READ

Recognition and Enrichment of Archival Documents

D7.13
Keyword Spotting Engines: QbE, QbS P1

Ioannis Pratikakis, Konstantinos Zagoris, George Barlas, DUTH
George Retsinas, George Sfikas, Basilis Gatos, NCSR
Alejandro H. Toselli, Joan Puigcerver, Enrique Vidal, Verónica Romero, Joan A. Sánchez, UPVLC

Distribution:

http://read.transkribus.eu/

READ
H2020 Project 674943

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 674943

Project ref no.	H2020 674943
Project acronym	**READ**
Project full title	**Recognition and Enrichment of Archival Documents**
Instrument	H2020-EINFRA-2015-1
Thematic Priority	EINFRA-9-2015 - e-Infrastructures for virtual research environments (VRE)
Start date / duration	01 January 2016 / 42 Months

Distribution	Public
Contractual date of delivery	31.12.2016
Actual date of delivery	04.12.2016
Date of last update	16.12.2016
Deliverable number	D7.13
Deliverable title	Keyword Spotting Engines: QbE, QbS P1
Type	Demonstrator
Status & version	Public & Version 1
Contributing WP(s)	WP7
Responsible beneficiary	DUTH
Other contributors	NCSR, UPVLC
Internal reviewers	Gundram Leifert, URO Sofia Ares Oliveira, EPFL
Author(s)	Ioannis Pratikakis, Konstantinos Zagoris, George Barlas, DUTH George Retsinas, George Sfikas, Basilis Gatos, NCSR Alejandro H. Toselli, Joan Puigcerver, Enrique Vidal, Verónica Romero, Joan A. Sánchez, UPVLC
EC project officer	Martin Majek
Keywords	Keyword Spotting, Query by Example, Query by String

Table of Contents

Executive Summary

Handwritten keyword spotting is the task of detecting query words in handwritten document image collections without involving a traditional OCR step. Recently, handwritten word spotting has attracted the attention of the research community in the field of document image analysis and recognition since it has been proved to be a feasible solution for indexing and retrieval of handwritten documents in the case where OCR-based methods fail to deliver proper results. This deliverable reports on the achievements concerning the tasks of keyword spotting for handwritten document image collections at the end of the first year of the READ project that have been realized by three distinct frameworks which correspond to partners DUTH, NCSR, and UPVLC, respectively.

I. The Query by Example (QbE) case engine

1. Introduction

A promising strategy to deal with unindexed documents is a keyword matching procedure that relies upon a low-level pattern matching called word spotting [Manmatha1996]. In the literature, word spotting appears under two distinct strategies wherein the fundamental difference concerns the search space which could be either a set of segmented word images (segmentation-based approach) or the complete document image (segmentation-free approach). The selection of the segmentation-based strategy is preferred when the layout is simple enough to correctly segment the words while the segmentation-free strategy performs better when there is considerable degradation on the document. Nevertheless both strategies use an operational pipeline where feature extraction and matching have prominent roles.

2. DUTH Keyword Spotting framework

2.1. Previous Work

Our proposed algorithm relies on Document-oriented Local Features (DoLF) [Zagoris2017, Zagoris2014] which take into account information around representative keypoints as well as a matching process that incorporates spatial context in a local proximity search without using any training data. Finally, it introduces a distance algorithm that incorporates spatial context and is employed under both segmentation-based and segmentation-free scenarios

The main novelties of the above aproach are:

i. Use of local features that takes in consideration the handwritten documents particularities. Therefore, it is able to detect meaningful points of the characters that reside in the documents independently of its scaling.
ii. It provides consistency between different handwritten writing variations.

iii. Use of the same operational pipeline in both segmentation-based and segmentation-free scenarios
iv. Incorporation of spatial context in the local search of the matching process.

Figures 1.1 and 1.2 present the segmentation-based and segmentation-free operational pipelines, respectively.

It shows considerable effectiveness against other local features under two different word spotting scenarios: segmentation - based and segmentation – free [Zagoris2017]. It is proven that the proposed framework achieves better performance after a consistent evaluation against 4 datasets and 13 different state of the art methods under two different keyword spotting scenarios (segmentation-based and segmentation-free) [Zagoris2017]. Finally, an implementation of the proposed keyword spotting method as a recommender system to a transcription process is available at http://vc.ee.duth.gr/ws [Zagoris2015].

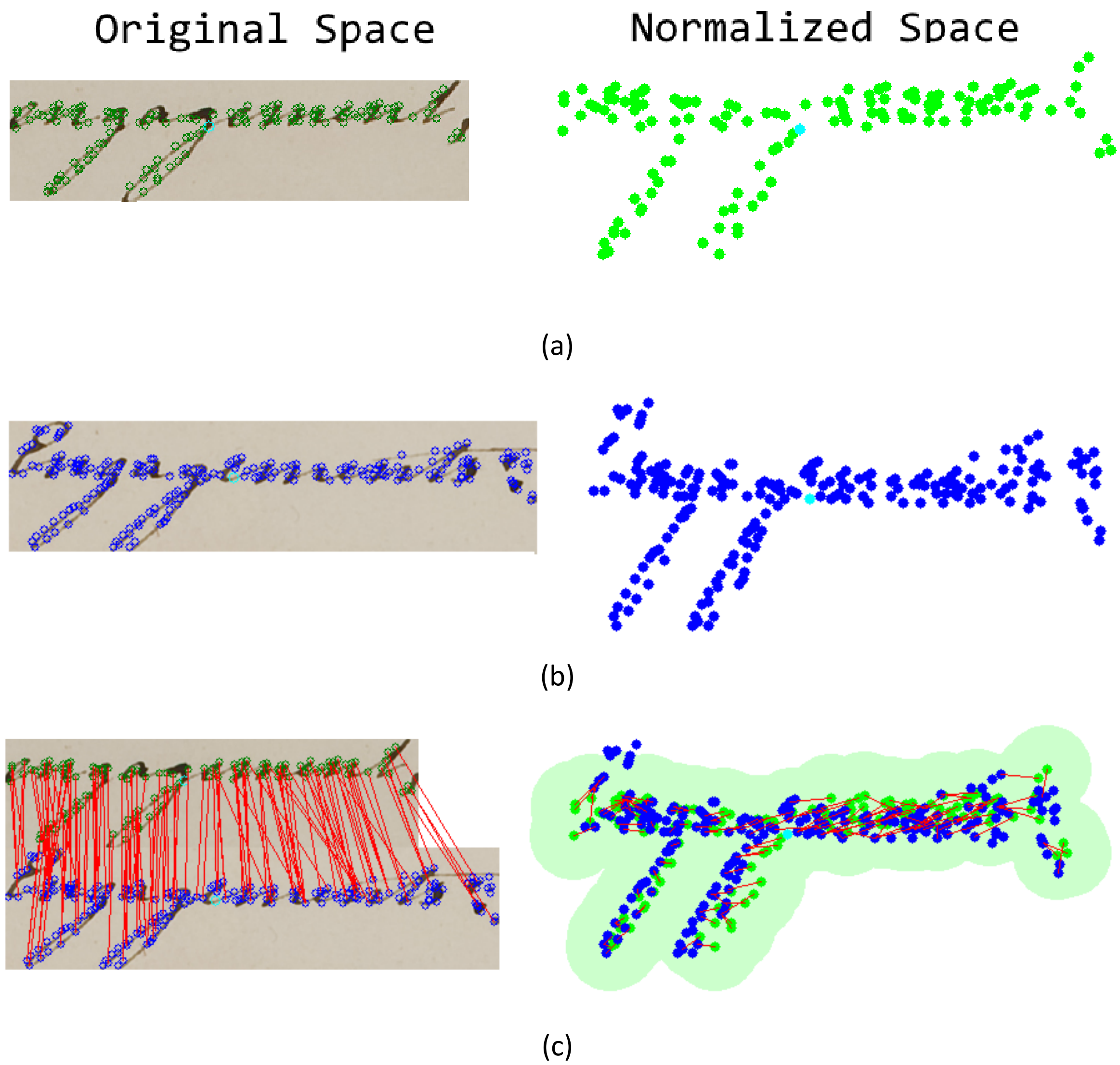
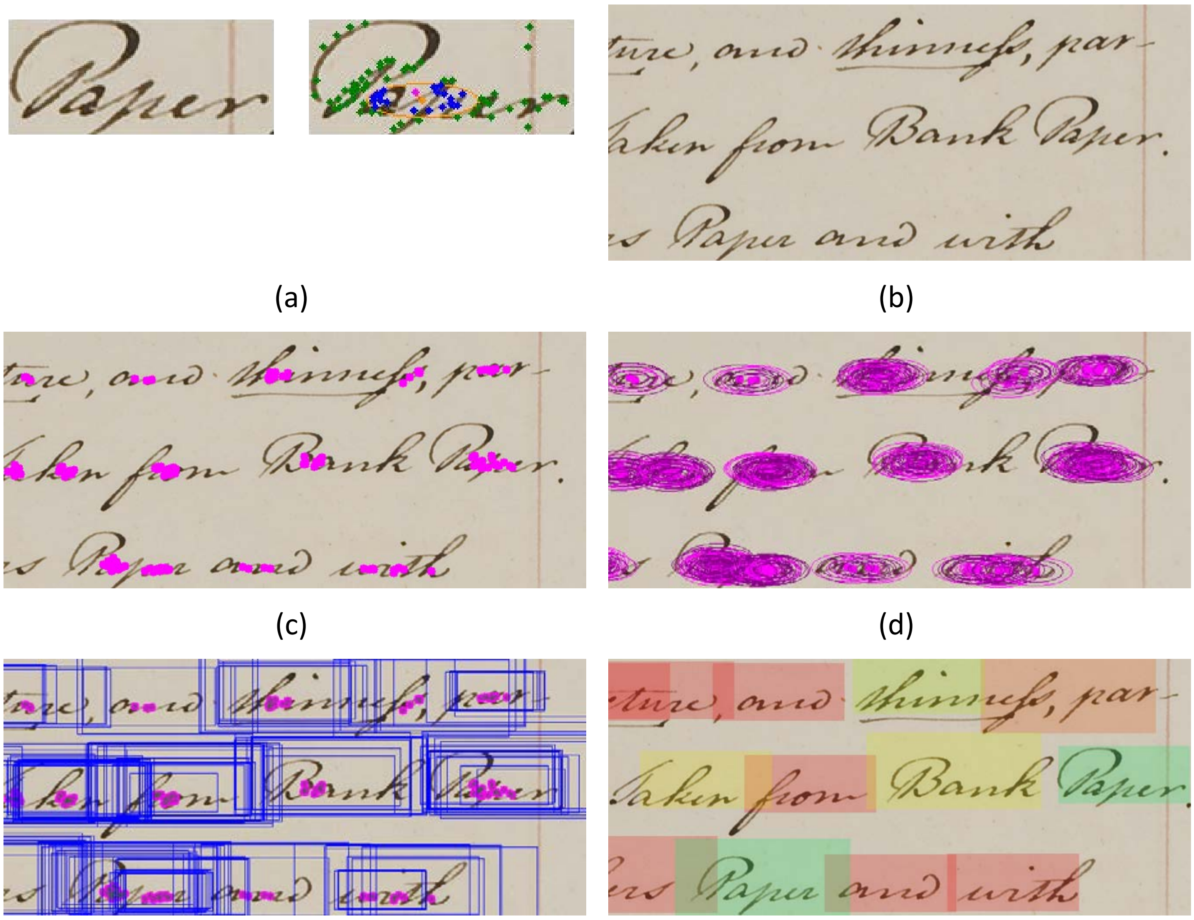


Figure 1.1: The segmentation-based operation pipeline (a) the query keypoints, (b) the word image keypoints, (c) the projection of query keypoints to the word image: the red lines connect the matched local points, the green area (right image) is the local proximity area of the nearest neighbour search in the normalized space.

2.2. Keyword Spotting in Big Data

The keyword spotting method discussed in Section 2.1 lowers its efficiency as the number of documents increases. This performance decrease is larger under a segmentation-free scenario which is arguably the most useful for handwritten historical documents whose word segmentation is nearly impossible.

The current effort under the READ framework is to make the keyword spotting, especially the segmentation-free approach, viable for big data. The first project year's efforts focused on optimizations for the local points extraction, matching algorithm and storage of the DoLFs.



(a) (b) (c) (d) (e) (f)

Figure 1.2 : The segmentation - free operational pipeline (a) the query image, the localpoints, the central location (shown in magenta colour) and its nearest keypoint (shown in orange colour), (b) the document image, (c) the candidate local points for the document coordinate origin, (d) multiple instances of word boundaries around each candidate coordinate origin, (e) multiple word detection (f) final result (the green colour denotes the most similar word).

In particular, the optimizations have been addressed by the following modifications:

i. local points reduction by broader Connecting Components filtering,
ii. constrained brute force search for the segmentation-free approach
iii. quantization of the local point descriptor

The resulting modified method is denoted as '**DUTH-DoLFs Optim'** for the segmentation-based case (see Table 4.1 – 4.2) while in the segmentation-free case, the corresponding method is denoted as '**DUTH-DoLFs Optim_noSeg'** (see Table 4.3 – 4.5).

In addition to the aforementioned modifications we encountered optimizations for the nearest neighbor search algorithms of the initial distance algorithm. To this end, we modified our segmentation-free matching procedure to include two efficient strategies dealing with the proximity search problem: hashing and random kD-trees.

Hashing is one of the popular solutions for approximate nearest neighbour search. In general, hashing is an approach of transforming the data item to a low-dimensional representation, or the equivalently of a short code consisting of a sequence of bits. This exploits the locality sensitive property that similar items have larger probability to be mapped to the same code than the dissimilar items. The main research efforts along this direction consist of designing hash functions satisfying the locality sensitive property as well as designing efficient search schemes using hash tables.

In order to incorporate a hashing function to our matching procedure we have modified the way we calculate the candidate local points for the document coordinate origin. Instead of using an Euclidean Distance to find the top N candidate local points for each document, we apply a 2-stable distribution function [Datar2004] to each local point descriptor in dataset in order to produce a locality-sensitive hashing code that corresponds to each local point. Local points with similar corresponding descriptors have the same hashing code. Therefore, the candidate local points are taken from the same hash bucket with the central query local point.

The kD-tree [Bentley1975, Friedman1977, Muja2014, Silpa2008] is one of the best known nearest neighbor algorithms. While very effective in low dimensionality spaces, its performance quickly decreases for high dimensional data.

To incorporate the kD-tree algorithm to the matching procedure we have modified the way we calculate the locations of the local points' nearest neighbors by representing the local points contained in every document as a kD-tree.

Summarizing, the latter optimizations comprise:

- A hashing technique that incorporates a ρ-stable distribution function for the detection of word location candidates in the dataset
- A kD-tree algorithm to calculate the near-neighbor local points

The resulting method is denoted as '**DUTH-Indexing Optim_NoSeg'** (see Table 4.3 – 4.5)

Both aforementioned methods have been implemented in C#/.NET and are available at GitHub under LGPL-3.0:

https://github.com/Transkribus/VCG-DUTH-Word-Spotting-by-Example

3. NCSR Keyword Spotting framework

Three novel methods were developed for the task of query-by-example segmentation-based keyword spotting. A brief description of these methods is given below:

1. NCSR-ZAH [Sfikas2016]
   A pre-trained Deep Convolutional Network (DCN) is used on a zoning of the word images. The DCN was trained on an independent set of typewritten characters. The retrieval was performed using the Euclidean distance.
2. NCSR-POG [Retsinas2016]
   A novel descriptor, first introduced in [Retsinas2015], is used on the word image. The descriptor, referred as Projections of Oriented Gradients (POG), is based on encoding the projections of the gradient orientation, similar to Radon transform. The retrieval is performed using the Euclidean distance.
3. NCSR-SeqPOG [Retsinas2017]
   Sequences of descriptors are generated by a zoning procedure on the word images, aiming to capture horizontal translations at the cost of a more complicated matching procedure. We chose POG as the local descriptor due to its performance ([Retsinas2015], [Retsinas2016]). The retrieval is performed using a novel sequence matching procedure, which finds the best subset pairing of the local descriptors.

The main goal of these approaches is minimizing the retrieval time in order to develop user-friendly keyword spotting applications. The first two methods (NCSR-ZAH and NCSR-POG) produce a single fixed-length descriptor for each image and consequently the retrieval is straightforward and efficient. It should be noted that the NCSR-ZAH method relies on DCN and without the proper hardware (GPUs) the extraction of the query descriptor is relatively slow. On the other hand, the NCSR-SeqPOG method has a significant performance gain (on every dataset that we experimented) at the expense of a slower retrieval time. Nonetheless, the involved sequence matching procedure is optimized using a dynamic programming approach and thus providing a satisfying compromise to the performance vs time dilemma. The NCSR keyword spotting methods are available, as console applications, at github:

https://github.com/Transkribus/NCSR_Tools

One straightforward and simple extension of these methods is their application on a segmentation-free keyword spotting task. This can be implemented by extracting candidate, possibly overlapping, word regions involving a word-segmentation procedure. Another extension towards segmentation-free task, is the modification of NCSR-SeqPOG method, which is sequence based, in order to be applicable at line-level sequences.

4. Evaluation

The proposed methods have been evaluated against two datasets (Figure 4.1):

- English Dataset which contains 109 Pages and 15923 words
- German Dataset named Konzilsprotokolle (https://doi.org/10.5281/zenodo.215383) which contains 100 Pages and 15579 words

We employ two evaluation modes:

- Mode 1: Exactly the Same with the query image
- Mode 2: Punctuation Marks (, .) and capitals are considered in the ground truth corpora

The queries consist of every word with length greater than 3 and frequency greater than 2. Therefore, the English dataset query set size is 4303 for Mode 1 and 4790 for Mode 2 and the German dataset query set size is 6875 for Mode 1 and 7119 for Mode2.

The performance of the word spotting methods was recorded in terms of the Precision at Top 5 Retrieved words (P@5) as well as the Mean Average Precision (MAP) [Pratikakis2014]. Time and memory requirements are recorded in terms of the following metrics which are self-explanatory: Retrieval Time per Query (RTpQ), Extraction Time per Document (ETpD) and Size per Document (SpD).

The evaluation of both DUTH and NCSR methods is performed on an 8-core Intel i7-4770K at 3.50GHz with 16Gb of RAM for parallel computation (4 cores). All DUTH methods are currently implemented in C#/.NET and all NCSR methods are currently implemented using MATLAB, while the feed-forward Deep Convolutional Network for NCSR-ZAH extraction is not optimized and runs on CPU instead of GPU.

We distinguish two main evaluation scenarios: segmentation-based and segmentation-free.

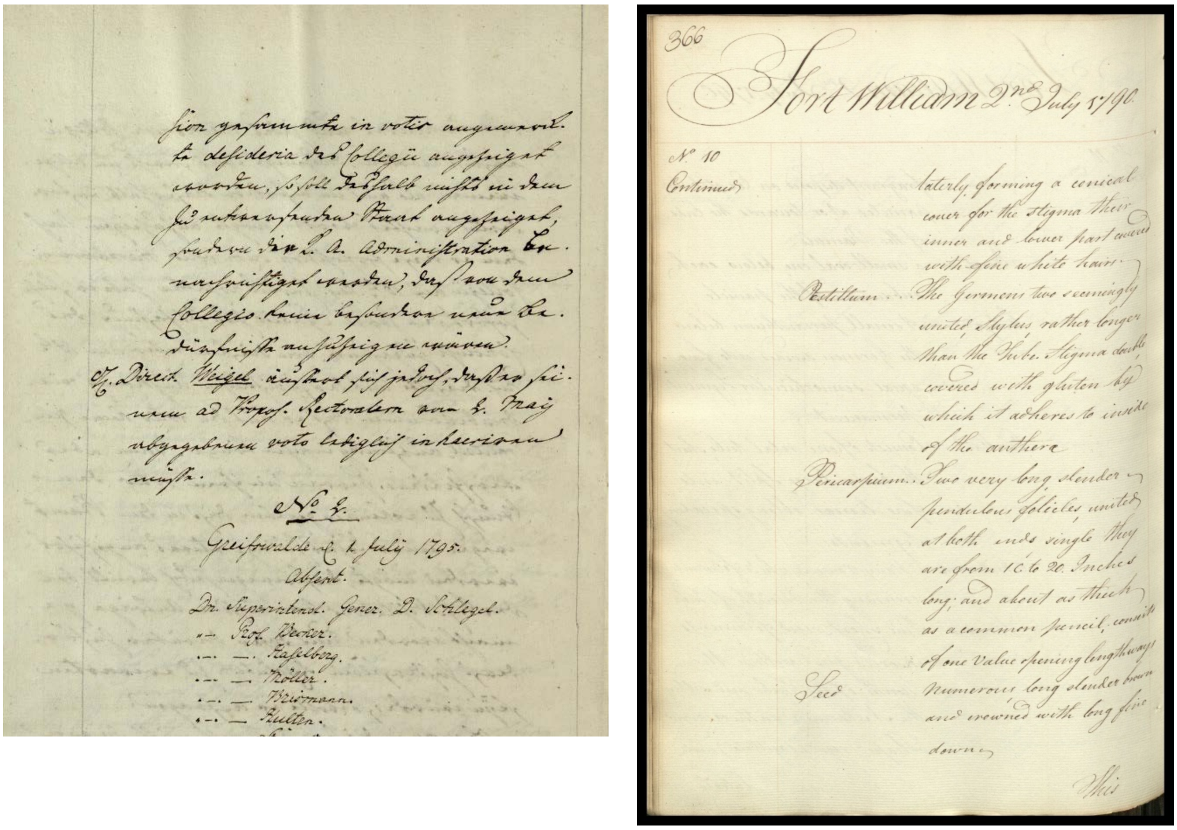


Figure 4.1 The German (left) and the English (right) Datasets

4.1. Conclusive remarks on the Segmentation-Based Scenario

Performance evaluation in terms of retrieval accuracy for DUTH and NCSR keyword spotting segmentation-based methods are presented in Table 4.1, while time and memory requirements are presented in Table 4.2 by averaging the corresponding metrics over the two datasets. In terms of performance, NCSR-SeqPOG method outperforms all the other methods, while at the same time the requirements with the respect to time and memory are retained low. The Euclidean-based distance matching, used in NCSR-ZAH and NCSR-POG methods, produces outstanding results in terms of time and memory. It should also be noted that NCSR-POG achieves good performance, comparable to NCSR-ZAH method, and it is only outperformed by NCSR-SeqPOG as well as DUTH-original method, while at the same time it is the faster of all methods and it requires the minimum storage space. Finally, it can be observed that DUTH's updated version (DUTH-DoLFs Optim) is faster and requires less storage space than their previous method. However, the gain in time and memory requirements leads to a significant drop in performance on both datasets.

Table 4.1. Evaluation Results for the Segmentation-Based Scenario

German Dataset				
	Mode 1		Mode 2	
Method	P@5	MAP	P@5	MAP
DUTH-Original	0.74	0.63	0.78	0.64
DUTH-DoLFs Optim	0.54	0.40	0.54	0.38
NCSR-ZAH	0.65	0.52	0.69	0.52
NCSR-POG	0.68	0.58	0.73	0.58
NCSR-SeqPOG	0.75	0.67	0.81	0.68
English Dataset				
	Mode 1		Mode 2	
Method	P@5	MAP	P@5	MAP
DUTH-Original	0.57	0.44	0.56	0.42
DUTH-DoLFs Optim	0.44	0.37	0.42	0.33
NCSR-ZAH	0.56	0.45	0.55	0.41
NCSR-POG	0.56	0.46	0.55	0.42
NCSR-SeqPOG	0.63	0.53	0.63	0.49

Table 4.2. Time and Memory Requirements for the Segmentation-Based Scenario

Method	*RTpQ(sec)*	*ETpD(sec)*	*SpD(KB)*
DUTH-Original	1.2550	6.57	15073
DUTH-DoLFs Optim	0.1025	4.47	242
NCSR-ZAH	0.0561	58.54	478
NCSR-POG	0.0076	3.24	97
NCSR-SeqPOG	0.0625	4.52	410

4.2. Conclusive remarks on the Segmentation-Free Scenario

Table 4.3 shows the segmentation-free evaluation results for the original method (DUTH-Original_NoSeg)[Zagoris2017], and the corresponding optimized versions as described in Section 2.2, namely, 'DUTH-DoLFs Optim_NoSeg' and 'DUTH-Indexing Optim_NoSeg'. Table 4.4 shows results when a text line-based ground truth is used. These experiments were conducted under the same parameters as those of the segmentation-based approaches. The time and memory requirements are presented in Table 4.5 by averaging the corresponding metrics over the two datasets.

Concerning the evaluation, the 'DUTH-Indexing Optim_NoSeg' method causes a minor drop in MAP evaluation only in German Dataset and achieves increased performance in terms of P@5 in both datasets as well as a sharp increase in terms of MAP for English Dataset. This occurs because the selection of candidate origin local points based on the same hash has filtered many local points. Also, it is worth noting that in the case of the 'DUTH-Indexing Optim_NoSeg' method, the retrieval time was reduced from 2.25 sec per query to 0.36 sec. This is a very optimistic sign with respect to the work which will be addressed in the second year of the project and will concern the access of big datasets.

Table 4.3 Experimental Results for word-based evaluation in a segmentation-free context

German Dataset				
	Mode 1		**Mode 2**	
Method	**P@5**	**MAP**	**P@5**	**MAP**
DUTH-Original_NoSeg	0.63	0.42	0.59	0.42
DUTH-DoLFs Optim_NoSeg	0.41	0.30	0.42	0.29
DUTH-Indexing Optim_NoSeg	0.43	0.25	0.46	0.24
English Dataset				
	Mode 1		**Mode 2**	
Method	**P@5**	**MAP**	**P@5**	**MAP**
DUTH-Original_NoSeg	0.39	0.28	0.35	0.22
DUTH-DoLFs Optim_NoSeg	0.25	0.18	0.25	0.15
DUTH-Indexing Optim_NoSeg	0.38	0.28	0.38	0.25

Table 4.4 Experimental Results for text line-based evaluation in a segmentation-free context

German Dataset				
	Mode 1		**Mode 2**	
Method	**P@5**	**MAP**	**P@5**	**MAP**
DUTH-Original_NoSeg	0.64	0.45	0.60	0.45
DUTH-DoLFs Optim_NoSeg	0.41	0.30	0.47	0.33
DUTH-Indexing Optim_NoSeg	0.44	0.27	0.47	0.26
English Dataset				
	Mode 1		**Mode 2**	
Method	**P@5**	**MAP**	**P@5**	**MAP**
DUTH-Original_NoSeg	0.41	0.30	0.38	0.25
DUTH-DoLFs Optim_NoSeg	0.27	0.20	0.29	0.19
DUTH-Indexing Optim_NoSeg	0.39	0.29	0.39	0.26

Table 4.5. Time and Memory Requirements for the Segmentation-free Scenario

Method	RTpQ(sec)	ETpD(sec)	SpD(KB)
DUTH-Original_NoSeg	15.84	12.85	19800
DUTH-DoLFs Optim_NoSeg	2.25	10.16	1410
DUTH-Indexing Optim_NoSeg	0.36	6.27	12300

II. The Query by String (QbS) case engine

For a set of text images, keyword spotting (KWS) consists in finding the images (and maybe the regions or locations within each image) where specific words may appear. Rather than deterministic results, KWS systems are expected to provide, for each detected spot of a query word, a confidence score which measures how sure is the system that the word appears in the spotted image or location. This allows the user to somehow establish a confidence threshold to specify the required "precision-recall trade-off"; that is the balance between the accuracy of the spotting results (referred to as "precision") and the amount of correct images actually retrieved (referred to as "recall").

In the QbS KWS setting, query words are given in the form of strings of letters, which is a very flexible and convenient form in many applications. Also for this very same reason, QbS KWS properly provides the basic technologies to develop indexing and search systems which aim at supporting fast free-text content access to (very) large collections of untranscribed handwritten text images.

Basic QbS techniques can also be used to very effectively deal with typical QbE tasks, as shown in [Vidal2015] and [Puigcerver2015].

1. UPVLC Keyword Spotting framework

UPVLC develops QbS KWS technologies within the information-retrieval domain and following well-funded statistical methodologies. The spotting confidence score is assumed to be the probability that an image, region, or location is "relevant" for the query keyword. An image is considered to be relevant if the word is actually written in it. Following this very general framework, several approaches are being developed by UPVLC. These different approaches aim at properly dealing with corresponding indexing and/or search problems raised by indexing and search applications involving hundreds of thousands or even millions of handwritten page images.

The work carried out by UPVLC under this framework during the first year of READ is described in the following subsections. Each subsection is associated with a publication in a scientific journal or in the proceedings of a major international conference. Therefore, only a brief summary of each work is provided, accompanied by the corresponding reference to the published paper.

The READ platform “Transkribus” already includes Indexing and Search tools based on these works.

1.1. A Probabilistic Formal Framework for QbS Word-Graph Based KWS

In this case, the spotting targets or image regions are considered to be text lines. The relevance probability is obtained by adequately combining the so called "frame-level word posterior probabilities". For each word, w, in a given vocabulary of words to be indexed, and for each horizontal position or “frame”, f, within a line, these probabilities measure the chance that the word w is written in a stretch of the image line which includes the frame f.

Frame-level posteriors are obtained using word graphs or lattices derived from the recognition process of a full-fledged handwritten text recognizer based on hidden Markov models and N-gram language models. This approach has several advantages. First, since it uses a holistic, segmentation-free technology, it does not require any kind of word or character segmentation. Second, the use of language models allows the context of each spotted word to be taken into account, thereby considerably increasing KWS accuracy. And third, the resulting KWS scores are based on true posterior probabilities, taking into account all (or most) possible word segmentations of the line image. Since these scores are properly bounded and normalized, this formulation lends itself to smooth, threshold-based keyword queries which, in turn, permit comfortable trade-offs between search precision and recall.

Experiments are carried out on several historic collections of handwritten text images, as well as a well-known data set of modern English handwritten text. According to the empirical results, the proposed approach achieves KWS results comparable to those obtained with the recently-introduced "BLSTM neural networks KWS" approach, which is much more expensive in terms of training requirements. On the other hand, it also clearly outperforms the popular, state-of-the-art "Filler HMM" KWS method. Overall, the results clearly support all the above-claimed advantages of the proposed approach. See more details in [Toselli2016a].

1.2. Assessing the impact of lattice size in QbS KWS based on word-lattices

Two document processing applications are considered: computer-assisted transcription of text images (CATTI) [Romero2012] and KWS [Toselli2016a], for transcribing and indexing handwritten documents, respectively. Instead of working directly on the handwriting images, both employ meta-data structures called word lattices or graphs (WG), which are obtained using segmentation-free handwritten text recognition technology based on N-gram language models and hidden Markov models. A WG contains most of the relevant information of the original text (line) image required by CATTI and KWS but, if it is too large, the computational cost of generating and using it can become non-affordable. Conversely, if it is too small, relevant information may be lost, leading to a reduction of CATTI or KWS performance.

We study the trade-off between WG size and performance in terms of effectiveness and efficiency of CATTI and KWS. Results show that small, computationally cheap WGs can be used without losing the excellent CATTI and KWS performance achieved with huge WGs. For more details, refer to [Toselli2016c].

1.3. New approaches for querying out-of-vocabulary words in lexicon-based KWS

Lexicon-based handwritten text KWS has proven to be a faster and more accurate alternative to lexicon-free methods. Nevertheless, since lexicon-based KWS relies on a predefined vocabulary, fixed in the training and indexing phase, it does not support queries involving non-indexed, out-of-vocabulary (OOV) keywords. In this paper, we outline previous work aimed at solving this problem and present a new approach based on smoothing the (null) scores of OOV keywords by means of the information provided by "similar" in-vocabulary words.

Good results achieved using this approach are compared with previously published alternatives on different data sets. See more details in [Puigcerver2016].

1.4. Comparing two Word-Lattice-based approaches for lexicon-free QbS KWS

Two methods are presented to improve word confidence scores for line-region QbS Lexicon-Free KWS in handwritten text images. The first one approaches true relevance probabilities by means of computations directly carried out on character lattices obtained from the lines images considered. The second method uses the same character lattices, but it obtains relevance scores by first computing frame-level character sequence scores which resemble the frame-level word posteriors used in previous approaches for lexicon-based KWS.

The first method results from a formal probabilistic derivation, which allow us to better understand and further develop the underlying ideas. The second one is less formal but, according to the experiments presented in the paper, it obtains almost identical results with much lower computational cost. Moreover, in contrast with the first method, the second one allows to directly obtain accurate bounding boxes for the spotted words.

To see more details, refer to [Toselli2016b].

1.5. QbS Keyword Spotting in Historical Daily Records Documents

In contrast with the works presented in the previous subsections, which were more fundamental and methodological, the work presented here deals with the application of previously described approaches to KWS (as well as to HTR) to a specific type of handwritten documents.

Historical records of daily activities provide an intriguing view of the historical life and contain interesting information useful for demography studies and genealogical research. However, so far automatic processing of historical documents has mostly been focused on single works of literature and less on daily records, which tend to have a distinct layout, structure, and vocabulary. This paper presents a study about the capability of state-of-the-art handwritten text recognition and KWS systems, when applied to this kind of documents.

A relatively small set of handwritten birth records registered in Wien in the 16th century is used in the experiments. A word accuracy of about 70% and an Average Precision of 0.74 are achieved for plain image transcription and KWS, respectively. Considering the many difficulties exhibited by these handwritten documents, these preliminary results are quite encouraging.

See details in [Romero2016].

References

[Bentley1975] J. L. Bentley, “Multidimensional binary search trees used for associative searching,” Commun. ACM, vol. 18, no. 9, pp. 509–517,1975.

[Datar2004] M. Datar, N. Immorlica, P. Indyk, and V. S. Mirrokni. Locality-sensitive hashing scheme based on p-stable distributions. In Symposium on Computational Geometry, pages 253–262, 2004

[Friedman1977] J. H. Friedman, J. L. Bentley, and R. A. Finkel, “An algorithm for finding best matches in logarithmic expected time,”ACM Trans. Math. Softw., vol. 3, no. 3, pp. 209–226, 1977.

[Manmatha1996] R Manmatha, Chengfeng Han, and Edward M Riseman. Word spotting: A new approach to indexing handwriting. In Computer Vision and Pattern Recognition, 1996. Proceedings CVPR'96, 1996 IEEE Computer Society Conference on, pages 631-637. IEEE, 1996.

[Muja2014] M. Muja and D. G. Lowe, "Scalable Nearest Neighbor Algorithms for High Dimensional Data," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 36, no. 11, pp. 2227-2240, Nov. 1 2014.

[Pratikakis2014] I. Pratikakis, K. Zagoris, B. Gatos, G. Louloudis and N. Stamatopoulos, “ICFHR 2014 Competition on Handwritten KeyWord Spotting (H-KWS 2014)”, International Conference on Frontiers in Handwriting Recognition, pp. 814-819, Greece, 2014.

[Puigcerver2015] Joan Puigcerver, Alejandro H. Toselli, Enrique Vidal: "ICDAR2015 Competition on Keyword Spotting for Handwritten Documents". In Proceedings of ICDAR-2015, pp. 1176-1180, 2015.

[Puigcerver2016] Joan Puigcerver, Alejandro H. Toselli and Enrique Vidal. "Querying out-of-vocabulary words in lexicon-based keyword spotting". Neural Computing and Applications, pages 1-10, 2016. ISSN: 1433-3058, DOI: 10.1007/s00521-016-2197-8.

[Retsinas2015] G. Retsinas, B. Gatos, N. Stamatopoulos and G. Louloudis, “Isolated Character Recognition using Projections of Oriented Gradients”, International Conference on Document Analysis and Recognition, pp. 336-340, France, 2015.

[Retsinas2016] G. Retsinas, G. Louloudis, N. Stamatopoulos and B. Gatos, “Keyword Spotting in Handwritten Documents using Projections of Oriented Gradients”, Workshop on Document Analysis Systems, pp. 411-416, Greece, 2016.

[Retsinas2017] G. Retsinas, G. Louloudis, N. Stamatopoulos and B. Gatos, “Efficient Learning-Free Keyword Spotting”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Under Review.

[Romero2012] Verónica Romero, Alejandro H. Toselli, and Enrique Vidal. "Multimodal interactive handwritten text transcription". Vol. 80. World Scientific, 2012.

[Romero2016] Verónica Romero, Alejandro H. Toselli, Joan A. Sánchez and Enrique Vidal, "Handwriting Transcription and Keyword Spotting in Historical Daily Records Documents" 2016 12th IAPR Workshop on Document Analysis Systems (DAS), Santorini, 2016, pp. 275-280. doi: 10.1109/DAS.2016.70

[Sfikas2016] G. Sfikas, G. Retsinas and B. Gatos, "Zoning Aggregated Hypercolumns for Keyword Spotting", International Conference on Frontiers in Handwriting Recognition, China, 2016.

[Silpa2008] C. Silpa-Anan and R. Hartley, "Optimised kD-trees for fast image descriptor matching," in Proc. IEEE Conf. Comput. Vis. Pattern Recog., 2008, pp. 1–8.

[Toselli2016a] Alejandro H. Toselli, Enrique Vidal, Verónica Romero, Volkmar Frinken: "HMM word graph based keyword spotting in handwritten document images", Information Sciences, Volumes 370-371, 20 November 2016, Pages 497-518, ISSN 0020-0255, http://dx.doi.org/10.1016/j.ins.2016.07.063.

[Toselli2016b] Alejandro H. Toselli, Joan Puigcerver and Enrique Vidal. "Two Methods to Improve Confïdence Scores for Lexicon-free Word Spotting in Handwritten Text". In Proceedings of the 15th International Confe rence on Frontiers in Handwriting Recognition (ICFHR 2016). Pages 349-354, Shenzhen, China (October 2016). Published by IEEE Computer Society, ISBN-13: 978-1-5090-0981-7.

[Toselli2016c] Alejandro H. Toselli, Verónica Romero, Enrique Vidal. Word graphs size impact on the performance of handwriting document applications. Neural Computing and Applications, 2016.

[Vidal2015] Enrique Vidal, Alejandro H. Toselli, Joan Puigcerver: "High performance Query-by-Example keyword spotting using Query-by-String techniques". In Proceedings of ICDAR-2015, pp741-745, 2015.

[Zagoris2017] K. Zagoris, I. Pratikakis, B. Gatos. 2017 Unsupervised Word Spotting in Historical Handwritten Document Images using Document-oriented Local Features. Transactions on Image Processing. Under Review.

[Zagoris2014] K. Zagoris, I. Pratikakis and B. Gatos, "Segmentation-Based Historical Handwritten Word Spotting Using Document-Specific Local Features," 2014 14th International Conference on Frontiers in Handwriting Recognition, Heraklion, 2014, pp. 9-14.

[Zagoris2015] K. Zagoris, I. Pratikakis, and B. Gatos, "A framework for efficient transcription of historical documents using keyword spotting," in Historical Document Imaging and Processing (HIP'15), 3rd International Workshop on, August 2015, pp. 9–14.