Combining Local Features For Offline Writer Identification

Rajiv Jain and David Doermann

Language and Multimedia Processing Laboratory

Department of Computer Science

University of Maryland, College Park

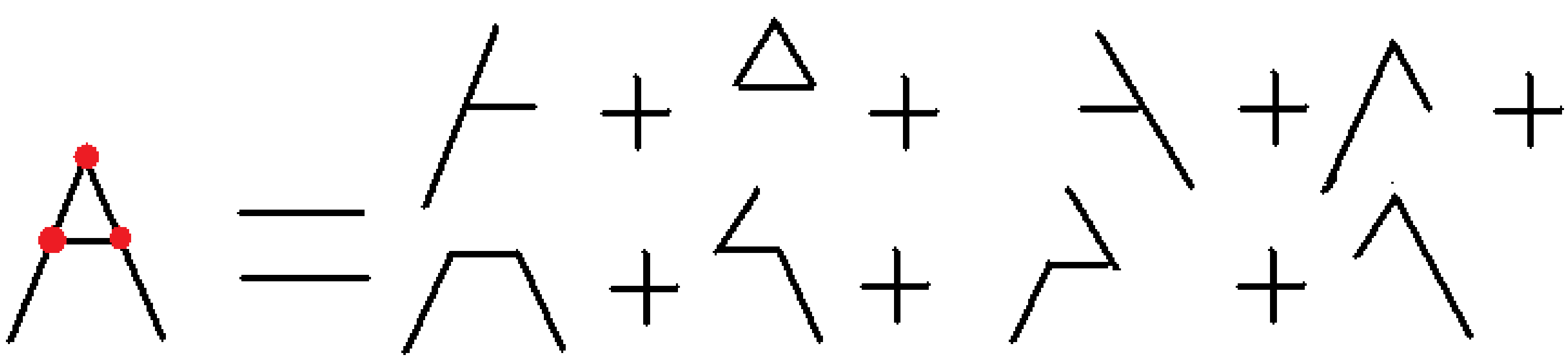
INTRODUCTION

- Writer Identification
  - Retrieve all handwriting samples in a collection from a given author
  - Assist Forensic Examiners / Historians with large collections
- Many strong local features recently proposed
  - Features desired that capture writer's variation
  - 7 different local features in ICDAR 2013 Writer ID contest
- Hypothesis
  - Combinations of complementary, strong features should provide state of the art performance
- We study 3 features
  - Edge / Stroke - K-Adjacent Segments
  - KeyPoint - SURF
  - Allograph Features – Contour Gradient Descriptor

K-Adjacent Segments (KAS)

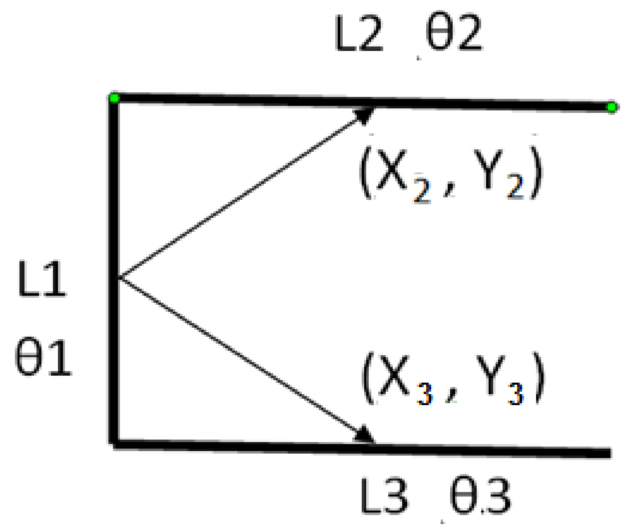
- Extract line segments from image contours

- Take all K-adjacent line segments (K=3)

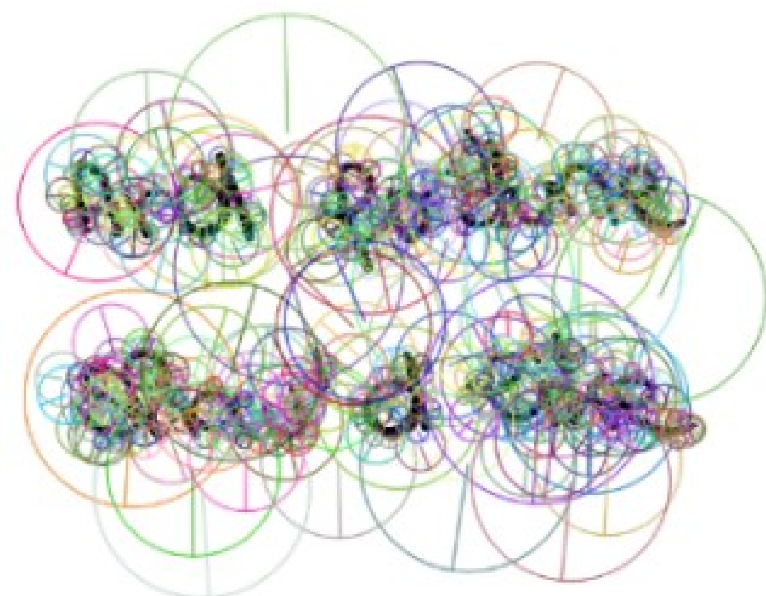


K-Adjacent Segments (KAS)

- Describe each line segment with a feature vector
  - $\frac{r^{x_2}}{N_d}, \frac{r^{y_2}}{N_d}, \dots, \frac{r^{x_k}}{N_d}, \frac{r^{y_k}}{N_d}, \frac{\theta_1}{.5*\pi}, \dots, \frac{\theta_k}{.5*\pi}, \frac{L_1}{N_d}, \dots, \frac{L_k}{N_d}$



SURF



- SURF (64-D) slightly better than SIFT

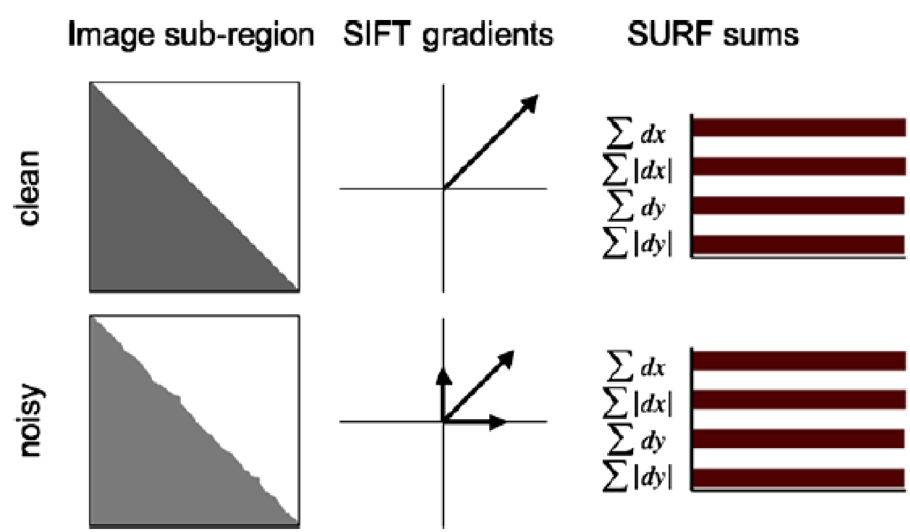
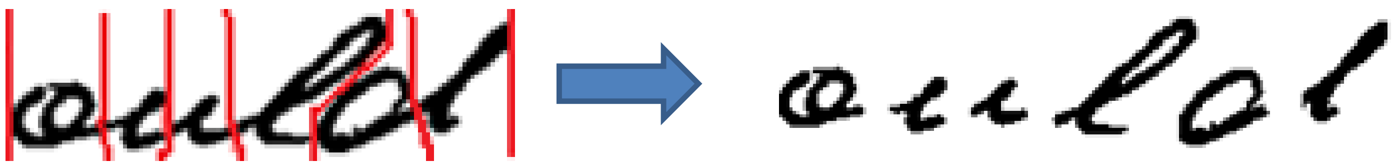


Image borrowed from
Speeded up Robust Features,
Bay et. al. 2008

CONTOUR GRADIENT DESCRIPTOR

- Allograph Approach
  - Segment Words in to character-like segments



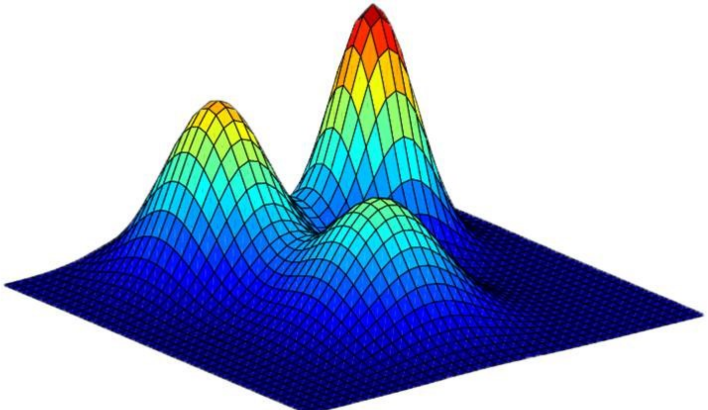
- Describe characters by contour gradients
  - 4x4 grid with 8 orientation bins

Fisher Vector

- Popular feature aggregation technique
  - Won 2011 ImageNet Challenge (Perronnin, 2011)
  - Outperforms BOF for Writer ID using SIFT (Fiel, ICDAR 2013)
- Step 1: Create a Gaussian Mixture Model

$$u_\lambda(x) = \sum_{k=1}^{K} w_k u_k(x)$$

$$u_k(x) = \frac{1}{2\pi^{D/2} * |\Sigma_k|^{1/2}} e^{-\frac{1}{2}(x-u_k)' * \Sigma_k^{-1} * (x-u_k)}$$



FISHER VECTOR

- Step 2: Accumulate partial derivatives for u and σ

$$\gamma_t(k) = \frac{w_k u_k(x_t)}{\sum_{j=1}^{K} w_j u_j(x_t)}$$

$$G_{u,k}^{X} = \frac{1}{T\sqrt{w_k}} \sum_{t=1}^{T} \gamma_t(k) \left(\frac{x_t - u_k}{\sigma_k} \right)$$

$$G_{\sigma,k}^{X} = \frac{1}{T\sqrt{2 * w_k}} \sum_{t=1}^{T} \gamma_t(k) \left(\frac{(x_t - u_k)^2}{\sigma_k^2} - 1 \right)$$

- Step 3: L2 and Power Normalization
- Final feature (super)vector is 2*K*D
- Number of Gaussians = 64 in experiments

Combining Features

- Cosine distance used for comparing Fisher Vectors
- Distance between samples A and B with K features:

$$D(A,B,K) = \sum_{k=1}^{K} w_k * \left(FV_k(A) \bullet FV_k(B) \right)$$

Where,

$$1 = \sum_{k=1}^{K} w_k$$

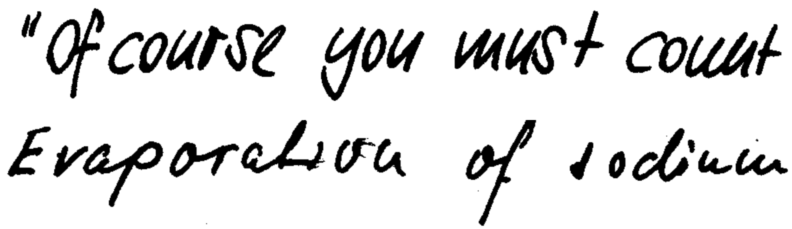
- Grid search on training set used to select w_k
- Baseline of setting $w_k = \frac{1}{3}$ for all three features

Experimental Setup

- Experimental Procedure
  - 4 Datasets spanning 3 scripts (Roman, Arabic, Greek)
  - Evaluated retrieval in a leave-one-out manner
  - GMM and w_k (feature weight) training
- Evaluation Measures
  - Soft Top N – At least one of Top N results by same writer
  - Hard Top N – All Top N results by same writer
  - Mean Average Precision

IAM

IAM Dataset
- English
- 650 Writers – 2 Samples



Features	Top-1	Top-2	Top-5	Top-10	MAP
K	88.8	91.1	95.0	96.4	0.914
S	90.0	92.4	96.2	97.6	0.926
C	91.3	93.8	96.6	97.6	0.936
K&S&C	94.1	96.0	98.2	98.5	0.958
K&S&C*	94.7	95.9	98.1	98.7	0.960
Chain Code	91	N/R	N/R	97%	N/R
Edge+ CO3	89	N/R	N/R	96%	N/R
SIFT+SOH	**98.5**	N/R	**99.1**	**99.5**	N/R

2013 ICDAR Writer ID Contest

- 250 Writers
  - Experiment A: 2 Greek Samples
  - Experiment B: 2 English Samples

Σωκράτης δίδασκε ότι η

Greek philosopher.

Greek

Features	Top-1	Top-2	Top-5	Top-10	MAP
K	93.2	95.6	98.0	99.0	0.952
S	94.6	97.2	98.8	99.2	0.964
C	97.2	98.6	99.2	99.6	0.984
K&S&C	98.2	99.0	99.4	**99.8**	0.988
K&S&C*	**99.2**	**99.6**	**99.8**	**99.8**	**0.995**
SIFT+FV	88.4	92.0	96.8	97.8	n/a
SIFT+SOH	93.8	96.4	97.2	97.8	n/a
Edge+Run Len	92.6	96.0	98.0	98.4	n/a

English

Features	Top-1	Top-2	Top-5	Top-10	MAP
K	92.4	94.4	96.4	97.2	0.942
S	94.6	96.2	97.6	98.0	0.959
C	96.4	97.2	98.0	98.6	0.971
K&S&C	97.0	**97.8**	98.0	98.6	0.976
K&S&C*	**97.4**	**97.8**	**98.6**	**98.8**	**0.979**
SIFT+FV	91.4	94.2	95.8	97.2	n/a
SIFT+SOH	92.2	94.6	96.4	96.6	n/a
Edge+Run Len	91.2	93.4	96.2	96.6	n/a

CVL Dataset

- 309 Writers
  - 4 English Samples
  - 1 German Sample

Lines, Triangles, Squares,

generally differ

Soft Criterion

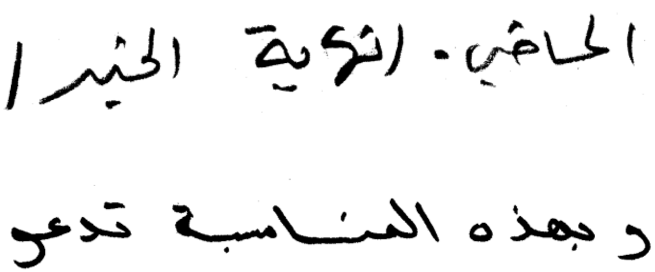
Features	Top-1	Top-2	Top-5	Top-10
K	98.5	99.1	99.2	99.5
S	98.7	99.2	99.4	99.5
C	97.0	98.1	99.0	99.4
K&S&C	**99.4**	**99.5**	99.5	**99.7**
K&S&C*	**99.4**	**99.5**	**99.6**	**99.7**
SIFT+FV	97.8	98.6	99.1	99.6
Edge + Run Len	97.6	97.9	98.3	98.5
Grid MicroStruc	97.7	98.3	99.0	99.1

Hard Criterion

Features	Top-2	Top-3	Top-4	MAP
K	94.3	85.9	66.2	0.927
S	96.1	88.5	70.7	0.941
C	91.0	77.8	52.3	0.881
K&S&C	**98.3**	**95.2**	80.8	0.966
K&S&C*	**98.3**	94.8	**82.9**	**0.969**
SIFT+FV	95.6	89.4	75.8	n/a
Edge + Run Len	94.3	88.2	73.0	n/a
Grid MicroStruc	95.3	94.5	73.0	n/a

DARPA MADCAT

- 300 Arabic Writers
  - 10 samples each



Soft Criterion

Features	Top-1	Top-2	Top-5	Top-10
K	96.8	98.1	99.1	99.4
S	92.8	94.2	95.5	96.0
C	96.9	98.2	99.4	**99.6**
K&S&C	97.1	98.0	99.0	99.3
K&S&C*	**97.8**	**98.6**	**99.4**	99.5

Hard Criterion

Features	Top-2	Top-3	Top-5	Top-7	Top-9	MAP
K	93.3	90.5	82.2	68.4	39.7	86.4
S	87.1	82.1	69.5	51.8	17.0	72.2
C	93.2	90.1	80.9	67.6	40.1	86.8
K&S&C	94.2	91.4	86.6	76.3	43.3	87.9
K&S&C*	**95.4**	**93.2**	**87.5**	**78.0**	**50.9**	**90.1**

CONCLUSION

- Conclusions
  - No one feature is optimal on all datasets
  - Learning weights to combine 3 local features provides state-of-the-art results
- Future Work
  - Graduate
  - Add more features?
    - Dozen(s) of other local features
    - Better Feature Fusion
    - Is this the right answer?
  - Feature Learning
    - Larger datasets are needed

QUESTIONS?