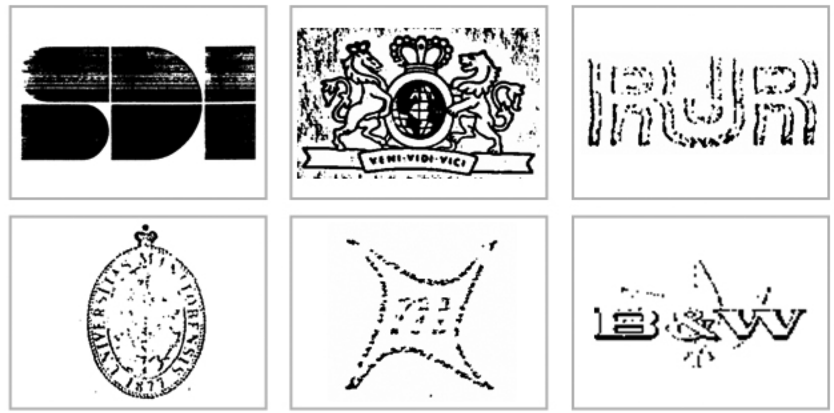
Examples of correctly detected logos from Tobacco-800.

Evaluation

(a) Over/under-segmented logos

(b) Non logos

Examples of incorrectly detected logos.



Examples of missed logos.

References

1. The Legacy Tobacco Document Library at University of California, San Francisco, http://legacy.library.ucsf.edu/.
2. The University of Maryland (UMD) Logo Database, http://documents.cfar.umd.edu/resources/database/UMDlogo.html.
3. F. Cesarini, E. Francesconi, M. Gori, S. Marinai, J. Sheng, and G. Soda. A neural-based architecture for spot-noisy logo recognition. In *Proc. Int'l Conf. Document Analysis and Recognition*, pages 175–179, 1997.
4. D. Doermann, E. Rivlin, and I. Weiss. Applying algebraic and differential invariants for logo recognition. *Machine Vision and Application*, 9(2):73–86, 1996.
5. J. Neumann, H. Samet, and A. Soffer. Integration of local and global shape analysis for logo classification. *Pattern Recognition Letters*, 23(12):1449–1457, 2002.
6. S. Seiden, M. Dillencourt, S. Irani, R. Borrey, and T. Murphy. Logo detection in document images. In *Proc. Int'l Conf. Imaging Science, Sys., and Tech.*, pages 446–449, 1997.
7. P. Suda, C. Bridoux, B. Kammerer, and G. Maderlechner. Logo and word matching using a general approach to signal registration. In *Proc. Int'l Conf. Document Analysis and Recognition*, pages 61–65, 1997.
8. T. Pham. Unconstrained logo detection in document images. Pattern Recognition, 36(12):3023–3025, 2003.
9. K. Turkowski. Filters for common resampling tasks. *Graphics Gems I*, pages 147–165, 1990.
10. R. Yager and D. Filev. Approximate clustering via the mountain method. *IEEE Trans. System, Man and Cybernetics*, 24(8):1279–1284, 1994.
11. K. Zyga, J. Schroeder, and R. Price. Logo recognition using retinal coding. In *Proc. 38th Asilomar Conf. Signals, Systems and Computers*, vol 2, pages 1549–1553, 2004.

Multi-scale detection

- Connected components are only meaningful over a very small range of image scales
- Using a multi-scale classification and refinement scheme gives more precise signature localization and reduces false alarms