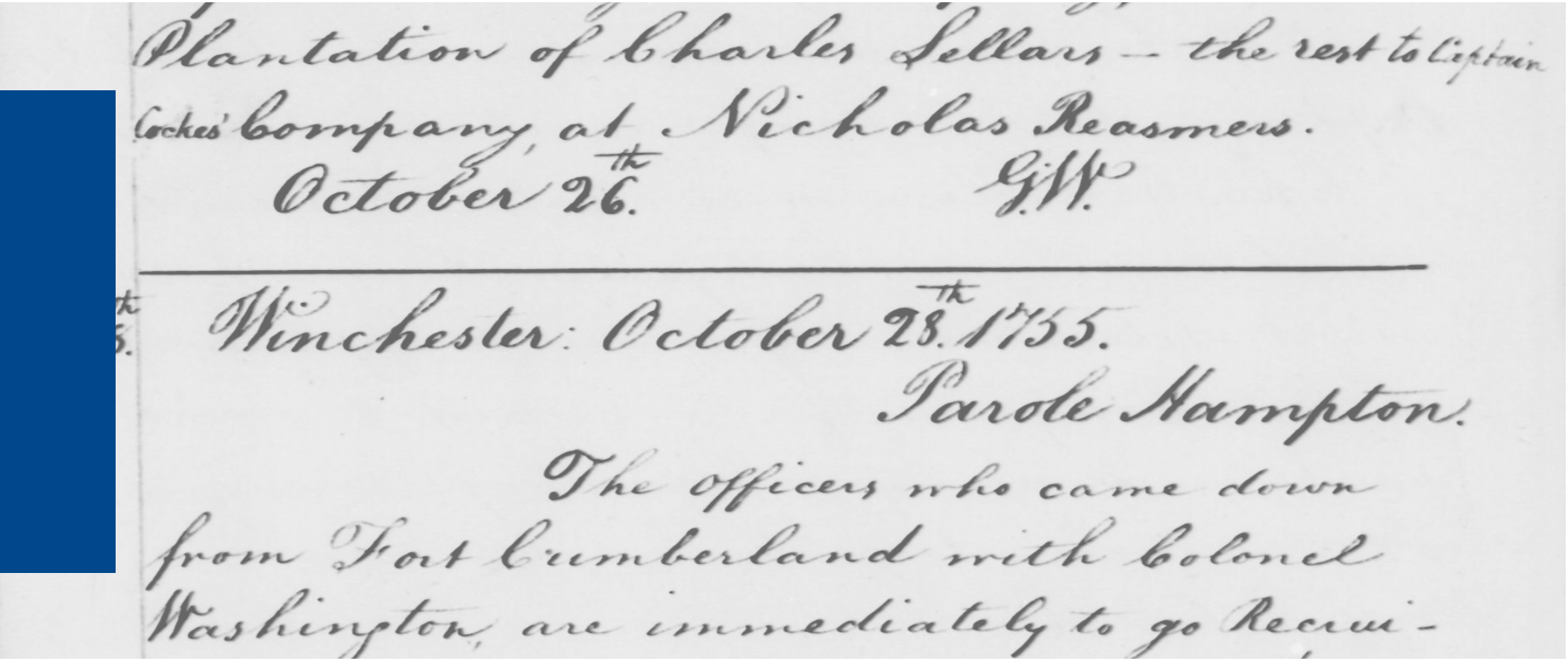
GBR 2017

Speeding-Up Graph-based Keyword Spotting in Historical Handwritten Documents

Michael Stauffer, Andreas Fischer, and Kaspar Riesen


Plantation of Charles Sellars — the rest to Captain
Cocke's 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 Recru-

Content

- What is Keyword Spotting
- Graph-based Keyword Spotting
  - Image Preprocessing
  - Graph Representation
  - Fast Rejection Method
- Experiments
- Conclusion + Future Work
- Q+A

What is Keyword Spotting

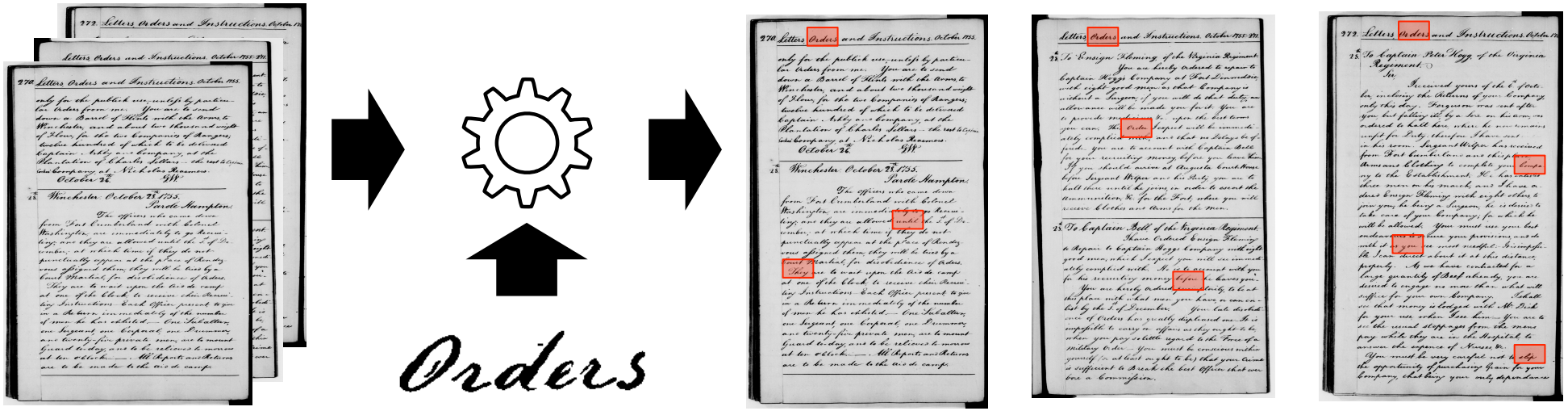
Keyword Spotting (KWS) is the task of **retrieving any instance** of a given **query** word in **speech recordings** or **text images**.

➡ Focus on handwritten, historical documents

Document

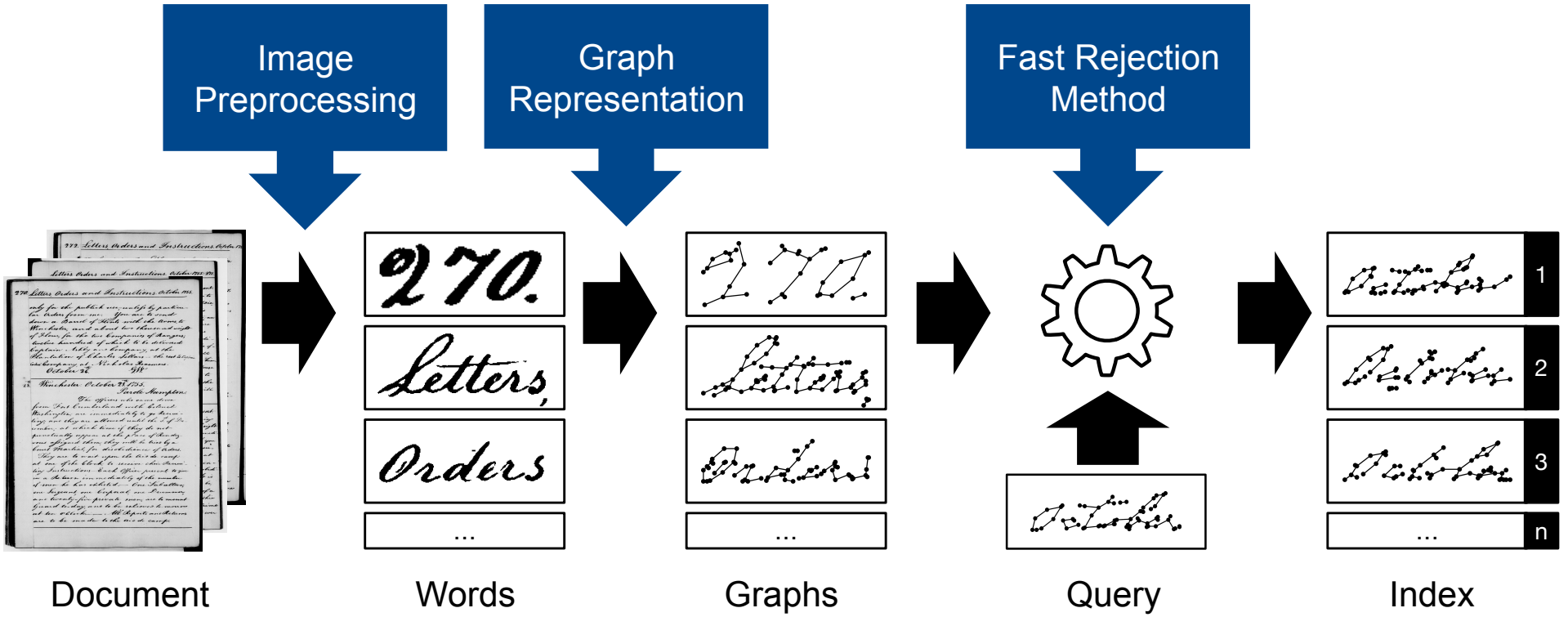
Query

Word Spotting



Graph-based Keyword Spotting – Overview

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

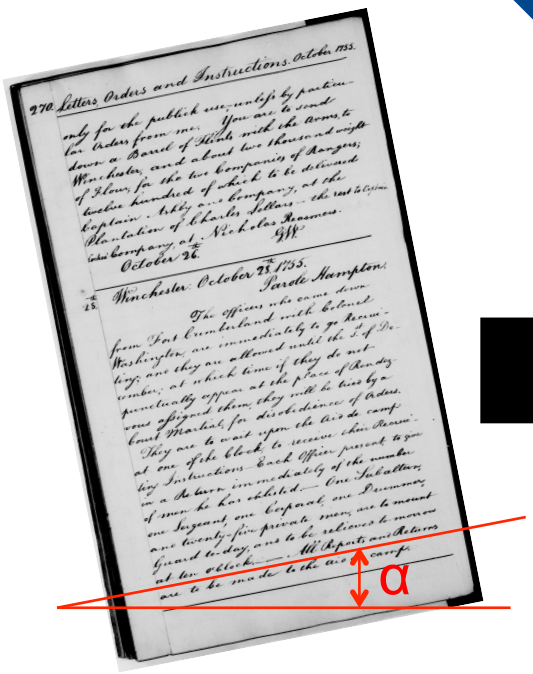


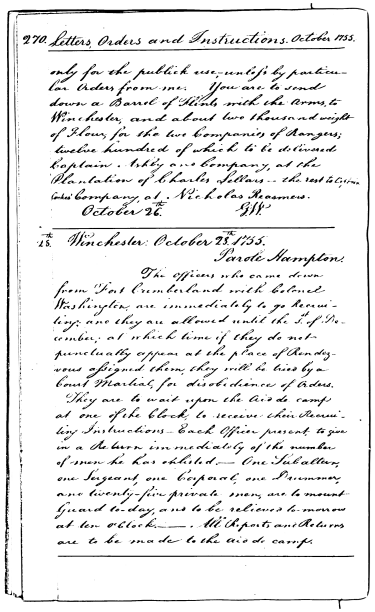
Graph-based Keyword Spotting – Image Preprocessing

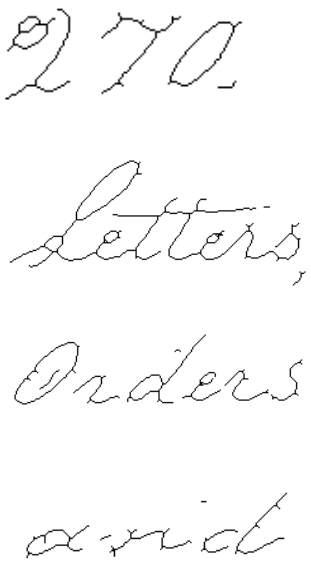
Binarisation / Difference of Gaussian / Skew Correction (α)

Word Segmentation

Skeletonisation (optional)







Document

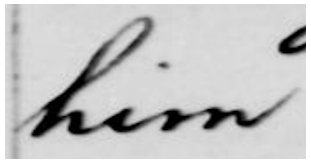
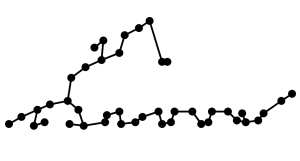
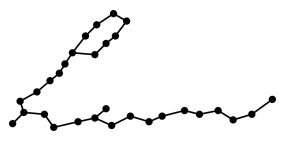
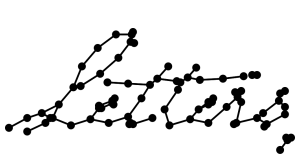
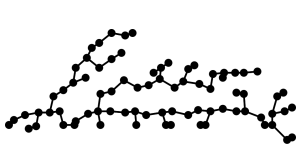
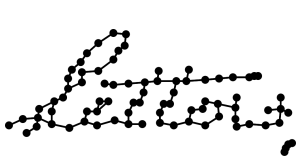
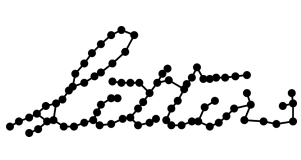
Preprocessed Document

Words

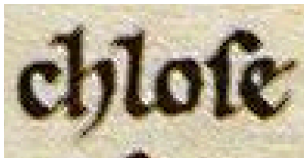
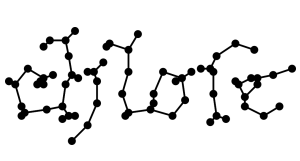
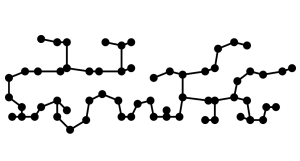
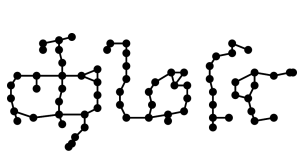
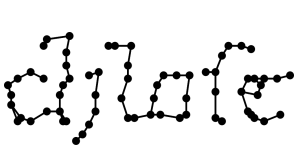
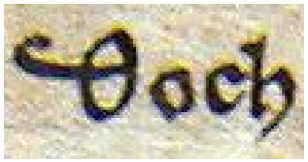
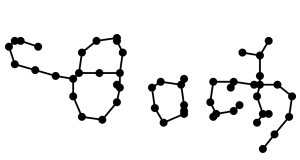
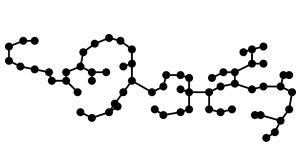
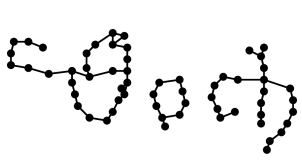
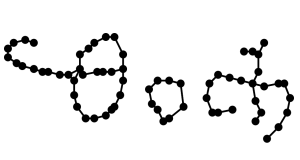
Skeletonised Words

Graph-based Keyword Spotting – Graph Representation

George Washington (GW)

Original	Preprocessed	Keypoint	Grid	Projection	Split
					
					

Parzival (PAR)

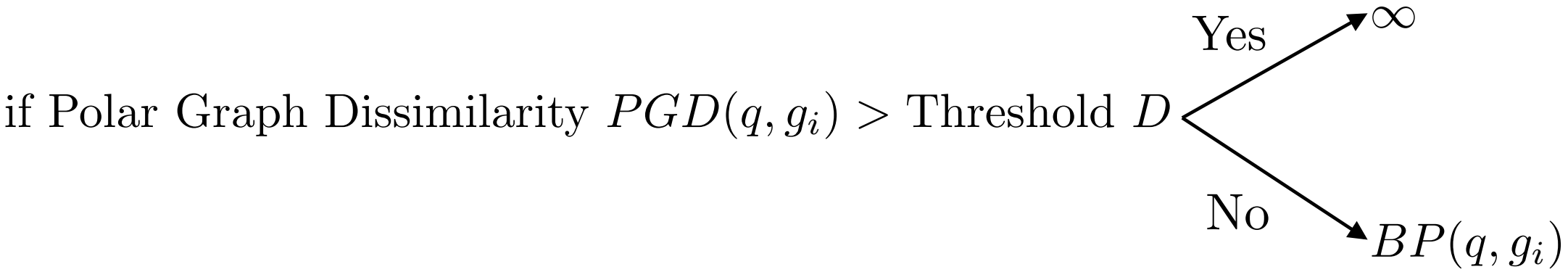
Original	Preprocessed	Keypoint	Grid	Projection	Split
					
					

Graph-based Keyword Spotting – Fast Rejection Method

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

➡ q x |G| matchings with cubic time complexity

Fast rejection = Filtering graphs with high dissimilarity and thus speeding up the KWS procedure without negatively affecting the retrieval accuracy.

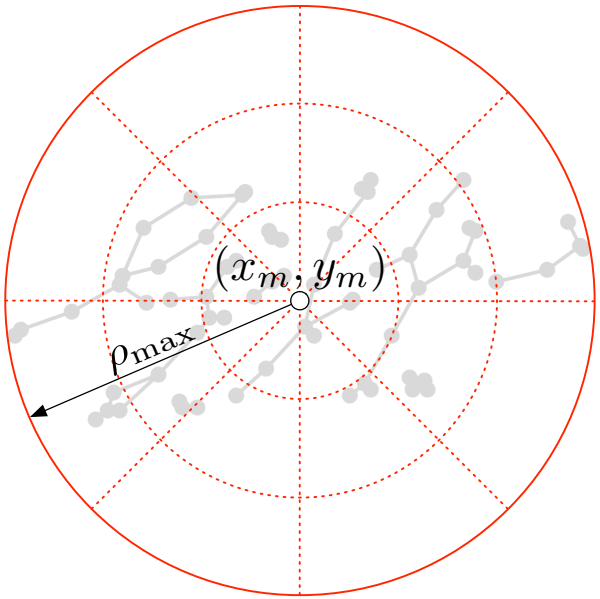


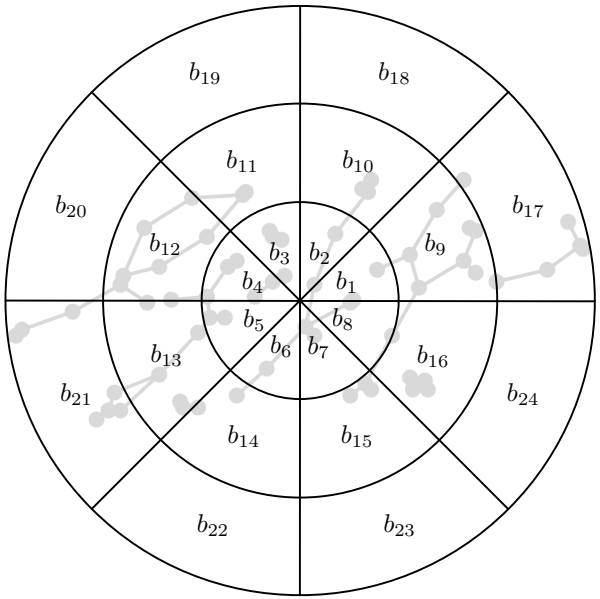
Graph-based Keyword Spotting – Fast Rejection (Construction of PGD)

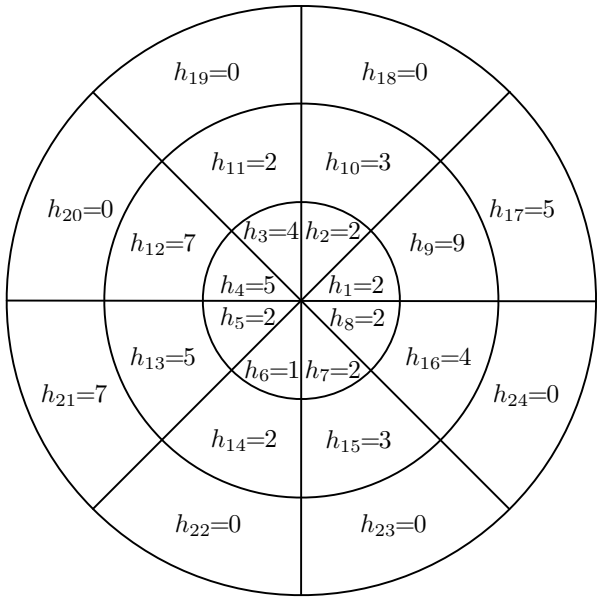
Define centre of mass (x_m,y_m) and radius ρ_{max}

Polar graph segmentation based on u_{max} and v_{max}, i.e. 3 and 8 (24 bins)

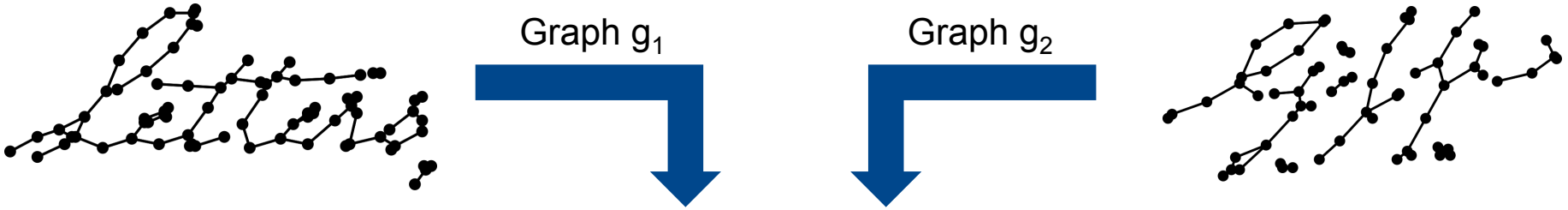
Create histogram H={h_1,…,h_n} by counting the number of nodes per segment h_i







Graph-based Keyword Spotting – Fast Rejection (Computation of PGD)



Algorithm 1 Polar Graph Dissimilarity (PGD)

Input: Graphs g_1 and g_2, recursion depth r
Output: Polar graph dissimilarity between graph g_1 and g_2

1: **function** PGD(l, g_1, g_2)
2: Create histogram H_1 based on g_1, and histogram H_2 based on g_2
3: Calculate χ^2-distance $d(H_1, H_2)$
4: **if** l equal r **then**
5: **return** d
6: Segment g_1 and g_2 based on quadtree to g_{1_1}, g_{1_2}, g_{1_3}, g_{1_4} and g_{2_1}, g_{2_2}, g_{2_3}, g_{2_4}
7: **return** $(\sum_{i=1}^{4} \mathrm{PGD}(l+1, g_{1_i}, g_{2_i})) + d$

$PGD(g_1, g_2) = 743$ actually used in $d(q, g_i) \begin{cases} \infty, & \text{if } PGD(q, g_i) > D \\ BP(q, g_i), & \text{otherwise} \end{cases}$

Experiments – Setup

KWS experiment is based on two datasets

- **George Washington**
- **Parzival**

Quality is measured by **Average Precision (AP)** for global thresholds, **mean Average Precision (mAP)** for local thresholds, and **Filter Rate (FR)**.

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

FR = Relative amount of pairwise matchings that is filtered

Experiments – Results KWS (mAP) / Validation of Polar Graph Dissimilarity (PGD)

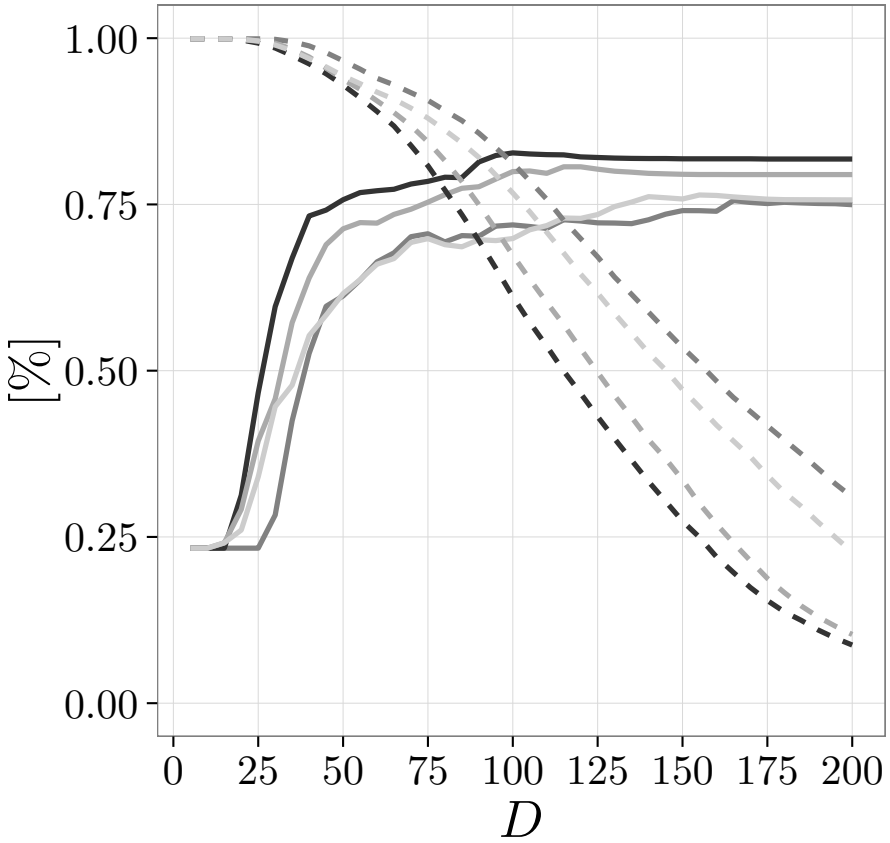
Optimal u_{max} and v_{max} for recursion level l = 1 and 2

We optimised the MAP for $\mathbf{u_{max}}$ = {1, 2, 3, 4, 5, 6} and $\mathbf{v_{max}}$ = {4, 8, 12, 16, 20, 24, 28, 32, 36, 40} for two recursion levels **l** = {1, 2}.

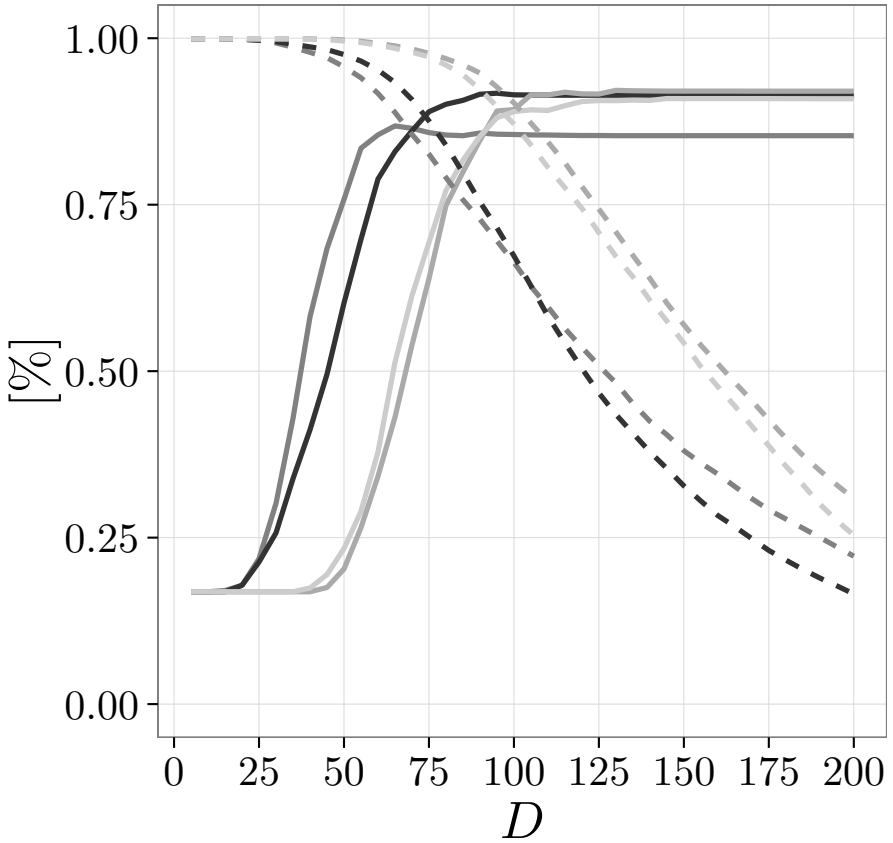
	GW				PAR			
	$l=1$		$l=2$		$l=1$		$l=2$	
Method	$u_{\max}$	$v_{\max}$	$u_{\max}$	$v_{\max}$	$u_{\max}$	$v_{\max}$	$u_{\max}$	$v_{\max}$
`Keypoint`	4	12	1	6	3	20	2	6
`Grid`	5	24	1	4	4	20	1	6
`Projection`	5	16	1	4	3	36	3	4
`Split`	4	20	1	4	3	40	2	6

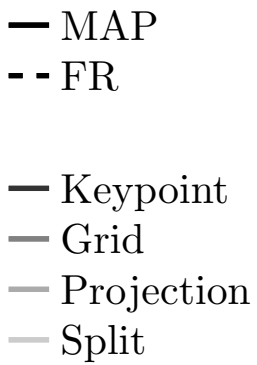
Experiments – Results KWS (mAP) / Validation of Rejection Threshold D

George Washington



Parzival





Experiments – Results KWS (mAP) / Testing GW

KWS without Fast Rejection BP vs. KWS with Fast Rejection BP-FR

	Method	MAP	±	AP	±	FR
BP	Keypoint	66.08		54.99		0.00
	Grid	60.02		46.44		0.00
	Projection	61.43		48.69		0.00
	Split	60.23		47.96		0.00
BP-FR	Keypoint	68.81	+4.12	55.68	+1.25	69.04
	Grid	62.59	+4.27	47.48	+2.23	54.65
	Projection	64.65	+5.25	50.41	+3.53	61.04
	Split	63.49	+5.41	46.95	−2.11	47.70

Experiments – Results KWS (mAP) / Testing PAR

KWS without Fast Rejection BP vs. KWS with Fast Rejection BP-FR

	Method	MAP	±	AP	±	FR
BP	Keypoint	62.04		60.74		0.00
	Grid	56.50		44.08		0.00
	Projection	66.23		60.61		0.00
	Split	59.44		55.46		0.00
BP-FR	Keypoint	67.70	+9.12	58.03	−4.46	58.72
	Grid	63.41	+12.23	38.59	−12.45	78.71
	Projection	72.02	+8.74	55.83	−7.89	58.10
	Split	65.65	+10.45	56.97	+2.72	39.24

Conclusion + Future Work

Conclusion

- Novel graph dissimilarity measure (PGD) for fast rejection of KWS matchings
- PGD reduces the amount of graph matchings by 50% or more
- KWS accuracy is not negatively affected

Future Work

- Consider not only nodes but also edges in PGD histograms
- Consider further graph matching algorithms (e.g. Hausdorff Edit Distance)

Q+A

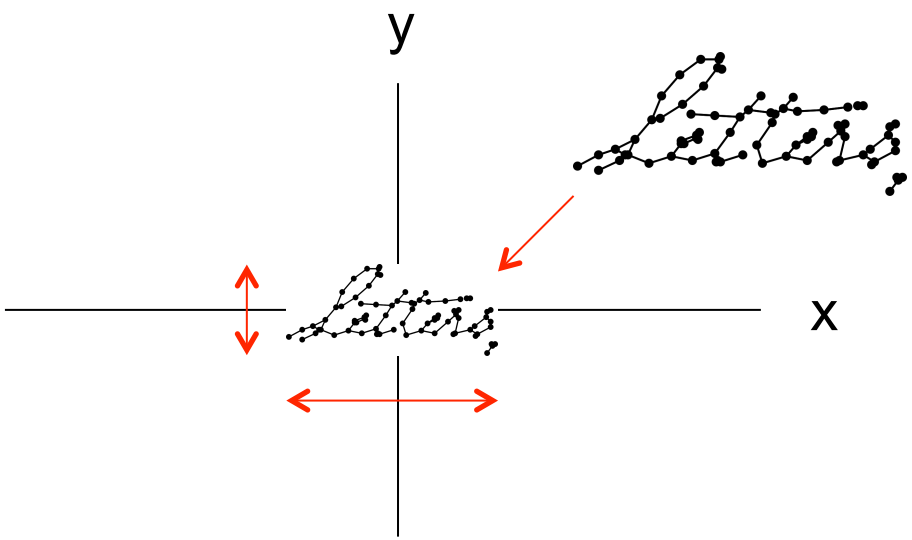
?!

Graph-based Keyword Spotting – Graph Representation

Graph Normalisation by Centering & Scaling

$x_n = (x - \mu_x) / \sigma_x$
$y_n = (y - \mu_y) / \sigma_y$

σ is standard deviation of node positions



Experiments – Setup (Number of Words per Dataset)

Validation

For both datasets (GW and PAR), the validation set consists **of 1000 different random words** including at least **10 instances** of all 10 **keywords**.

Testing

Dataset	Keywords	Train	Test
GW	105	2,447	1,224
PAR	1,217	11,468	6,869

Experiments – Results KWS (mAP) / Validation and Testing

Optimal D and corresponding Filter Rate FR

	GW			PAR		
Method	D	MAP	FR	D	MAP	FR
`Keypoint`	100	82.8	61.1	95	91.7	71.5
`Grid`	165	75.6	46.0	70	86.5	85.6
`Projection`	115	80.7	56.9	130	92.2	70.9
`Split`	155	76.4	44.6	145	90.9	57.5

KWS with PGD

	GW		PAR	
	MAP	AP	MAP	AP
BP	66.08	54.99	62.04	60.74
PGD	58.54	44.77	42.65	31.63