SEMINAR TALK LA ROCHELLE, JUNE 28 HANDWRITING RECOGNITION IN HISTORICAL DOCUMENTS

Andreas Fischer Post-doctoral researcher, DIVA group Department of Informatics, University of Fribourg andreas.fischer@unifr.ch

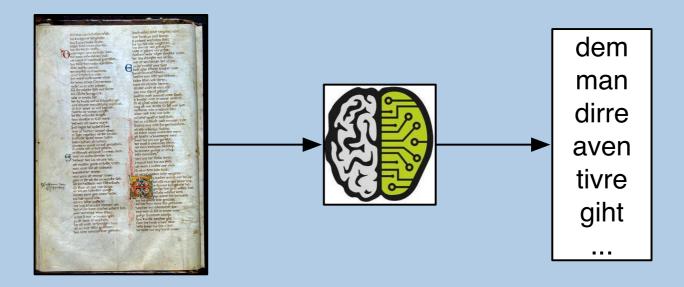
Motivation

- Integration of historical manuscripts in digital libraries.
- Making our cultural heritage accessible to researchers and the public.



Objective

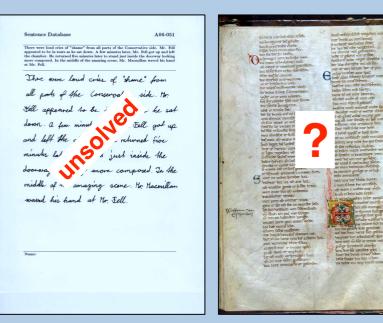
- Transcription of digital images into machine-readable text.
- Automatic reading needed for processing millions of manuscripts.



State of the Art

- Industrial solutions available for printed documents.
- Unsolved problem for handwriting images.



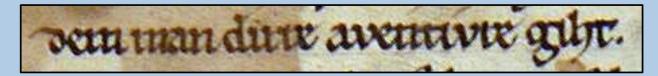


4

Learning by Samples

D UNIVERSITÄT BERN

- Recognition rules are learned from samples.
- Difficult and costly to obtain in case of historical documents.

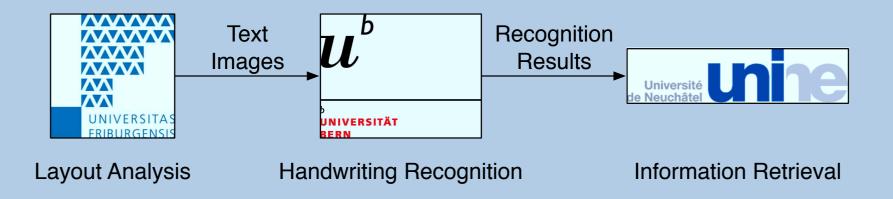


dem man dirre aventivre giht.

HisDoc Project

UNIVERSITÄT BERN

- Synergy research project of the Swiss National Science Foundation.
- Pioneering work on historical handwriting recognition.



Outline

b UNIVERSITÄT BERN

• Database

- Ground Truth Creation
- Automatic Transcription
 - Graph Similarity Features
 - Fast Recognition Algorithm
- Keyword Spotting
- Transcription Alignment
- Conclusions

8

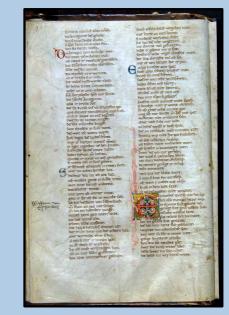
UNIVERSITÄT BERN

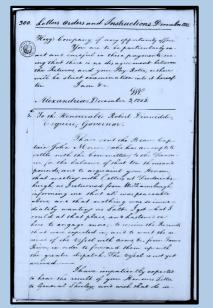
• One of the first comprehensive research databases in the field.

• Freely available on the Internet.

IAM-HistDB

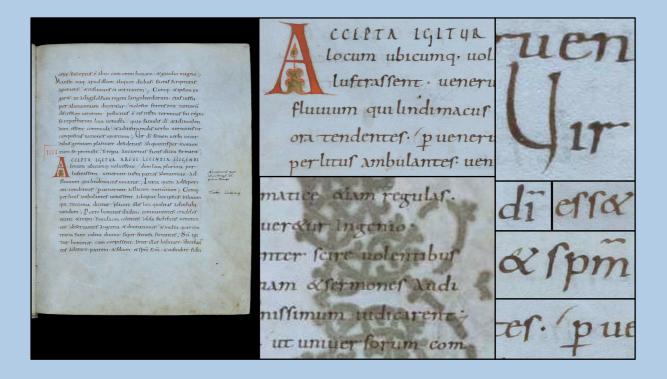






Saint Gall Database

- Vita Sancti Galli by Walafrid Strabo, Cod. Sang. 562, 9th century.
- 60 pages, Latin, Carolingian script.



Parzival Database

b UNIVERSITÄT BERN

- Parzival by Wolfram von Eschenbach, Cod. 857, 13th century.
- 47 pages, German, Gothic script.



George Washington Database

UNIVERSITÄT BERN

- Letters of George Washington, Library of Congress, 18th century.
- 20 pages, English, longhand script.

of Flour, for the two bompanies of Mangers; Letters orders and Instructions De twelve hundred of which to be delivered yearry if any opportunity of Captain Ashby and bompany at the reful in these pays Plantation of Charles Sellars - the rest to letter S. lockes bompany at Nicholas Reasmers. and your Pay Rollis; as there ill be shict examination into it hereaf October 26 Jam to. Gil Alexandria: December 3, 1753. you are hereby ordered to Ren 5. To the Honourable Robert Dinmiddle, dervous at Alexandria the first Day of Osquire; Governor. December, with all the men you can Thave sent the Bear Cap. John Mercer (whe has accompts to raise by that time. settle with the Committees to the Preasu the balance of that ton thousa Given &c. at Predericksburg, November 3. 1955. ens to acquaint your Honour as Ireland from Williamsburgh IM. aid de camp. me that all was peaceable that nothing wanting as Salt. Igot what I much complained of in the last Detachment I at that place, and hastened on you are to apply to Ab ajor Lewis for the In accted in; and to want the structions which I left with him and see of the Vefsel with arms to. from . in order to forward them up with hat they are observed. the greater dispatch. The refall is not yet Given under my Hand Vc. Thave impatiently expected Th to hear the result of your Aonours Letter to General Shiley; and wish that the do giv? Winchester: October 29.1755.

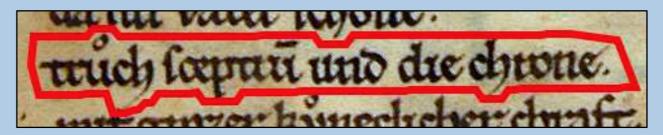
Outline

- Database
 - Ground Truth Creation
- Automatic Transcription
 - Graph Similarity Features
 - Fast Recognition Algorithm
- Keyword Spotting
- Transcription Alignment
- Conclusions

Ground Truth

b UNIVERSITÄT BERN

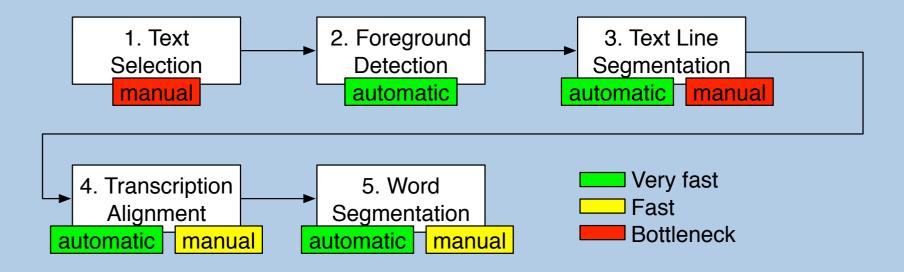
- Supervised manuscript annotation with correct transcriptions.
- Enables learning by samples as well as performance evaluation.



trůch sceptrū(sceptrum) und die chrone.

Semi-Automatic Proceeding

- Automatic: image segmentation (Indermühle et al., 2009) & alignment.
- Manual: text area selection and corrections.



$\textbf{Human} \leftrightarrow \textbf{Machine Interface}$

- Graphical user interface for interactive correction.
- Laypersons spent about 10 minutes for one page.

000		
Open Save Close		(Align Font) + -
Difiv aventivite.	שווי אירוודוידי.	disiv aventivre.
Pvert ane der brehe stivre.	vert ane der brehe flavre.	vert ane der b [°] che stivre.
e man fi hete for ein breh.	e man fi here for ein boch.	ê man si hete fvr ein bůch.
o 1ch were e nachet ane tich.	1ch w.ere e nachet ane rych.	ich wære ê nachet ane tvch.
fo 1ch 111 dem bade fære.	fo 1ch 111 dem lade Lexe.	so ich in dem bade sæze.
ob ich del questen nibt vergeze.	ebicipel quelter nibt veratte.	ob ich des questen niht vergæze.
a object ou ditater in the uvrich mur ben lup.	z macher uvtich mar ben hy.	z machet trvrich mir den lip.
ar also marian bazer wip.	Dix alfo muriai beizer wip.	daz also mængiv heizet wip.
Thinme fint geliche bel.	and ano nation of the bel.	ir stimme sint geliche hel.
genivge fint gein vallebe fnel.	genoge fint geniche hel.	genvge sint gein valsche snel.
etfliche valiches lære.	etfliche valichef lære.	etsliche valsches lære. svs teilent sich div mære.
fys teilent fich div mæver.	fvo teilent fich div mate.	daz die geliche sint genamt.
dar die geliche firtt genantt.	dar die genelle futt genante:	des hat min herze sich geschamt.
ber hat min herze fich geschantt.	Der har min herze fich grichantr.	ipheit din ordenlicher site.
? Snipheit din ordenlicher fite.	Suppert din ordenlicher fite.	dem vert vī fŷr ie triwe mite.
dem vert vn fvr ie triwe mite.	dem vert vn for ie triwe mite.	gnýge sprechent armýt.
D gnyge fprechent armyt.	anvae sprechent armyt.	daz div sîe zenihte gvt.
Say ow fie zentite gut.	bar div fie zenilite qut.	swer die dvrch triwe lîdet.
fiver die ovreh triwe lider.	fwer die durch triwe lider.	helle fiwer die sele nîdet.
helle fiwer die sele nidet.	helle fimer die sele midet.	die dolte ein wip dvrch triwe.
Die dotte ein wip durch trawe.	Die dotte ein wip durch triwe.	

Outline

- Database
 - Ground Truth Creation
- Automatic Transcription
 - Graph Similarity Features
 - Fast Recognition Algorithm
- Keyword Spotting
- Transcription Alignment
- Conclusions

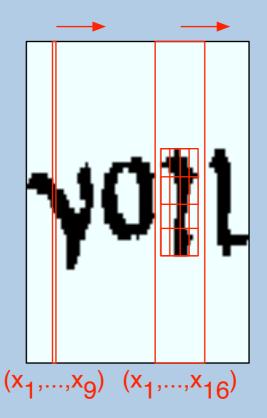
Sayre's Paradox (1973)

- Recognition requires character segmentation.
- Character segmentation requires recognition.



Sliding Window Approach

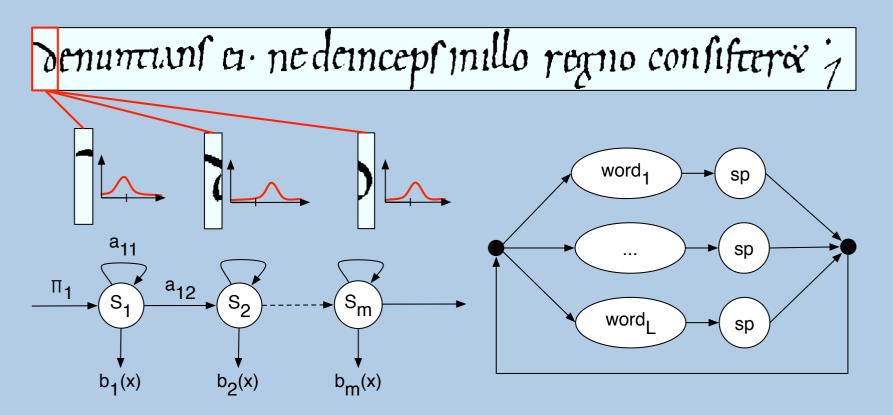
- Slice into sub-character parts, extract features (Marti & Bunke, 2001).
- Reconstruct characters during recognition.



Hidden Markov Models (HMM)

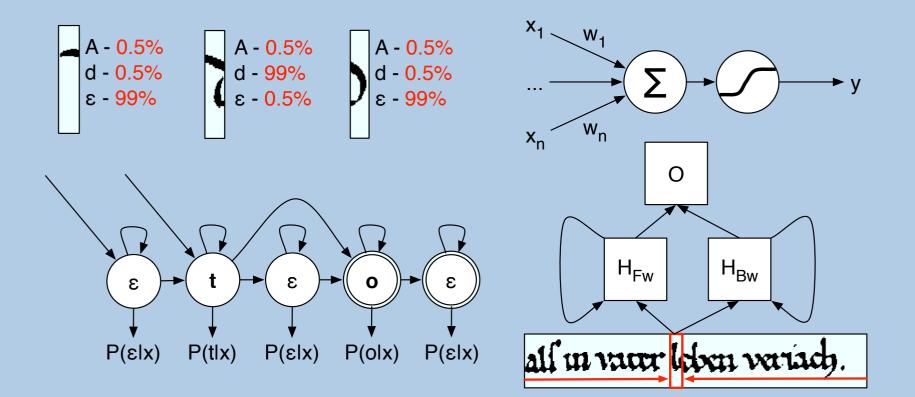
UNIVERSITÄT BERN

- Generative approach to model feature distributions (Rabiner, 1989).
- Efficient algorithms for training (Baum-Welch) and recognition (Viterbi).



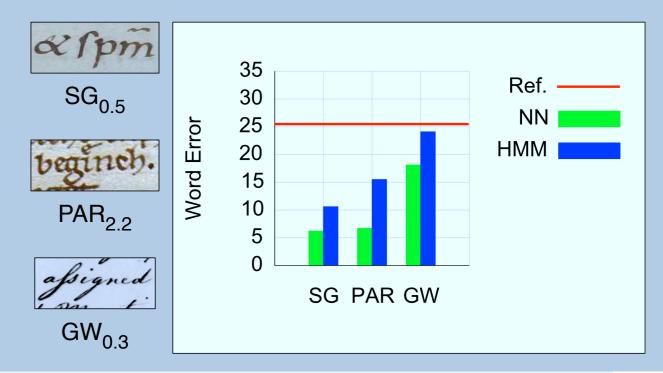
Neural Networks (NN)

- Discriminative approach to distinguish characters (Graves et al., 2009).
- Analysis reveals: training and recognition closely related to HMMs.



Text Line Recognition Results

- Single manuscript recognition, closed vocabulary, word bigrams.
- NN word error: 6% SG_{0.5}, 7% PAR_{2.2}, 18% GW_{0.3}
- \rightarrow Significantly better than modern scripts, promising for digital libraries.

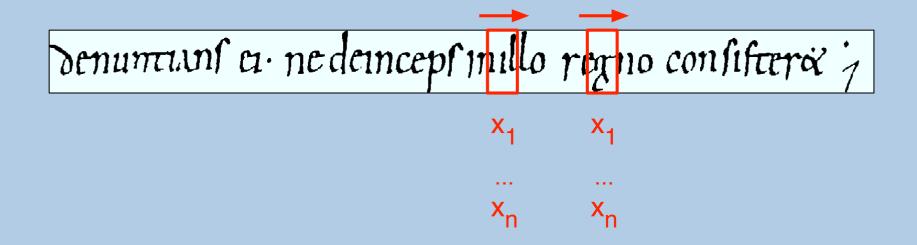


Outline

- Database
 - Ground Truth Creation
- Automatic Transcription
 - Graph Similarity Features
 - Fast Recognition Algorithm
- Keyword Spotting
- Transcription Alignment
- Conclusions

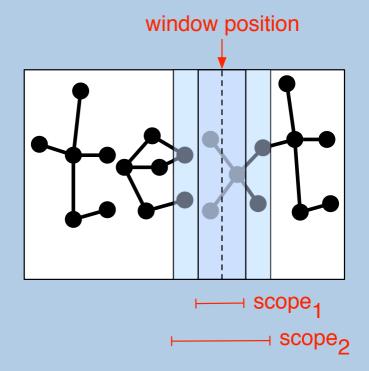
Standard Statistical Approach

- Fixed-size window, fixed-size vectorial features.
- Disregards pattern complexity and sub-part relations.



Proposed Structural Approach

- Dynamic window size, graph-based handwriting representation.
- However: most basic mathematical operations not available for graphs.

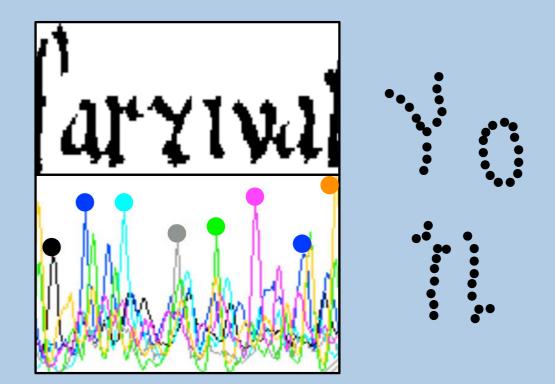


Graph Similarity Features

UNIVERSITÄT BERN

- Bridge the gap with vector space embedding (Riesen & Bunke, 2010).
- Calculate character prototype similarity at each window position.





Word Recognition Results

- Significantly outperforms statistical features on the Parzival database.
- HMM error reduction with respect to (Marti & Bunke, 2001): 50%
- \rightarrow Enables statistical recognition of arbitrary structural representations.

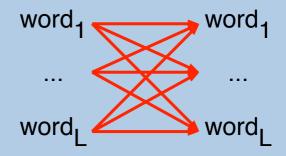
Features	Acc.
Marti & Bunke, 2001	88.69
Vinciarelli <i>et al.</i> , 2004	90.49
Fischer et al., 2011	94.51

Outline

- Database
 - Ground Truth Creation
- Automatic Transcription
 - Graph Similarity Features
 - Fast Recognition Algorithm
- Keyword Spotting
- Transcription Alignment
- Conclusions

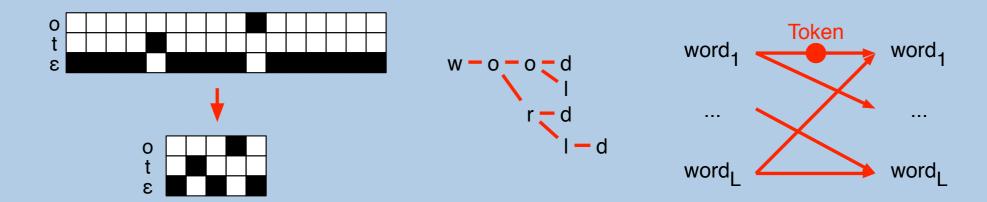
Computational Complexity

- Recognition: computational time \propto word lexicon size²
- *Minutes* per text line, *years* for the George Washington collection.



Search Space Optimization

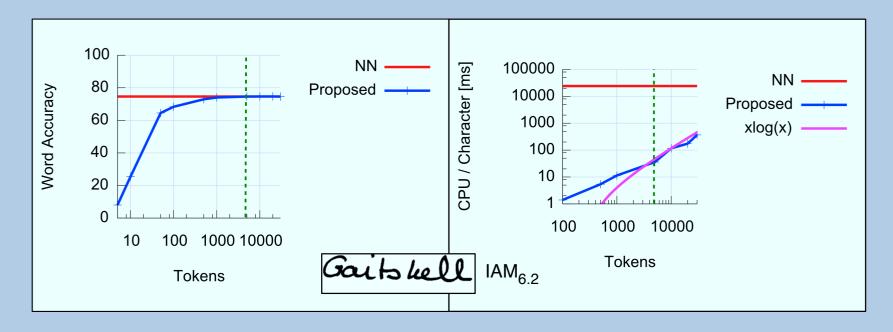
- Compression of subsequent "no character" windows for NN.
- Word prefix based recognition, keep only few promising partial results.



Speedup Results

UNIVERSITÄT BERN

- Over 600 times faster without accuracy loss for 20,000 lexicon words.
- Seconds per text line, weeks for the George Washington collection.
- \rightarrow One of the most accurate and fastest recognition systems worldwide. \rightarrow Furthermore: versatile recognition output in form of word lattices.



Outline

- Database
 - Ground Truth Creation
- Automatic Transcription
 - Graph Similarity Features
 - Fast Recognition Algorithm
- Keyword Spotting
- Transcription Alignment
- Conclusions

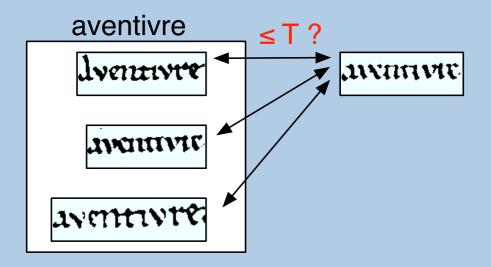
Keyword Spotting

- Identify search terms without transcription (Manmatha et al., 1996).
- Lexicon-free indexing of manuscript images for digital libraries.

270. Letters Orders and Instructions. October 1955.

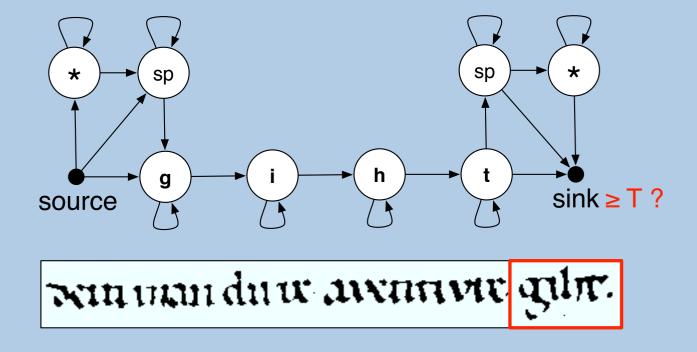
Standard Template-Based Approach

- UNIVERSITÄT BERN
- Image matching between keyword templates and the manuscript.
- Dynamic Time Warping (DTW) distance < threshold ?



Proposed Learning-Based Approach

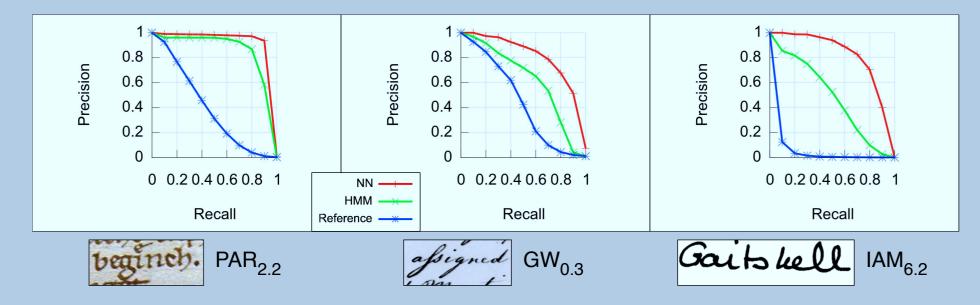
- Concatenate trained character models to keywords.
- Recognition confidence at the most likely position \geq threshold ?



34

Spotting Results

- NN average precision: 95% PAR_{2.2}, 81% GW_{0.3}, 83% IAM_{6.2}
- *Milliseconds* per text line, *hours* for the George Washington collection.
- \rightarrow Best performance on the George Washington database worldwide.



Outline

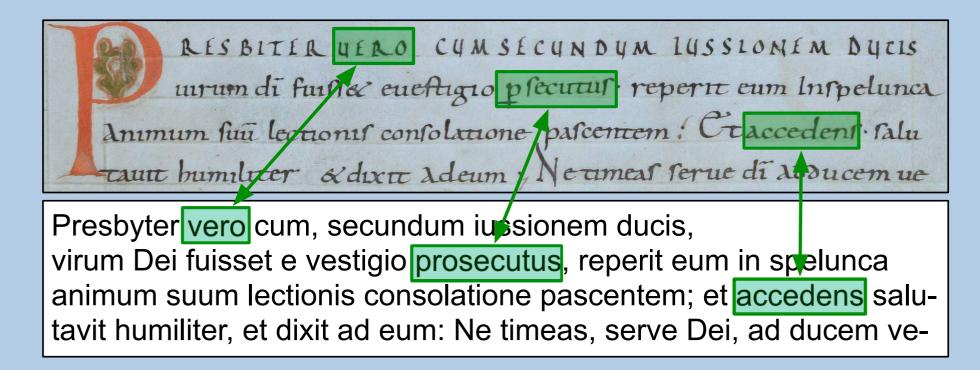
b UNIVERSITÄT BERN

- Database
 - Ground Truth Creation
- Automatic Transcription
 - Graph Similarity Features
 - Fast Recognition Algorithm
- Keyword Spotting
- Transcription Alignment
- Conclusions

Transcription Alignment

UNIVERSITÄT BERN

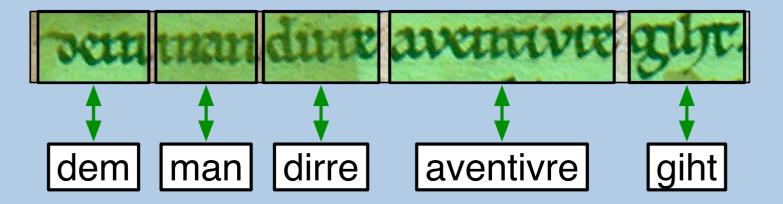
- Existing text editions often deviate from the manuscript image.
- Automatic alignment needed to extract training samples.



Standard Segmentation Problem

UNIVERSITÄT BERN

- 1:1 correlation between text and image.
- Segmentation of text line data sets (Zimmermann & Bunke, 2002).



Inaccurate Transcriptions

UNIVERSITÄT BERN

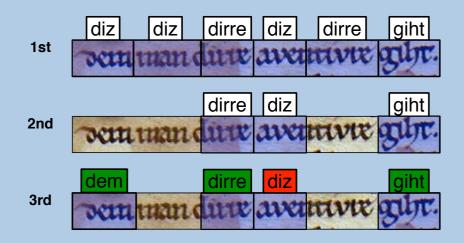
- n:m correlation between text and image.
- Images and texts without correspondence have to be rejected.



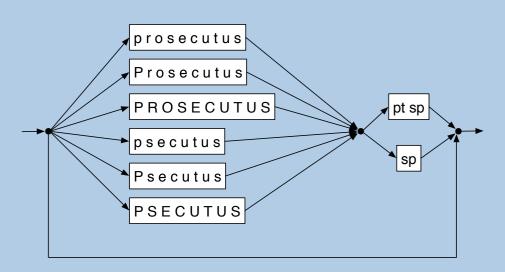
Proposed Alignment Approaches

UNIVERSITÄT BERN

- All text errors: multi-pass HMM recognition with a page lexicon.
- Spelling errors: HMM segmentation with spelling variants.



Transcription: dem man dirre aventivre giht.Edition: demdirre diz willegiht.



Alignment Results

- 50% artificial text errors: 6% PAR_{2.2} alignment error.
- Spelling errors: 8% SG_{0.5}, 17% SG_{0.02} alignment error.

 \rightarrow Automatic database creation feasible with few initial learning samples.

ebynde ir <mark>mynt vil wol gemaebett.</mark> deilwar daz <mark>taz</mark> ieb ane haz.	RESBITER YERO CUMSECUNDYM LUSSLONEM DUCIS
min zorn 1A immer ^u nawe: 1eb bin <mark>Volfram</mark> von Adpenbach.	urun di fuisse euchigio psecutus reperit cum inspelunce
ow but mune libe. owe warvabe time fidaz.	Animum suu lectionis consolutione pascentem : Ot accedens salu
er af idoch ar waybeit. doch enfvin filich vergaben mint:	TAUT humiliter of dixit Adeum : Netimear Verue di 200 ucem ue
me elwnne wol geinezzen. mer ut von berzen let ur pin. Gwer atten fröwen fpricher mat.	RESDITER HERO CHMSECUNDYM LUSSLONEM DUCIS
orene folich ruhr betoren, Some folich ruhr betoren, Sobricht mithur witze chenich.	urum di fusser euchigio psecutus reperit cum inspelunce.
mag ich mit fehilde vi ber mit fper; der an nittebaft nach minnen vilt;	Animum suu lecuonis consolutione pascentein : Et accedens salu
al prate warentive vor. der enzel fi zeohemen biehe,	
verr ane der bvehe hvre.	TAUT humiliter & dixit Adeum ;, Netimear rerue di 200 ucem ue

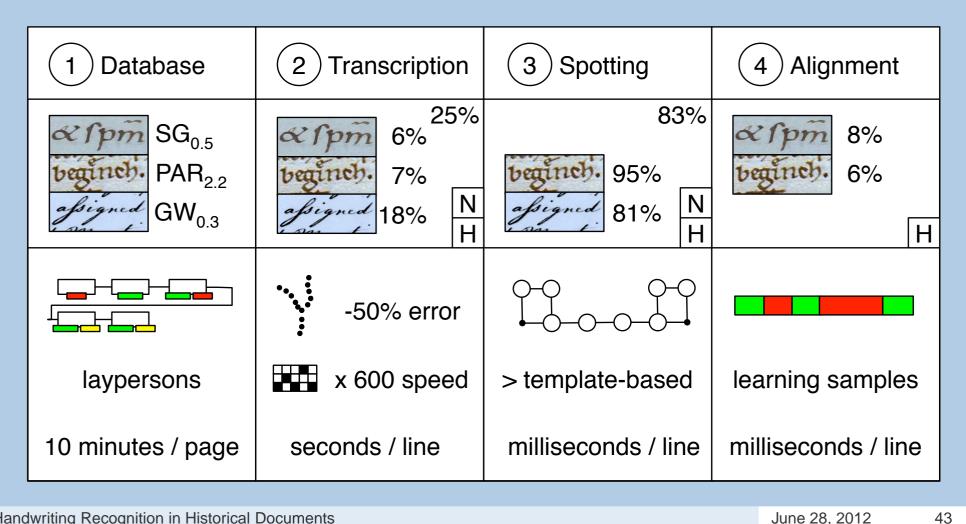
41

Outline

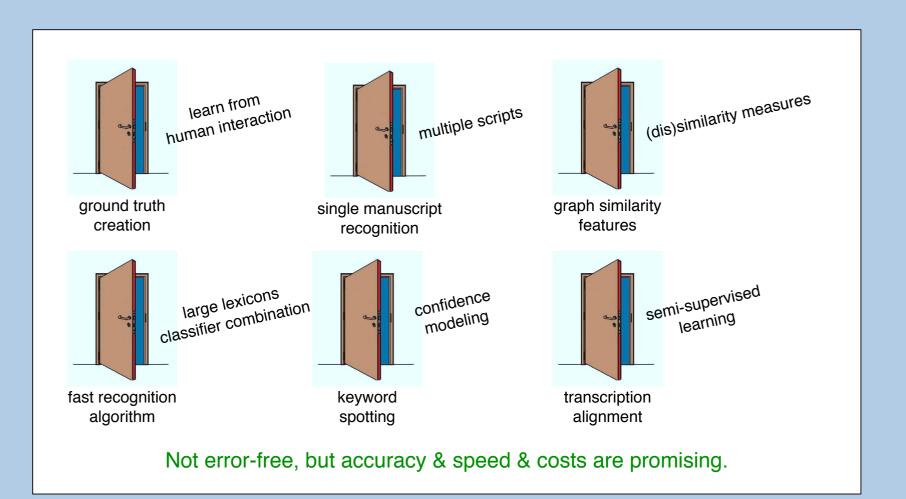
b UNIVERSITÄT BERN

- Database
 - Ground Truth Creation
- Automatic Transcription
 - Graph Similarity Features
 - Fast Recognition Algorithm
- Keyword Spotting
- Transcription Alignment
- Conclusions

Contributions



Outlook



Questions

6 UNIVERSITÄT BERN

 $u^{\scriptscriptstyle b}$

