

Writer Recognition by Computer Vision

Jeffrey P. Woodard MITRE CORP

Christopher P. Saunders MITRE CORP & South Dakota State University

Mark J. Lancaster USG

Measurement Science and Standards in Forensic Handwriting Analysis Conference

June 4-5, 2013, NIST Gaithersburg, MD

Approved by MITRE for Public Release: 10-2431

Approved for Public Release: 12-4194

Distribution Unlimited

MITRE

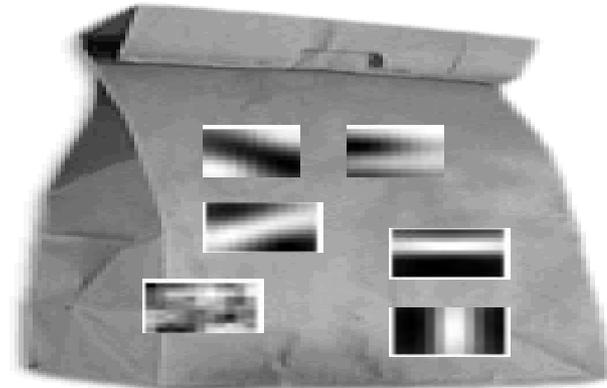
MITRE Computer Vision Approach

- Algorithms used common in science and engineering
- First known application to forensic handwriter recognition
- Public domain intellectual property
- No linguistic analysis
- No segmentation into units
- Supervision only for classification

Overview

■ Methods

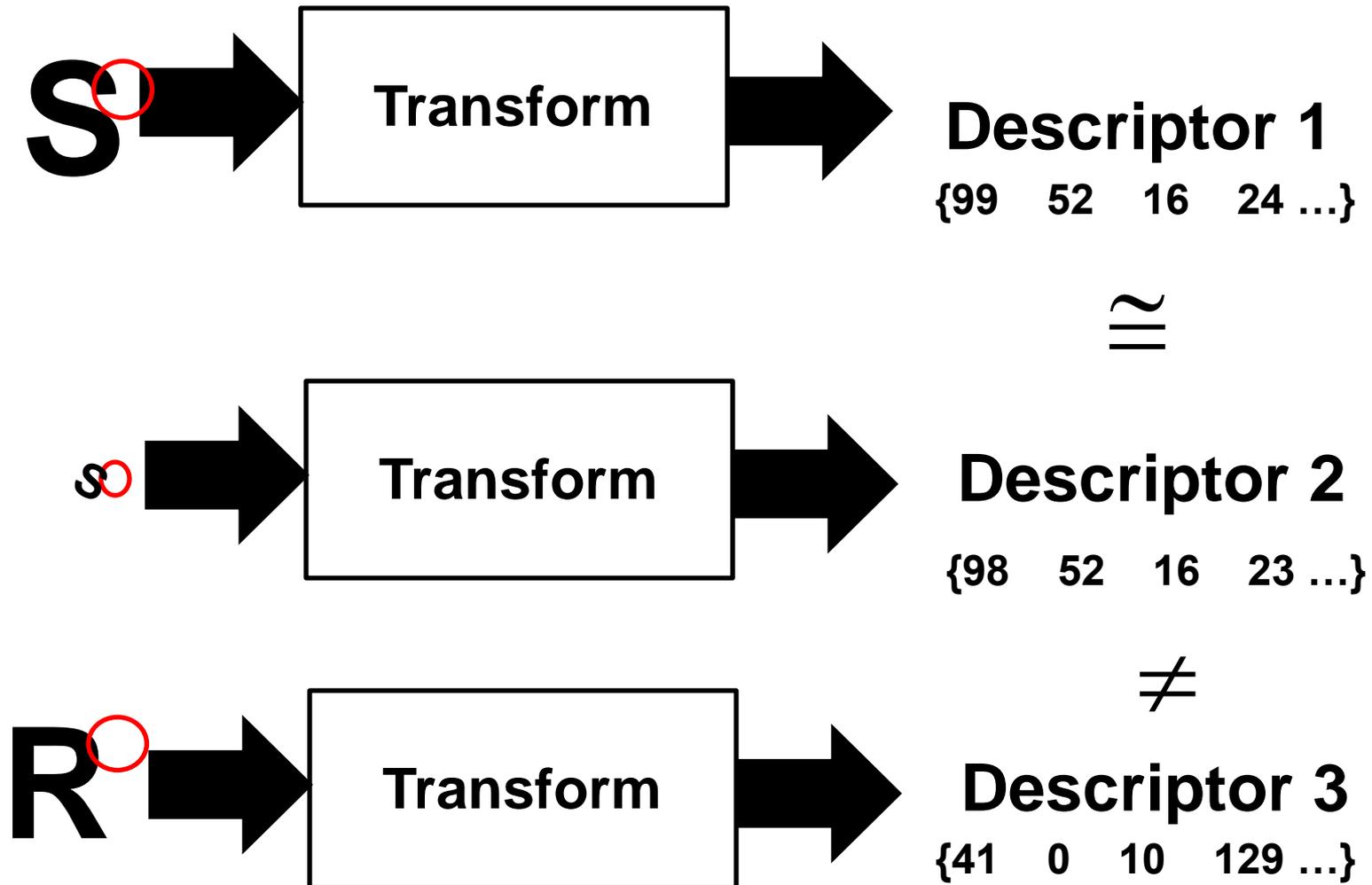
- Features
- Clustering
- Bag of Words
- Latent Approaches



■ Experiments

- Arabic
- Dutch

Local Features: distinctiveness & invariance



Lowe's Scale Space

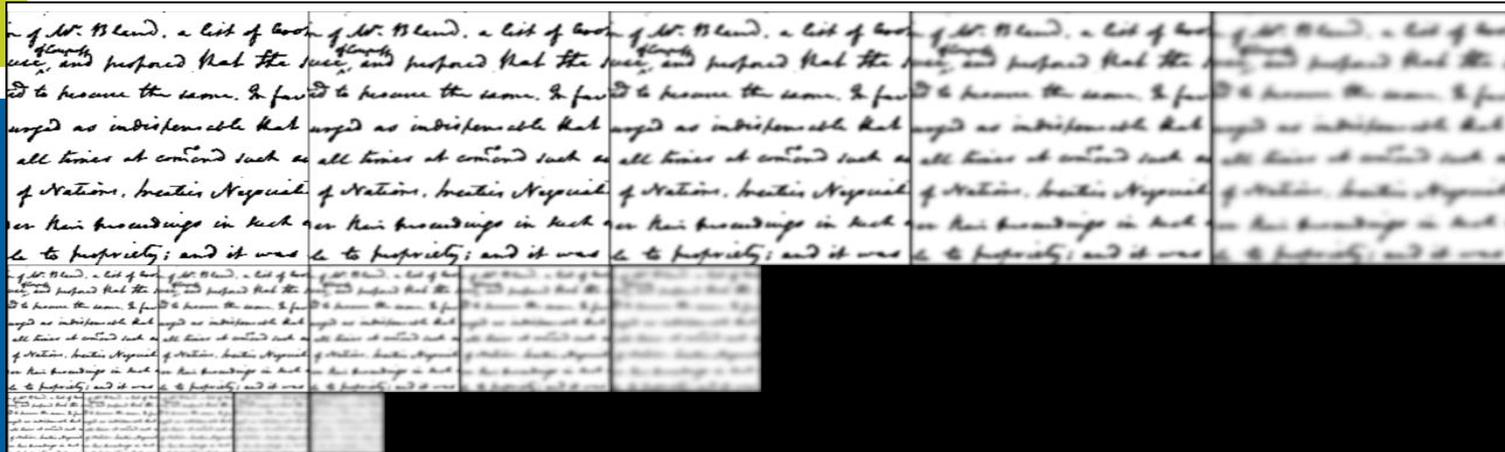
Scale 1

Scale 2

Scale 3

Scale 4

Scale 5

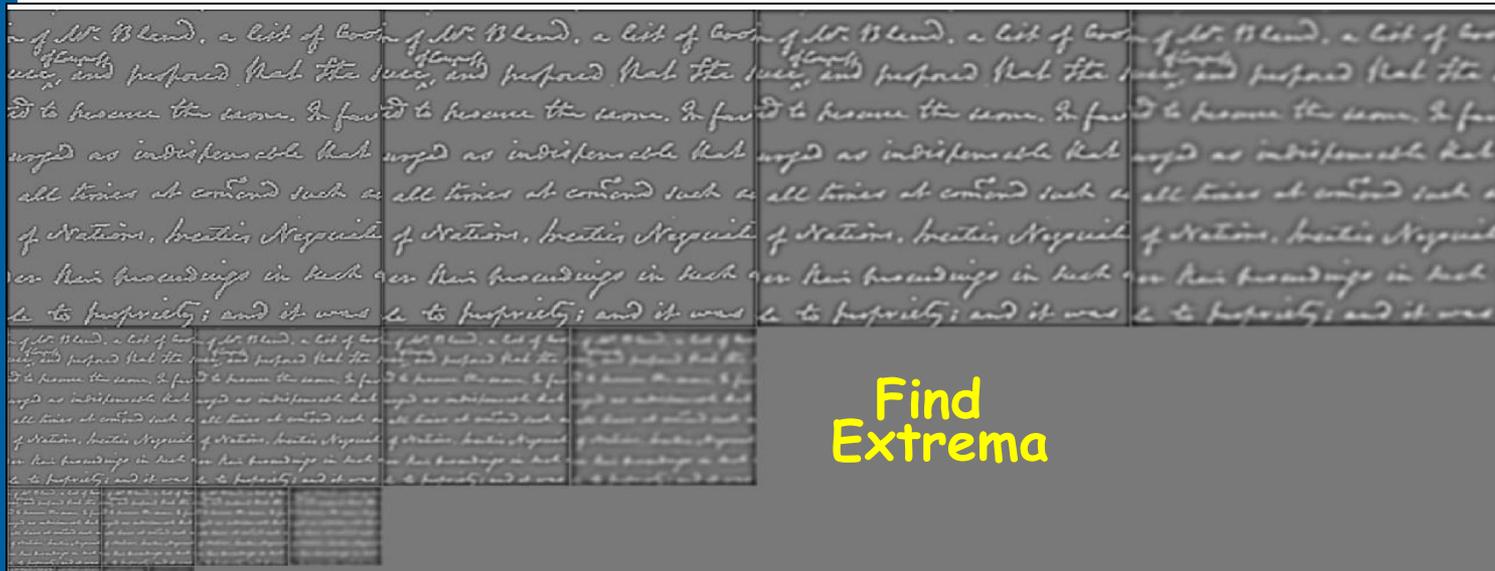


Octave 1

Octave 2

Octave 3

Gaussian Representation



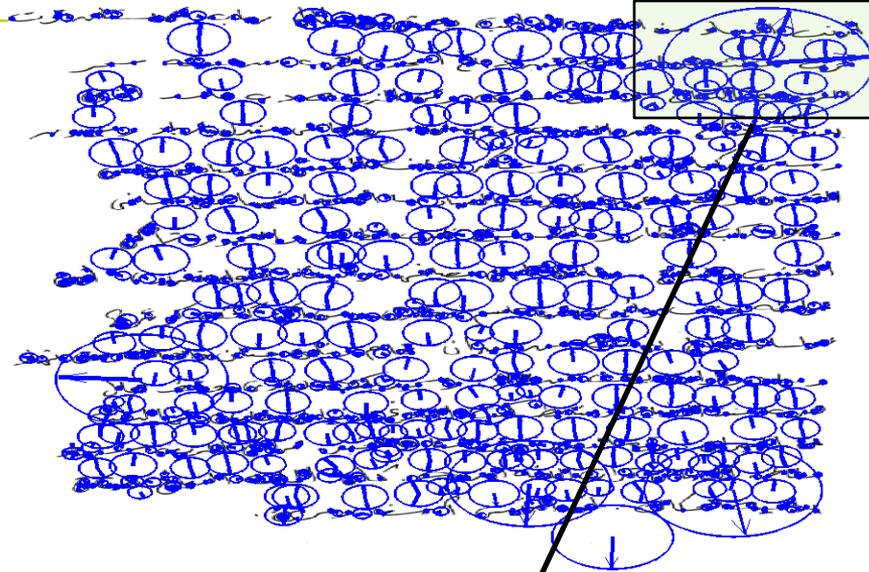
Find Extrema

Difference of Gaussians Representation

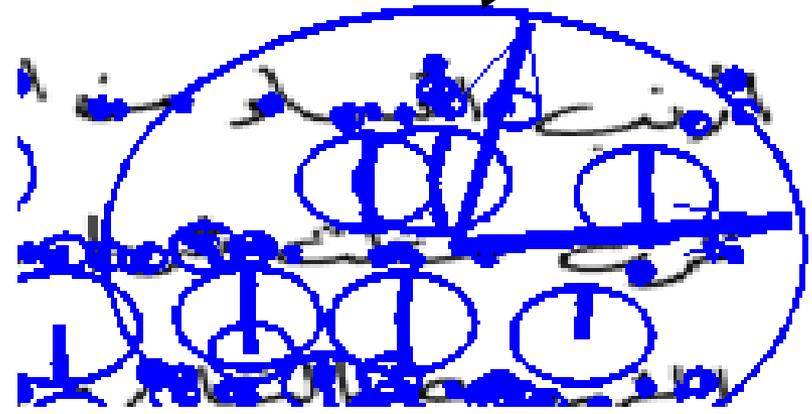
Local Features

Circle center: feature x-y position
 Circle radius: feature scale
 Circle arrow: dominant orientation(s)

أرنب أسود من أقوى الأراب يعيش مع بلبل رباب ثم يستأنه قوت
 ترب مثلث شريا رفيع قبه دجاج أجد بالقرب سرمدية سر
 المزرعة بالتقاع . وقوف غفري من المزرعة دج حامد الذي
 يدرب ذراع مستر القوية رزم على سريها ثم مثل نريد وسبور
 رساس بيكرية الثور وشاكر يقطف الشمش . ومدق يقمن
 القمصن والضخم يقول لجه العناد ما أهدوا فيها نعل حاشي
 وظلال ركب القطار رمانف شتيع نلمت الظاهر والعيور تأكل
 العنكب والنفع . بخانة تلبس ماضها والبيغار تقول تفريد ، أين
 تراشه حفاف وأين بقرة تاسم وطابق ومن أكل كعكه كوشر؟
 ءولسه فليل بييمك سره ميزان قام وساي حنان والصلت يدق نهار
 وسباد والداري بييمك عهد اللؤلؤ . وسعيد كيت لغادي وقط حمال
 صرد منذ الصباح وصاحب كني البلاغ ونيان شارح من اليعك
 مع الجيران . ربع الحجج من نسي نمتات الضفر . نرائع الميهية شرب
 مطاطم . الضامن ترات اللغه الدية . كان الملك الدياغ الدواق
 براميه الحكاك . ان الرئيش ؟ صف ومرغ .



أرنب أسود هنا
 ترب مثلث شريا
 المزرعة بالتقاع .



Vector Quantization (VQ)

■ K-means, Generalized Lloyd clustering

Gersho, A. & Gray, R. (1992) *Vector Quantization and Signal Compression*, Boston: Kluwer Academic Press.

- Many different local descriptors quantized to small codebook -- “visual words”
- Represent each image by histogram of visual words

$$X \equiv [x_1, x_2, x_3, \dots, x_B] \in R^{B \times 128} \quad \text{Local Features}$$

Iterate

$$m_k = \frac{\sum_{i:C(i)=k} x_i}{N_k}, \quad k = 1, \dots, L$$

$$C_i^* = \arg \min_{1 \leq k \leq L} \|x_i - m_k\|_2^2, \quad i = 1, \dots, B$$

Training

$$i^{\hat{a}} \equiv \arg \min_i \|x_t - C_i^{\hat{a}}\|$$

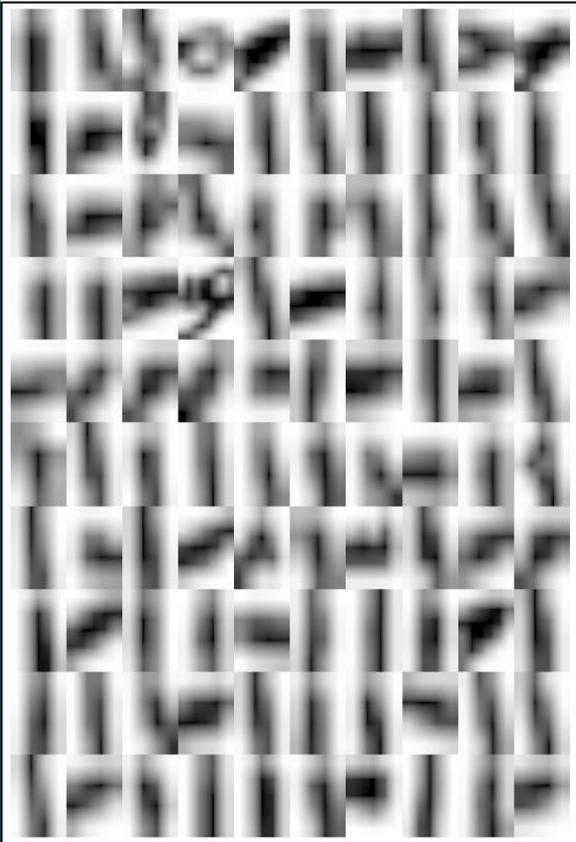
Testing

VQ “VISUAL WORDS”

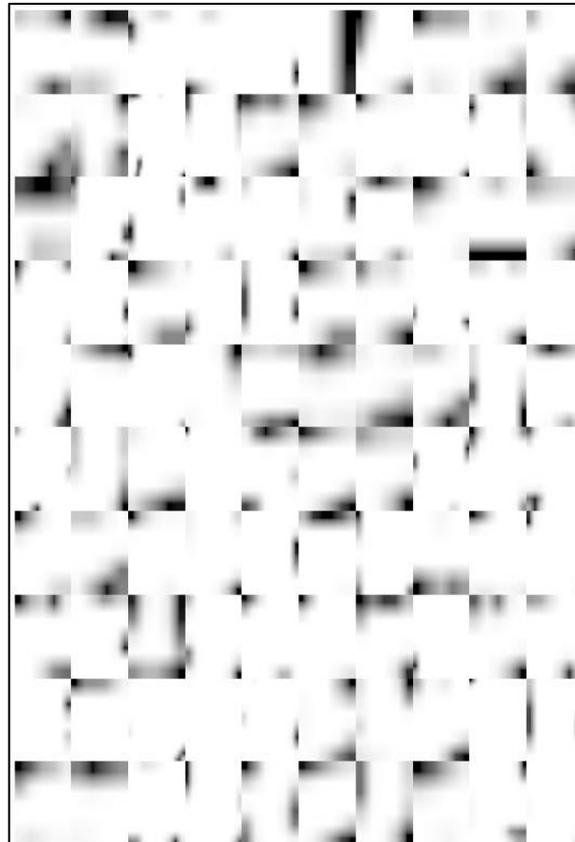
100 random image regions from Arabic, DoG detector, SIFT descriptor, 512 codewords

Region sizes vary: all displayed identically

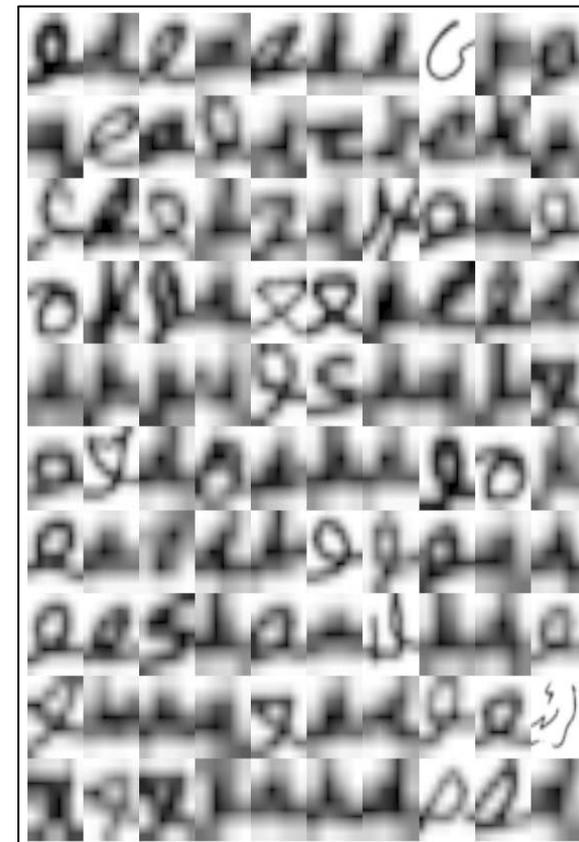
Regions for VQ codeword 1



Regions for VQ codeword 2



Regions for VQ codeword 3



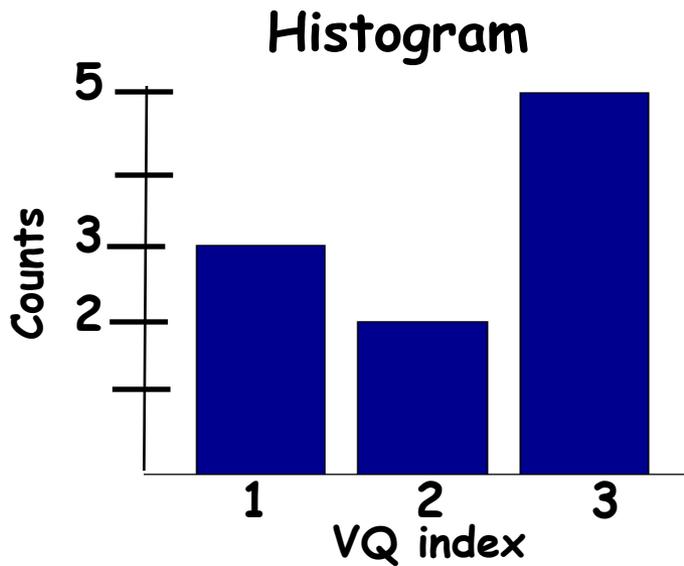
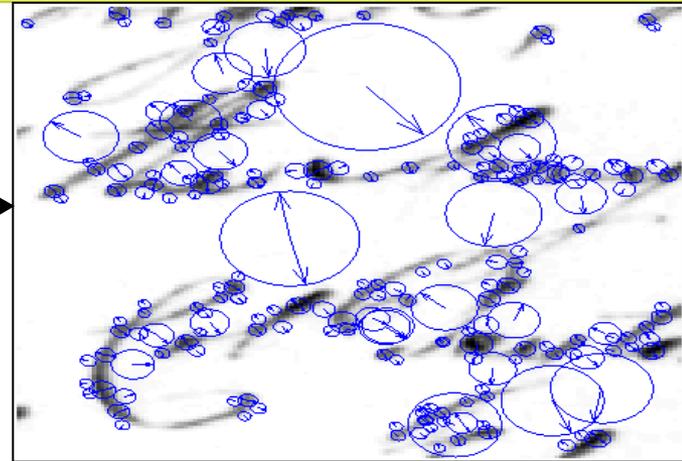
Visual words” represent fundamental properties associated with early to middle vision

Bag of Words Representation

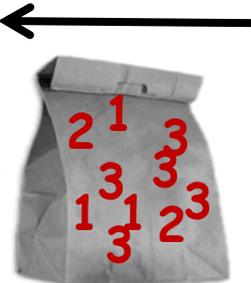
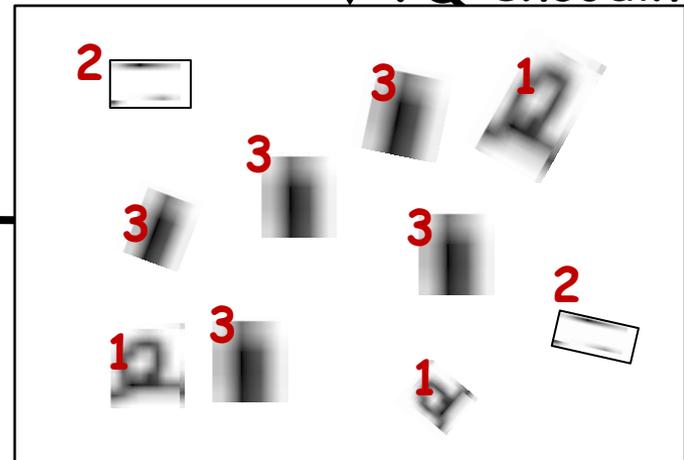
Image



Local Features



VQ encoding



First learning...

Images: Class 1

Images: Class N



Local Features



Clustering Codebook

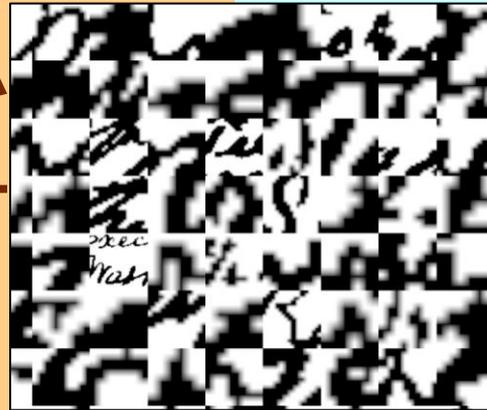


image representations

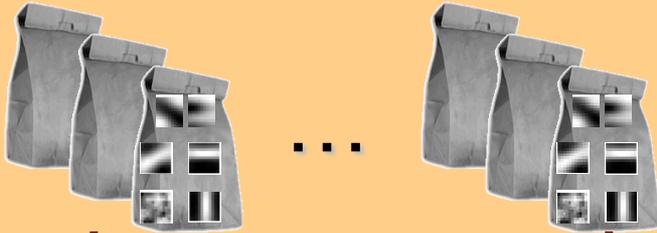


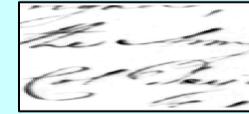
Image or Class Models

Class 1

Class N

...Then recognition

Class Unknown

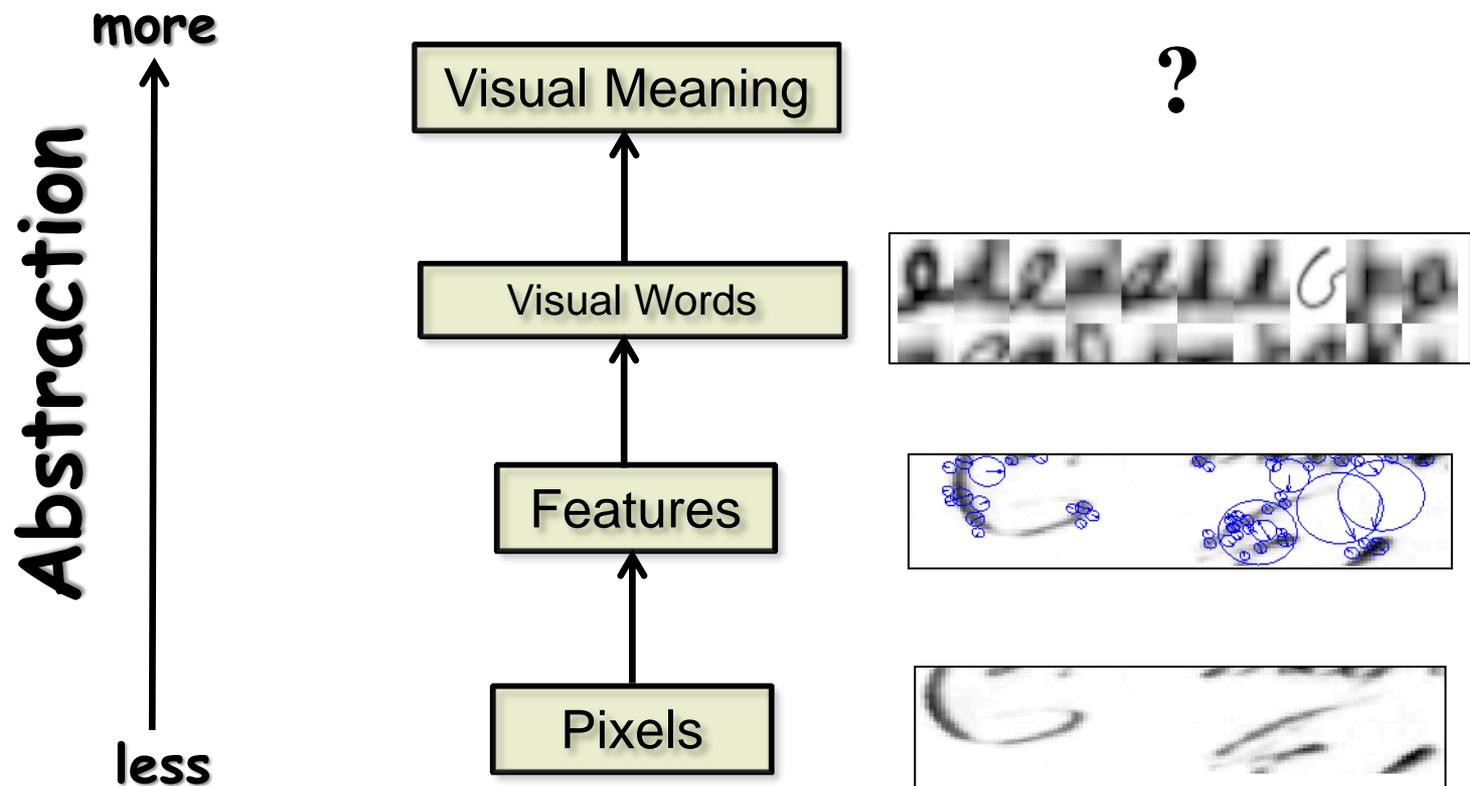


Local Features



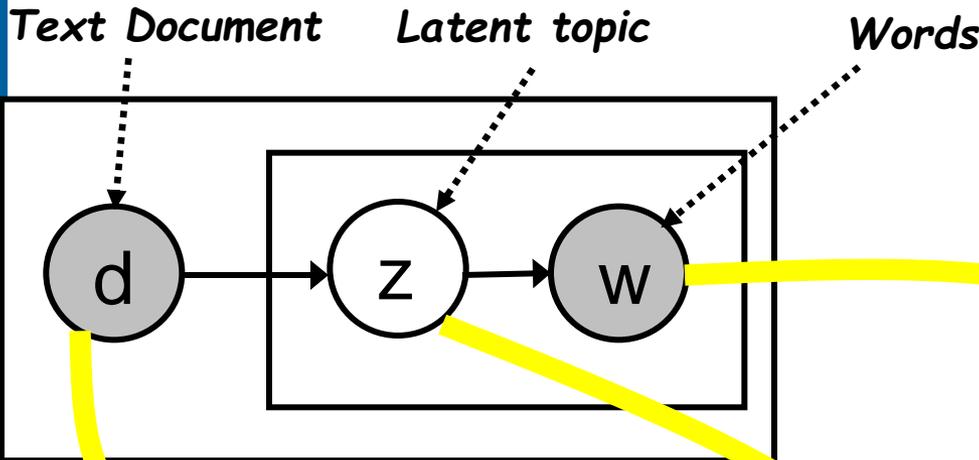
Class Decision

A Hierarchy of Visual Abstraction



probabilistic Latent Semantic Analysis (pLSA): Text Documents

documents modeled as combinations of latent topics



Generative View

- Select a document d_i with prob $P(d_i)$
- Pick latent class z_k with prob $P(z_k|d_i)$
- Generate keyword w_j with prob $P(w_j|z_k)$
- Boxes replicate

IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 50, NO. 3, MARCH 2002

A Survey of Convergence Results on Particle Filtering Methods for Practitioners

Dan Crisan and Arnaud Doucet

Abstract—Optimal filtering problems are ubiquitous in signal processing and related fields. Except for a restricted class of models, the optimal filter does not admit a closed-form expression. Particle filtering methods are a set of flexible and powerful sequential Monte Carlo methods designed to solve the optimal filtering problem numerically. The posterior distribution of state is approximated by a large set of Dirac-delta masses (samples/particles) that evolve randomly in time according to the dynamics of the model and the observations. The particles are interacting; thus, classical limit theorems relying on statistically independent samples do not apply. In this paper, our aim is to present a survey of recent convergence results on this class of methods to make them accessible to practitioners.

Index Terms—Bayesian estimation, optimal filtering, particle filtering, sequential Monte Carlo, state-space models.

I. INTRODUCTION

MANY models in signal processing can be cast in a statespace form. In most applications, prior knowledge of the system is also available. This knowledge allows us to adopt a Bayesian approach, that is, to combine a prior distribution for unknown quantities with a likelihood function relating these quantities to the observations. Within this setting, one performs inference on the unknown state according to the posterior distribution. *Often, the observations arrive sequentially in time and*

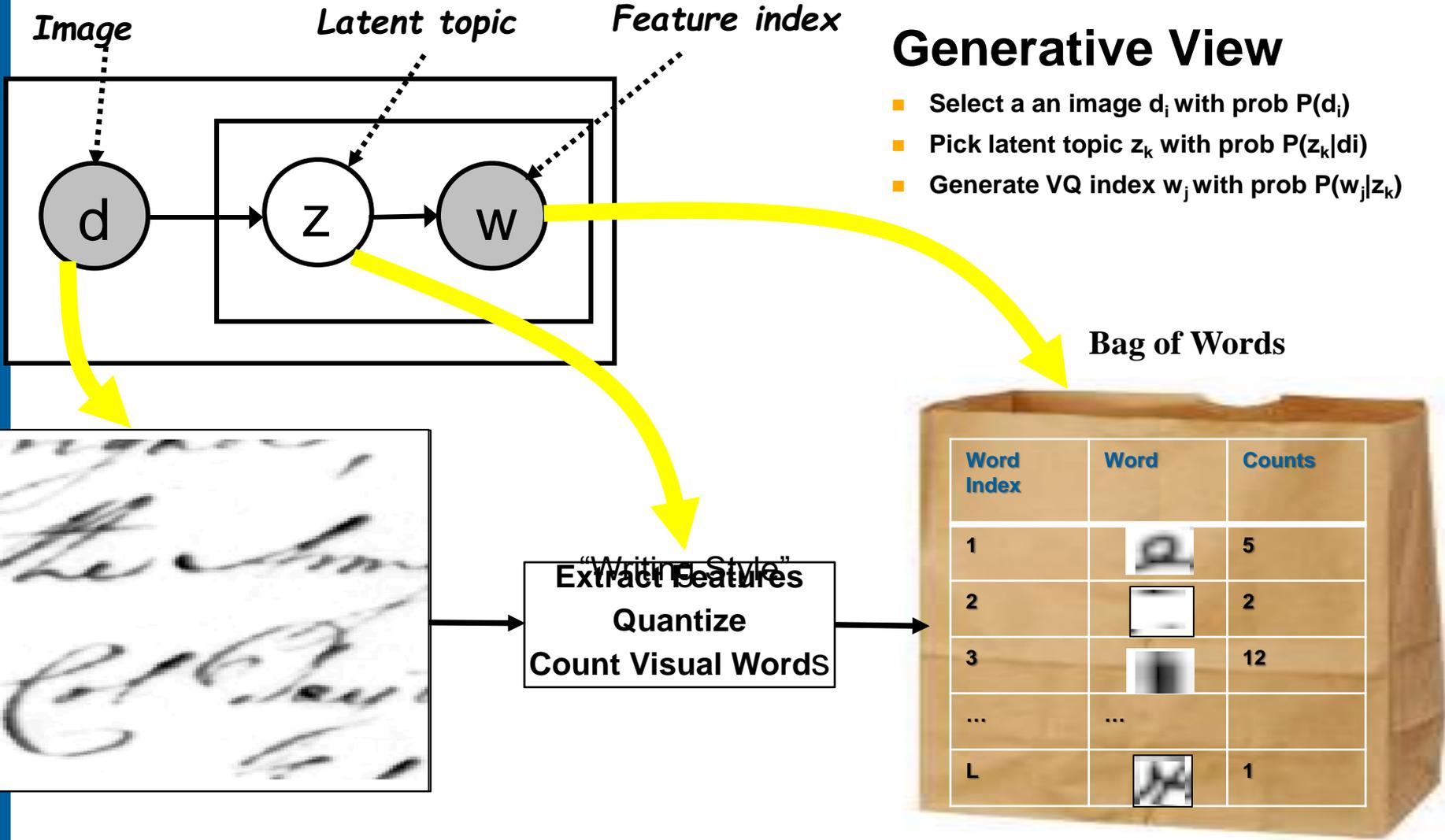
“Particle Filter”
Count Keywords

Bag of Words

Word Index	Word	Counts
1	Filter	5
2	Particle	2
j	Method	12
...	...	
N	Result	1

pLSA for Images

images modeled as combinations of latent topics



pLSA Object "Discovery"

Ranked by max $P(d|z_k) = \operatorname{argmax} P(\text{Image} | \text{Object}_k)$

Topic 1 Style = "tight"

Topic 2 Style = "loose"

أرشد أسامة من أصول الأردنية، يعيش مع أهل أبيه في شمال غرب سوريا. تربى في بيئة دينية إسلامية، حيث كان والده يقرأ القرآن الكريم في كل يوم. بدأ في تعلم اللغة العربية في سن مبكرة، وكان والده يقرأ له القرآن الكريم في كل يوم. بدأ في تعلم اللغة العربية في سن مبكرة، وكان والده يقرأ له القرآن الكريم في كل يوم. بدأ في تعلم اللغة العربية في سن مبكرة، وكان والده يقرأ له القرآن الكريم في كل يوم.

50

39

29

Writer indexes shown in green

أرشد أسامة من أصول الأردنية، يعيش مع أهل أبيه في شمال غرب سوريا. تربى في بيئة دينية إسلامية، حيث كان والده يقرأ القرآن الكريم في كل يوم. بدأ في تعلم اللغة العربية في سن مبكرة، وكان والده يقرأ له القرآن الكريم في كل يوم. بدأ في تعلم اللغة العربية في سن مبكرة، وكان والده يقرأ له القرآن الكريم في كل يوم.

39

13

39

أرشد أسامة من أصول الأردنية، يعيش مع أهل أبيه في شمال غرب سوريا. تربى في بيئة دينية إسلامية، حيث كان والده يقرأ القرآن الكريم في كل يوم. بدأ في تعلم اللغة العربية في سن مبكرة، وكان والده يقرأ له القرآن الكريم في كل يوم. بدأ في تعلم اللغة العربية في سن مبكرة، وكان والده يقرأ له القرآن الكريم في كل يوم.

13

13

9

أرشد أسامة من أصول الأردنية، يعيش مع أهل أبيه في شمال غرب سوريا. تربى في بيئة دينية إسلامية، حيث كان والده يقرأ القرآن الكريم في كل يوم. بدأ في تعلم اللغة العربية في سن مبكرة، وكان والده يقرأ له القرآن الكريم في كل يوم. بدأ في تعلم اللغة العربية في سن مبكرة، وكان والده يقرأ له القرآن الكريم في كل يوم.

46

أرشد أسامة من أصول الأردنية، يعيش مع أهل أبيه في شمال غرب سوريا. تربى في بيئة دينية إسلامية، حيث كان والده يقرأ القرآن الكريم في كل يوم. بدأ في تعلم اللغة العربية في سن مبكرة، وكان والده يقرأ له القرآن الكريم في كل يوم. بدأ في تعلم اللغة العربية في سن مبكرة، وكان والده يقرأ له القرآن الكريم في كل يوم.

46

أرشد أسامة من أصول الأردنية، يعيش مع أهل أبيه في شمال غرب سوريا. تربى في بيئة دينية إسلامية، حيث كان والده يقرأ القرآن الكريم في كل يوم. بدأ في تعلم اللغة العربية في سن مبكرة، وكان والده يقرأ له القرآن الكريم في كل يوم. بدأ في تعلم اللغة العربية في سن مبكرة، وكان والده يقرأ له القرآن الكريم في كل يوم.

46

أرشد أسامة من أصول الأردنية، يعيش مع أهل أبيه في شمال غرب سوريا. تربى في بيئة دينية إسلامية، حيث كان والده يقرأ القرآن الكريم في كل يوم. بدأ في تعلم اللغة العربية في سن مبكرة، وكان والده يقرأ له القرآن الكريم في كل يوم. بدأ في تعلم اللغة العربية في سن مبكرة، وكان والده يقرأ له القرآن الكريم في كل يوم.

48

أرشد أسامة من أصول الأردنية، يعيش مع أهل أبيه في شمال غرب سوريا. تربى في بيئة دينية إسلامية، حيث كان والده يقرأ القرآن الكريم في كل يوم. بدأ في تعلم اللغة العربية في سن مبكرة، وكان والده يقرأ له القرآن الكريم في كل يوم. بدأ في تعلم اللغة العربية في سن مبكرة، وكان والده يقرأ له القرآن الكريم في كل يوم.

46

أرشد أسامة من أصول الأردنية، يعيش مع أهل أبيه في شمال غرب سوريا. تربى في بيئة دينية إسلامية، حيث كان والده يقرأ القرآن الكريم في كل يوم. بدأ في تعلم اللغة العربية في سن مبكرة، وكان والده يقرأ له القرآن الكريم في كل يوم. بدأ في تعلم اللغة العربية في سن مبكرة، وكان والده يقرأ له القرآن الكريم في كل يوم.

5

أرشد أسامة من أصول الأردنية، يعيش مع أهل أبيه في شمال غرب سوريا. تربى في بيئة دينية إسلامية، حيث كان والده يقرأ القرآن الكريم في كل يوم. بدأ في تعلم اللغة العربية في سن مبكرة، وكان والده يقرأ له القرآن الكريم في كل يوم. بدأ في تعلم اللغة العربية في سن مبكرة، وكان والده يقرأ له القرآن الكريم في كل يوم.

12

5

Ranked left-to-right, top to bottom

Woodard, J. Computer Vision for Forensic Applications, February Fourier Talks, Norbert Wiener Center, Department of Mathematics, University of Maryland, Feb. 17-18, 2011



Results: Dutch Uppercase Database

250 Subjects
 2 documents per subject
 500 total documents
 Style: Mixed Printed and Cursive
 Gray Scale
 Text Independent

Strict leave one out (SLOO) Cross-validation

MITRE³

Approach	Cross-Validation	Features	Outside VQ Codebook	Codebook size	Rank-1 %Error	TIME (hrs)
Spatial Pyramid	SLOO	Hessian-Affine	IAM ¹ partial	512	3.8	1.04
Two-level probabilistic	SLOO	Texture & Allographic	IAM ¹ partial	400	14	?

Univ. Groningen²

¹<http://www.iam.unibe.ch/fki/databases/iam-handwriting-database>.

²Brink, A., Smit, J., Bulacu, M. & Schomaker, L. (2012) Writer identification using directional ink-trace width measurements. *Pattern Recognition*, Vol. 45, pp. 162-171.

³Woodard, J., Saunders, C. & Lancaster, M. Computer Vision and Statistical Learning (on a Budget), *Defense Threat Analysis/National Science Foundation/National Geospatial Intelligence Agency Workshop on Algorithms*, San Diego, CA, November 17-22, 2012

Results: Dutch Lowercase Database

250 Subjects
 2 documents per subject
 500 total documents
 Style: Printed Uppercase
 Gray Scale
 Text Independent

Strict leave one out (SLOO) Cross-validation

	Approach	Cross-Validation	Features	Outside VQ Codebook	Codebook size	Rank-1 %Error	TIME (hrs)
MITRE ³	pLSA	10-fold cross	Hessian-Affine	IAM ¹ partial	512	11.8	15.4
	Spatial Pyramid	SLOO	Hessian-Affine	IAM ¹ partial	512	4.4	1.31
Univ. Groningen ²	Bulacu & Schomaker (2007)	SLOO	Texture & Allographic	IAM ¹ partial	400	16	?

¹<http://www.iam.unibe.ch/fki/databases/iam-handwriting-database>.

²Bulacu, M. & Schomaker, L. (2007) Text-independent writer identification and verification using textural and allographic features. *IEEE Trans. on Pattern Analysis and Machine Intelligence, Special Issue - Biometrics: Progress and Directions*, vol. 29, no. 4, pp. 701-71, April

³Woodard, J., Saunders, C. & Lancaster, M. Computer Vision and Statistical Learning (on a Budget), *Defense Threat Analysis/National Science Foundation/National Geospatial Intelligence Agency Workshop on Algorithms*, San Diego, CA, November 17-22, 2012

Summary

- **Fully automated methods handwriter recognition methods based on general computer vision methods**
- **Reasonable performance achieved on Dutch, Arabic, and other languages with little or no re-engineering**
- **Little or no human supervision required**
- **All algorithms are believed to be public domain**
- **Much more work remains!**

Backups

Querying

object overlap: probability that chosen objects in first and second images are similar: “cosine similarity”

$$\text{sim}(d_i, d_m) \approx \sum_{k=1}^K P(z_k | d_i) P(z_k | d_m)$$

$$+ \alpha \sum_{j=1}^L \left(\sum_{k=1}^K P(z_k | d_i, w_j) P(z_k | d_m, w_j) \right) \frac{n_{ij} n_{mj}}{\sum_j n_{ij} \sum_j n_{mj}}$$

VQ index sense overlap: do both VQ indexes refer to the same object?

VQ index overlap: do both images contain common indexes?

$$\text{retrieved class } d^* \equiv \arg \max_j \{ \text{sim}(d_i, d_j) \}$$