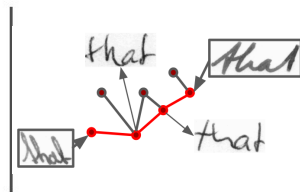


Improving Information Retrieval in Multiwriter Scenario by Exploiting the Similarity Graph of Document Terms

Pau Riba, Anjan Dutta, Sounak Dey, Josep Lladós, Alicia Fornés

Computer Vision Center, Univ. Autònoma de Barcelona

ICDAR, November 14th, 2017.



Outline

Introduction

Graph Construction

Graph of terms

Graph Analytics

Retrieval Strategies

Experimental Validation

Symbol Retrieval

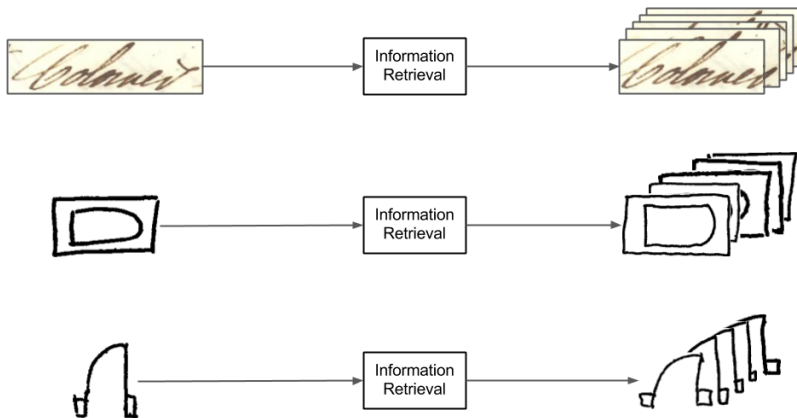
Word Retrieval

Conclusion and Future Work

Introduction

Introduction

Document retrieval

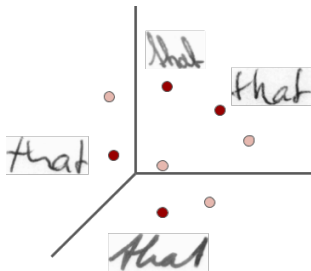


Introduction

Motivation

Classical approaches:

- ▶ Single author.
- ▶ Sensitivity to noise.

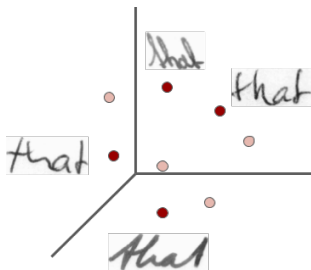


Introduction

Motivation

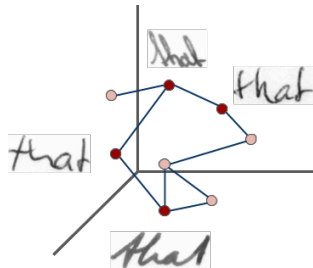
Classical approaches:

- ▶ Single author.
- ▶ Sensitivity to noise.



Contribution:

- ▶ Multi writer.
- ▶ Noise tolerance.
- ▶ New retrieval strategies.



Introduction

Motivation

- ▶ A **graph** approximates the feature space **topology**.
- ▶ **Improve** image retrieval using **graph analytics**.
- ▶ Query expansion.

Graph Construction

Attributed graph

- ▶ $G = (V, E, L_V, L_E)$.
- ▶ V is the set of nodes;
- ▶ $L_V : V \rightarrow \Sigma_V \times A_V^k$ labelling function for nodes;
 - ▶ Σ_V set of symbolic labels for vertices;
 - ▶ A_V set of attributes for vertices
 - ▶ $k \in \mathbb{N}$;
- ▶ $E \subseteq V \times V$ is the set of edges;
- ▶ $L_E : E \rightarrow \Sigma_E \times A_E^l$ labelling function for edges
 - ▶ Σ_E set of symbolic labels for edges;
 - ▶ A_E set of attributes for edges
 - ▶ $l \in \mathbb{N}$;

Graph of terms

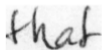
- ▶ $G_t = (V, E, L_V, L_E)$.
- ▶ V is the set of terms (*i.e. words, symbols, images...*);
- ▶ $L_V : V \rightarrow \Sigma_V \times A_V^k$ term descriptor;
 - ▶ LBP
 - ▶ SIFT + FV
 - ▶ etc.
- ▶ $E \subseteq V \times V$ term relations;
- ▶ $L_E : E \rightarrow \Sigma_E \times A_E^l$ relation label;
 - ▶ Attribute distance
 - ▶ Attribute similarity
 - ▶ etc.

Overview

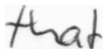
Overview



that



that



that



that

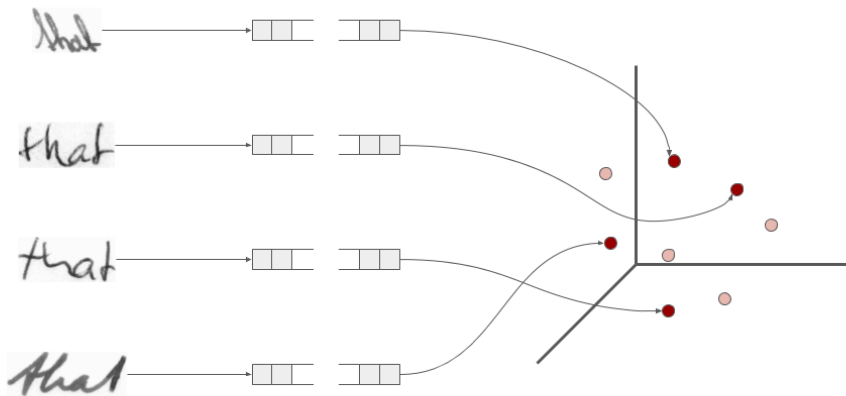
Overview



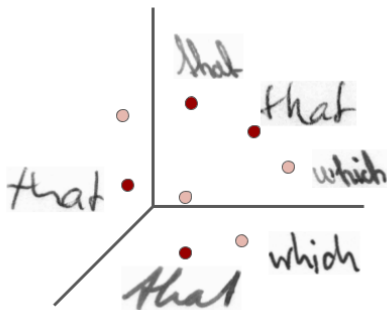
Attribute functions

- ▶ *BSM* [7]: Weighed zoning descriptor.
- ▶ *LBP* [5]: Pooling of the original LBP in the adaptive regions.
- ▶ *SIFT* [10]: Regular grid filtered by high gradient response. Fisher Vectors.
- ▶ *PHOC* [2]: Joint embedding for word images and strings.

Overview



Overview



Overview

Edge construction

- ▶ ϵ -threshold Graph: Edge if and only if distance is smaller than ϵ .
- ▶ k -NN Graph: Each node is connected with its k nearest neighbours
 - ▶ Directed
 - ▶ Undirected



Affinity graph

- ▶ Distance \leftrightarrow Similarity

$$A(v_i, v_j) = e^{-\frac{d^2}{2\sigma}}$$

Graph Analytics

Graph Analytics

Shortest Path:

- ▶ Path minimising the sum of edge weights
- ▶ New distances following the graph

Graph Analytics

Shortest Path:

- ▶ Path minimising the sum of edge weights
- ▶ New distances following the graph

Graph Diffusion [18]:

- ▶ Propagates affinity information
- ▶ Geodesic paths

$$Q^{(t+1)} = EQ^{(t)}E^T + I$$

[18] X. Yang, L. Prasad, and L. Latecki, "Affinity learning with diffusion on tensor product graph," IEEE TPAMI, 2013

Retrieval Strategies

Retrieval Strategies

Single query:

- ▶ Unseen in train data
- ▶ Nearest neighbour
- ▶ Distance follow topology

Retrieval Strategies

Single query:

- ▶ Unseen in train data
- ▶ Nearest neighbour
- ▶ Distance follow topology

Double query:

- ▶ Different authors
- ▶ Both unseen in train data
- ▶ Distances are combined

Retrieval Strategies

Single query:

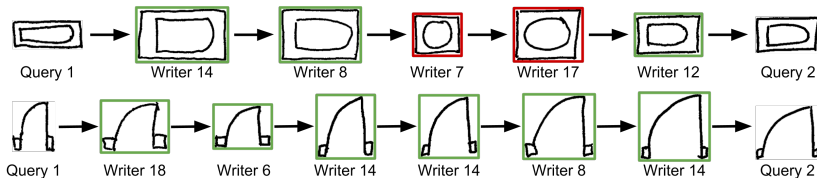
- ▶ Unseen in train data
- ▶ Nearest neighbour
- ▶ Distance follow topology

Double query:

- ▶ Different authors
- ▶ Both unseen in train data
- ▶ Distances are combined

Query expansion:

- ▶ Shortest path

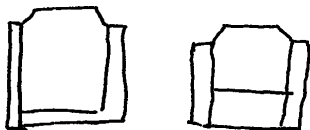


Experimental Validation

Datasets

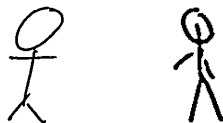
Architectural Symbol (GREC [14])

- ▶ Hand drawn graphical symbols
- ▶ 50 classes
- ▶ 22 writers
- ▶ **5000 instances**



NicIcon [17]

- ▶ Hand drawn graphical symbols
- ▶ 14 classes
- ▶ 34 writers
- ▶ **26163 instances**
- ▶ Writer independent



Symbol Retrieval

Single query

Table: MAP and Std. Shortest Path (SP) and Diffusion (Diff). BSM descriptor.

Dataset:	Architectural Symbols			Nicolcon			
	Graph:	4-NN	8-NN	16-NN	4-NN	8-NN	16-NN
Baseline		68.37±7.92			47.06±3.78		
SP		89.39 ±3.23	88.07 ±3.89	85.32 ±5.00	71.18 ±2.43	69.38 ±2.69	67.06 ±2.96
Diff		93.61 ± 2.24	93.31 ±2.26	91.24 ±2.95	84.98 ± 2.81	84.32 ±2.95	82.96 ±2.92

Symbol Retrieval

Double query

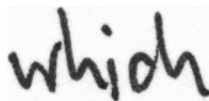
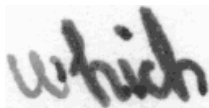
Table: MAP and Std. Shortest Path (SP) and Diffusion (Diff). Query Expansion (QE). BSM descriptor.

Dataset:	Architectural Symbols			Nicolcon		
Graph:	4-NN	8-NN	16-NN	4-NN	8-NN	16-NN
Baseline	77.16±8.81			55.35±2.16		
SP	90.93 ±3.71	90.24 ±4.26	88.38 ±5.23	73.81 ±1.41	72.65 ±1.59	70.78 ±1.71
Diff	95.21 ± 1.73	95.11 ±1.85	93.42 ±2.70	86.38 ± 1.40	85.88 ±1.40	84.88 ±1.41
SP+QE	90.80 ±3.73	90.34 ±4.15	88.69 ±5.00	74.51 ±1.36	73.50 ±1.44	71.72 ±1.63
Diff+QE	94.69 ±2.00	94.96 ±1.91	93.45 ±2.68	85.97 ±1.42	85.70 ±1.44	84.90 ±1.45

Dataset

IAM Handwriting [11]

- ▶ Handwritten English text
- ▶ Select:
 - ▶ 15 classes
 - ▶ 20 authors



Word Retrieval

Single query

Table: MAP and Std. Shortest Path (SP) and Diffusion (Diff). IAM dataset.

Descriptor:	LBP			SIFT + FV			PHOC		
	4-NN	8-NN	16-NN	4-NN	8-NN	16-NN	4-NN	8-NN	16-NN
Baseline	29.92±4.06			89.05±3.88			91.28±3.62		
SP	30.12 ±4.11	29.79 ±3.90	29.88 ±3.92	93.41 ±2.39	94.08 ±2.46	93.95 ±2.44	90.88 ±3.37	93.72 ±2.94	95.27 ±2.67
Diff	28.62 ±3.78	28.25 ±3.71	27.32 ±3.31	95.45 ±2.03	96.57 ±2.08	96.61 ±2.03	92.58 ±3.18	95.69 ±2.91	96.93 ±2.41

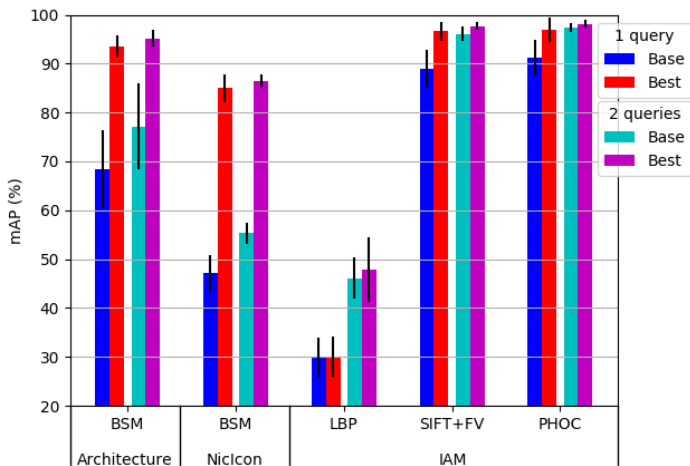
Word Retrieval

Double query

Table: MAP and Std. Shortest Path (SP) and Diffusion (Diff). Query Expansion (QE). IAM dataset.

Descriptor:	LBP			SIFT + FV			PHOC		
Graph:	4-NN	8-NN	16-NN	4-NN	8-NN	16-NN	4-NN	8-NN	16-NN
Baseline	46.08±4.18			96.07±1.45			97.46±0.92		
SP	46.85 ±5.18	45.66 ±4.67	45.12 ±4.33	96.89 ±1.09	97.33 ±0.97	97.42 ±0.90	96.04 ±1.26	97.21 ±0.90	97.88 ±0.85
Diff	44.36 ±4.40	43.90 ±4.01	42.58 ±3.90	96.74 ±1.12	97.59 ±1.02	97.73 ±0.84	95.51 ±1.28	97.54 ±0.88	98.17 ±0.77
SP+QE	47.92 ±6.63	45.60 ±5.34	45.85 ±5.33	96.83 ±1.02	97.32 ±0.93	97.45 ±0.83	95.61 ±1.61	97.02 ±1.12	97.84 ±0.99
Diff+QE	46.05 ±5.89	44.77 ±4.77	43.79 ±4.84	96.61 ±1.26	97.53 ±1.08	97.70 ±0.89	95.23 ±1.66	97.34 ±1.11	98.12 ±0.84

Summary



Conclusion and Future Work

Final thoughts

Conclusions

- ▶ Information retrieval \uparrow
- ▶ Feature space approximated by a graph
- ▶ Distance in graph \longleftrightarrow Distance in features
- ▶ One query \longleftrightarrow Two queries
- ▶ Shortest path as Query Expansion

Future Work

- ▶ More authors
- ▶ Less examples

Thank you for your attention!

Pau Riba
Computer Vision Center
priba@cvc.uab.cat
www.cvc.uab.cat/people/priba



