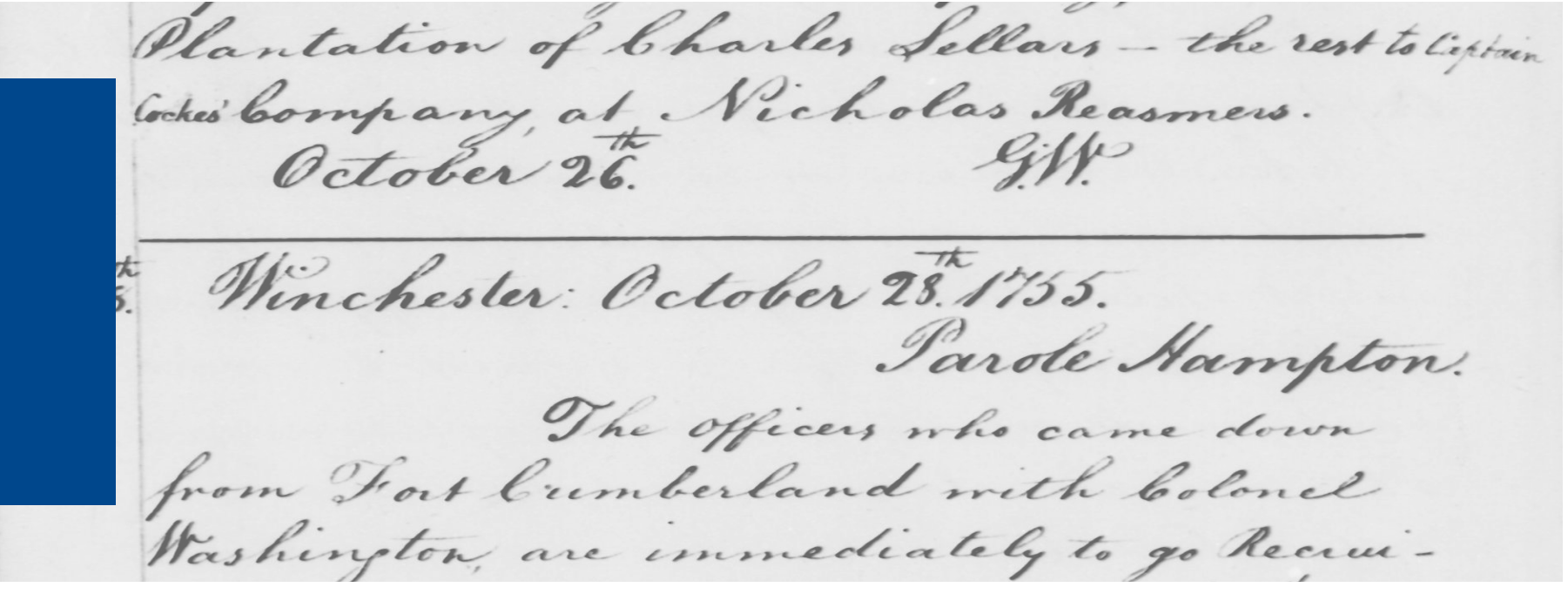
CAIP 2017

Speeding-Up Graph-based Keyword Spotting by Quadtree Segmentations

Michael Stauffer, Andreas Fischer, and Kaspar Riesen


Plantation of Charles Sellars — the rest to Captain
Cockes' Company, at Nicholas Reasmers.
October 26th G.W.

Winchester: October 28th 1755.
Parole Hampton.
The officers who came down
from Fort Cumberland with Colonel
Washington, are immediately to go Recrui-

Content

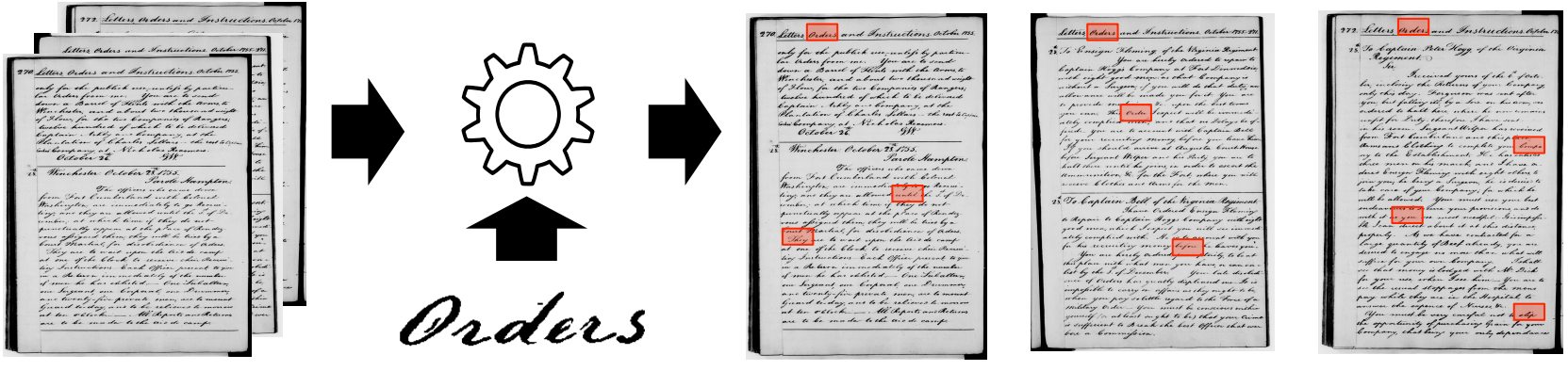
- What is Keyword Spotting
- Graph-based Keyword Spotting
  - Image Preprocessing
  - Graph Representation
  - Quadtree Graph Matching
- Experiments
- Conclusion + Future Work
- Q+A

What is Keyword Spotting

Keyword Spotting (KWS) is the task of **retrieving any instance** of a given **query** word in **speech recordings** or **text images**.

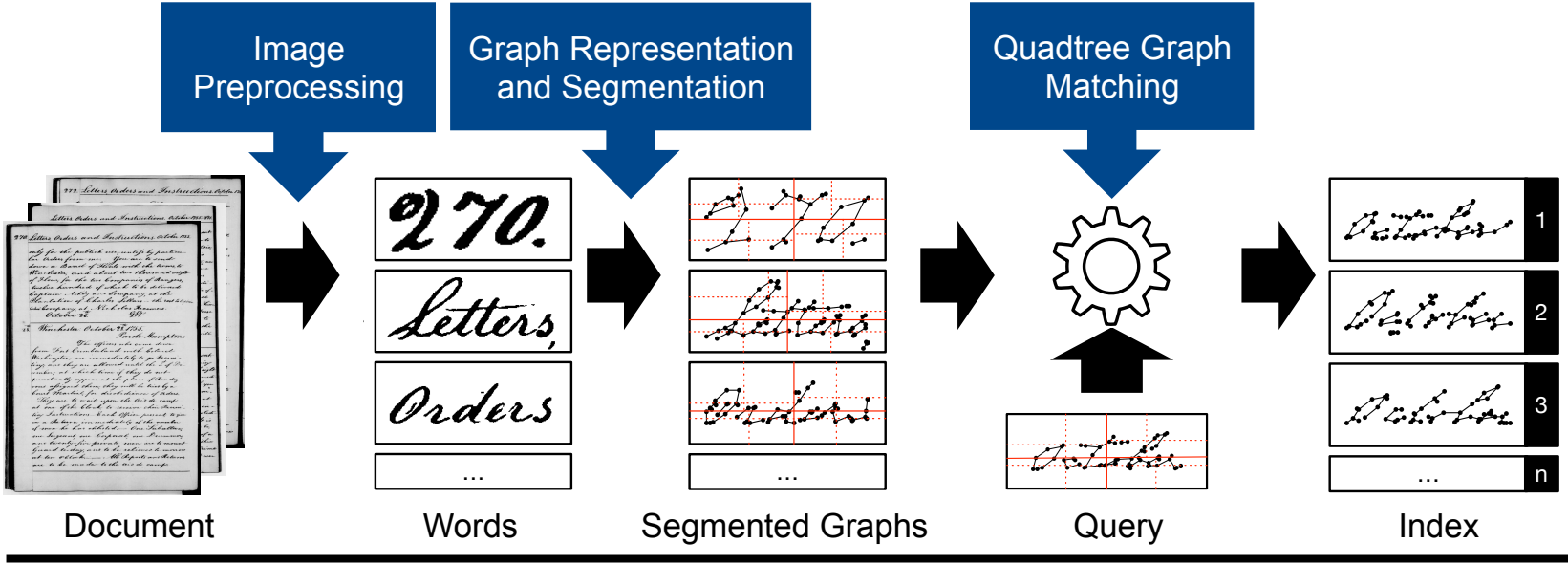
➡ Focus on handwritten, historical documents

Document | Query | Word Spotting



Graph-based Keyword Spotting – Overview

Graph-based KWS is based on the **representation of words** by means of different quadtree segmented **graphs**. This representations are eventually used to **retrieve a keyword** by **matching a query** graph **with** all **document graphs**.

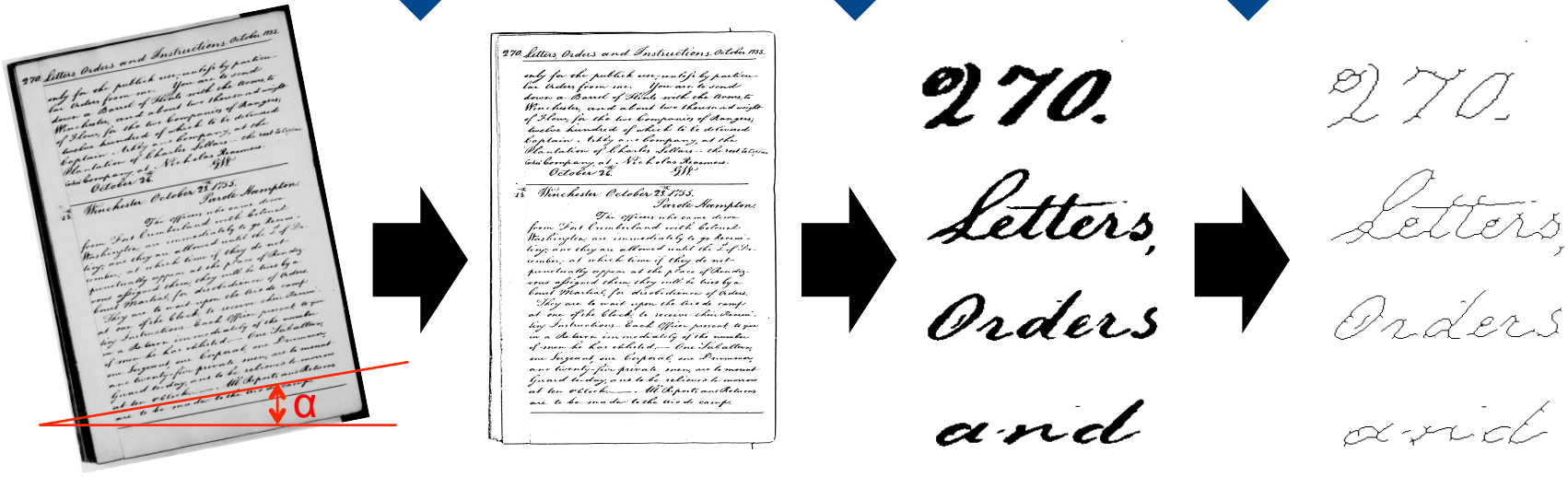


Graph-based Keyword Spotting – Image Preprocessing

Binarisation / Difference of Gaussian / Skew Correction (α)

Word Segmentation

Skeletonisation (optional)



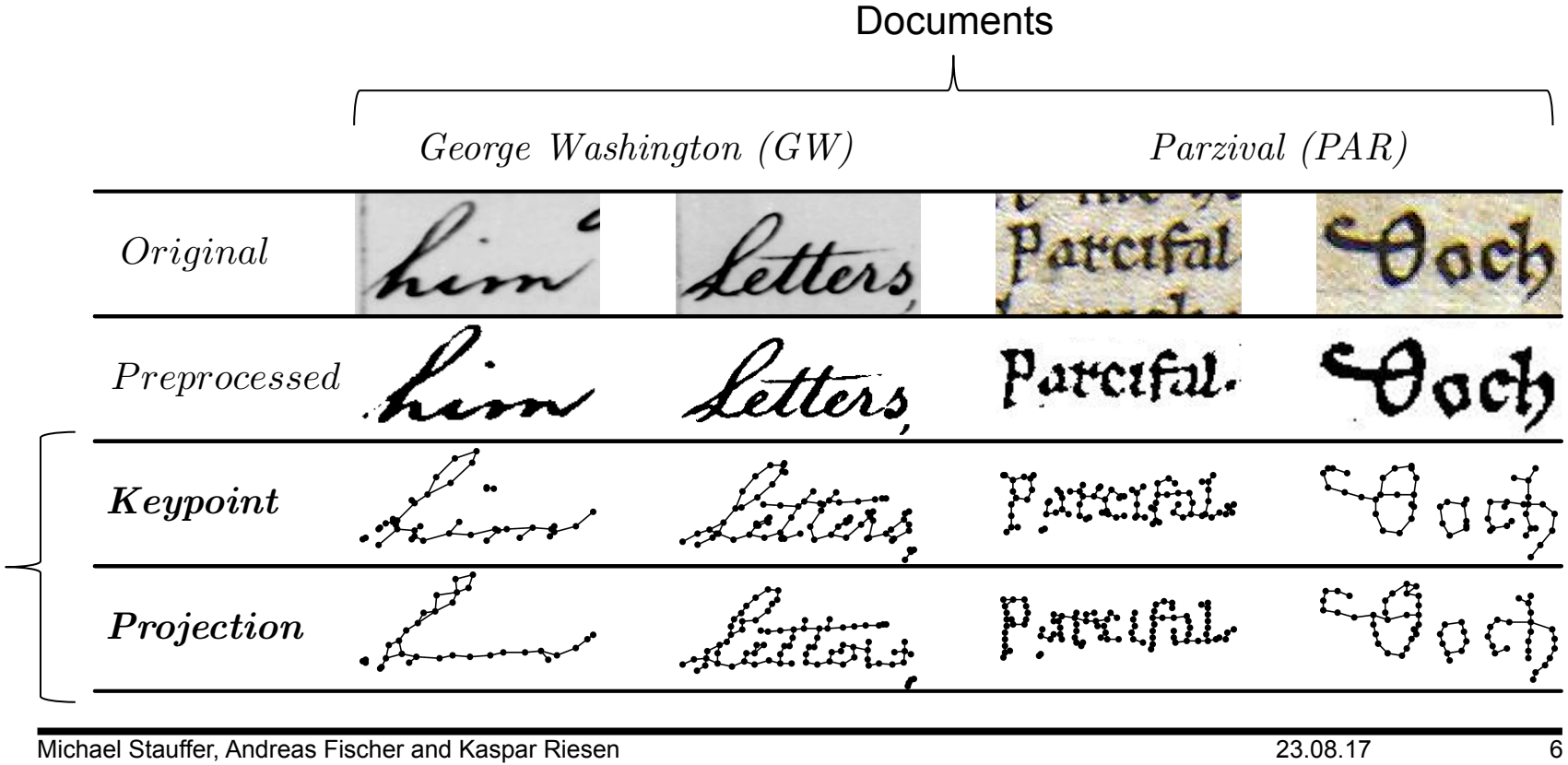
Document

Preprocessed Document

Words

Skeletonised Words

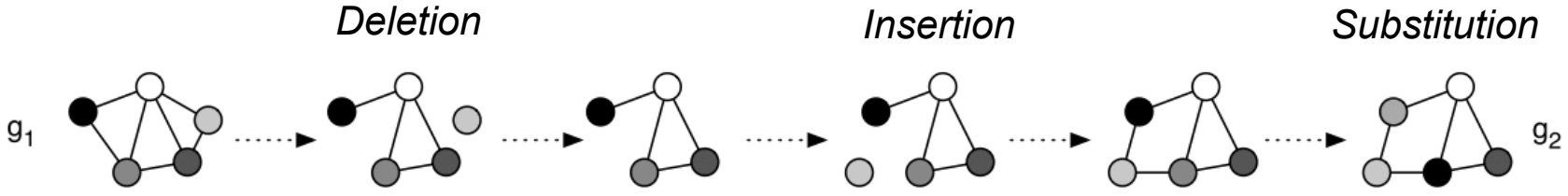
Graph-based Keyword Spotting – Graph Representation



Graph-based Keyword Spotting – Quadtree Graph Matching

The actual KWS is based on **matching** a **query graph** q **with** a set of **document graphs** G = $\{g_{1,\dots,}g_N\}$ by means of **Bipartite Graph Edit Distance (BP)**.

BP **approximates** the **minimum amount of distortion** that is needed to **transform** graph $\mathbf{g_1}$ **into** graph $\mathbf{g_2}$.

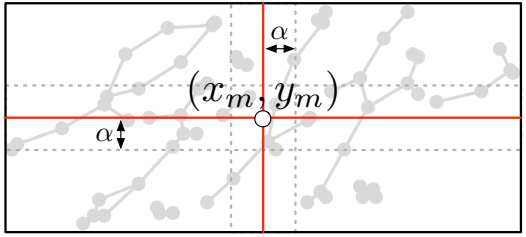


We have **q x |G| matchings** with **cubic** time complexity with respect to the **number of** involved **nodes**.

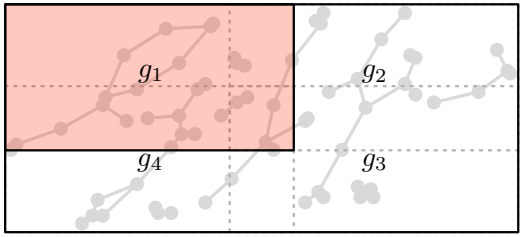
Graph-based Keyword Spotting – Quadtree Graph Matching (Segmentation)

The graph matching procedure can be speed up by **matching smaller subgraphs** rather than complete graphs. Thus, we first **segment graphs** by means of a **quadtree procedure**.

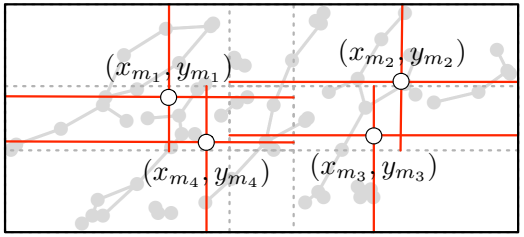
Recursion Level l = 1 with Centre of Mass (x_m,y_m) and Overlap Factor α



Recursion Level l = 1 with Subgraphs g1, g2, g3, and g4



Recursion Level l = 2 with Centres of Mass (x_{m1},y_{m1}), ... , (x_{m4},y_{m4})



...

Graph-based Keyword Spotting – Quadtree Graph Matching (Algorithm)

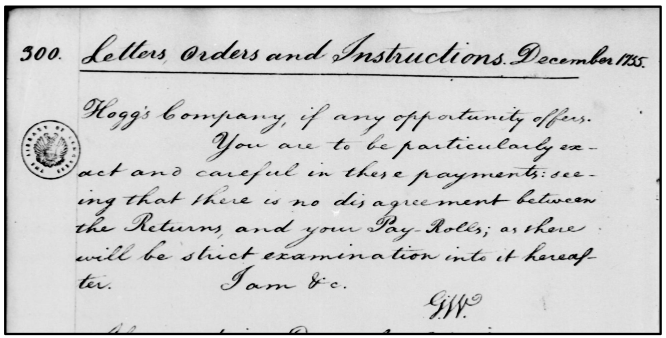
Algorithm 1 Quadtree Graph Matching

Input: Graphs g and g', overlap factor α, maximum recursion depth $r > 0$
Output: Graph distance $d_{\mathrm{BP_Q}}$ between graph g and g'

1: **function** BPQ(l, g, g')
2: Quadtree segment g and g' to g_1, g_2, g_3, g_4 and g'_1, g'_2, g'_3, g'_4
3: $d_{\mathrm{BP_Q}} = \sum_{i=1}^{4} d_{\mathrm{BP}(g_i, g'_i)}$
4: **if** l equal r **then**
5: **return** $d_{\mathrm{BP_Q}}$
6: **return** $d_{\mathrm{BP_Q}} + \left(\sum_{i=1}^{4} \mathrm{BPQ}(l+1, g_i, g'_i) \right)$

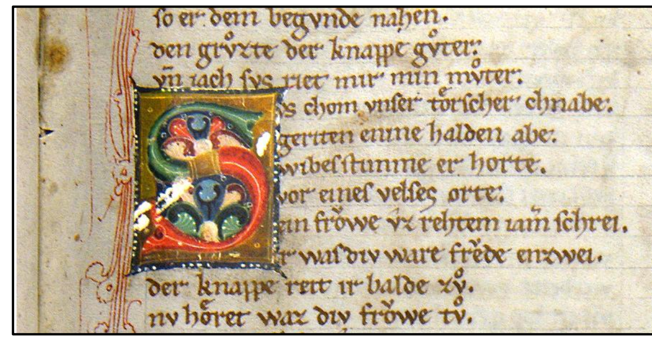
Experiments – Setup

Quality is measured by **Mean Average Precision (MAP)** for local thresholds and **Average Precision (AP)** for global thresholds.


300. Letters, Orders and Instructions. December 1755.

Hogg's Company, if any opportunity offers.
You are to be particularly exact and careful in these payments: seeing that there is no disagreement between the Returns, and your Pay-Rolls; as there will be strict examination into it hereafter.
I am &c.
GW.

George Washington (GW)


so er dem begunde nahen.
den grvzte der knappe gvter;
vn iach sus riet mir min mvter.
s chom vnser torscher chnabe.
geriten eine halden abe.
wibes stimme er horte.
vor eines velses orte.
ein frowe vz rehtem iamer schrei.
r was div ware frede enzwei.
der knappe reit ir balde zv.
nv horet waz div frowe tv.

Parzival (PAR)

Recall = True Positives / (True Positives + False Negatives)

Precision = True Positives / (True Positives + False Positives)

AP = Area under the curve of the Recall-Precision curve

MAP = Average area under the curve of Recall-Precision curves

Experiments – Validation

Optimisation of maximum **recursion depth r** = {1, 2, 3, 4, 5}
in combination with **overlap factor α** = {0.00, 0.01, ... , 0.20}.

	GW				PAR			
	Keypoint		Projection		Keypoint		Projection	
r	α	MAP	α	MAP	α	MAP	α	MAP
1	**0.01**	**81.41**	**0.02**	**77.94**	**0.01**	**90.03**	**0.01**	**89.58**
2	0.00	81.25	0.04	77.16	0.00	89.23	0.01	89.19
3	0.00	80.70	0.04	73.91	0.00	88.71	0.04	87.62
4	0.00	78.44	0.00	70.67	0.00	87.99	0.00	86.96
5	0.00	77.03	0.00	69.67	0.00	87.33	0.00	85.88

Experiments – Testing

KWS with original BP vs. KWS with Fast Rejection BP-FR[1] and KWS with Quadtree Segmentations BP-Q.

		GW					PAR				
	Method	MAP	±	AP	±	SF	MAP	±	AP	±	SF
BP	Keypoint	66.08		55.22			62.04		60.76		
	Projection	61.43		49.34			66.23		62.38		
BP-FR	Keypoint	68.81	+4.1	54.10	−2.0	3.2	67.70	+9.1	63.01	+3.7	2.4
	Projection	64.65	+5.2	48.94	−0.8	2.6	72.02	+8.7	63.49	+1.8	2.3
BP-Q	Keypoint	65.92	−0.2	54.91	−0.6	17.1	56.83	−8.4	54.66	−10.0	21.2
	Projection	59.57	−3.0	48.13	−2.5	15.0	64.62	−2.4	61.72	−1.1	21.5

Speed-up factors of **15 to 21** with marginal affection of KWS accuracy.

[1] M. Stauffer, A. Fischer, K. Riesen, Speeding-Up Graph-based Keyword Spotting in Historical Handwritten Documents, in: Graph-Based Represent. Pattern Recognit., 2017.

Conclusion + Future Work

Conclusion

- Graph-based keyword spotting for historical documents
- Quadtree segmentation and matching for graphs
- Speed-up factors of 15 to 21
- KWS accuracy is only marginally affected

Future Work

- Implement weighting of recursion level (i.e. higher level = more relevant)
- Combine fast-rejection with quadtree segmentations
- Consider further graph matching algorithms (e.g. Hausdorff Edit Distance)

Q+A

?!