



Wordspotting (segmentierungsfrei) mit einem patchbasierten Ansatz

Frederik Heerde, Ben Wilkes, Steven Brodziak, Berat Özdemir, Maximilian Brand

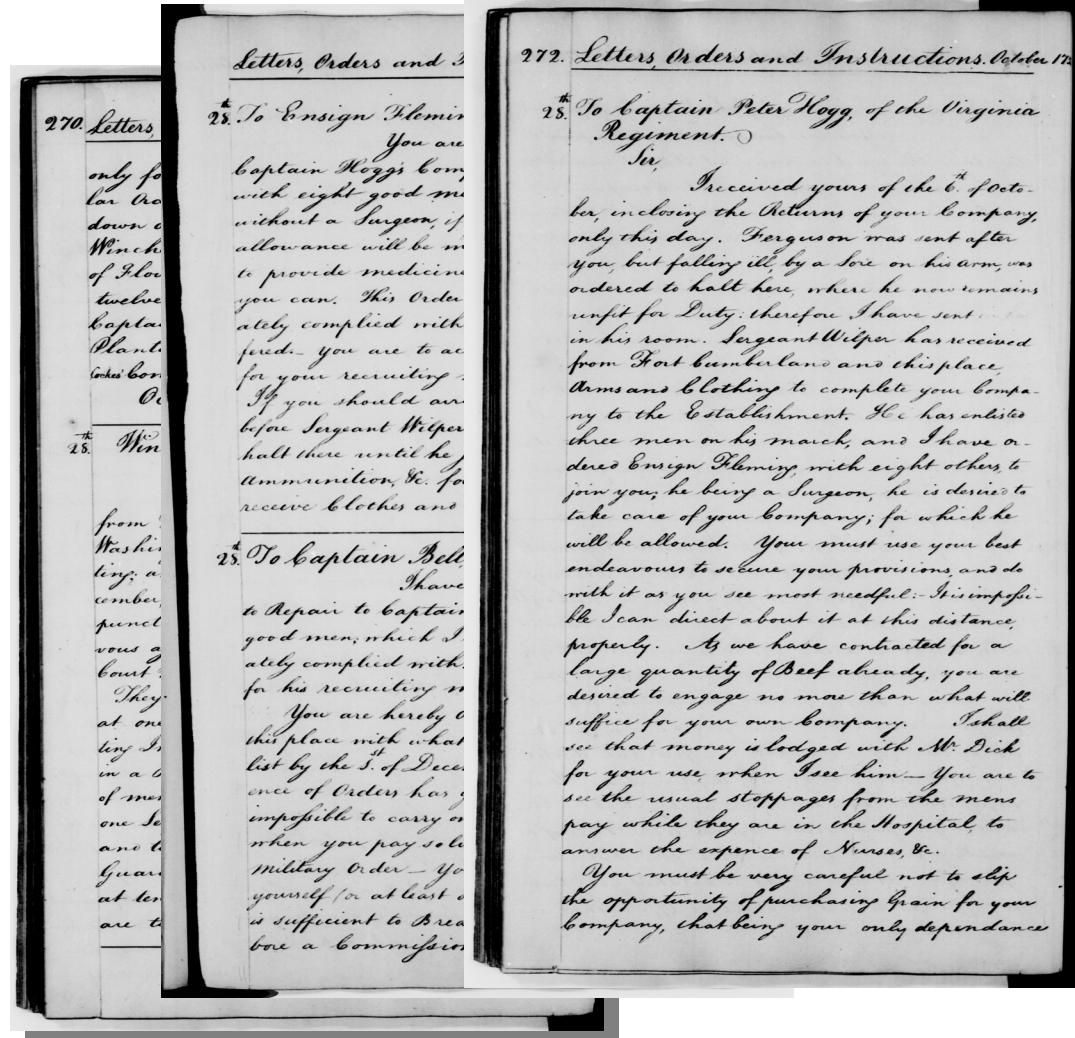
Unser Problem

Letters,

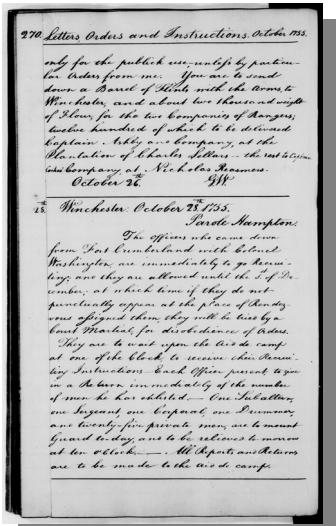
Suchen in:



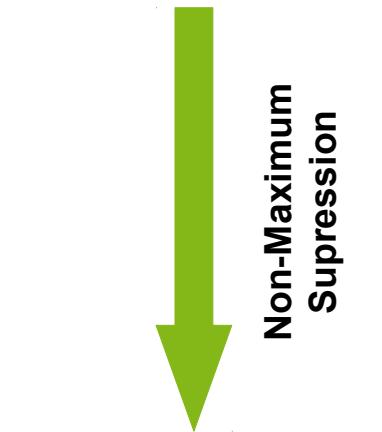
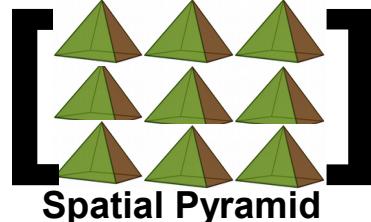
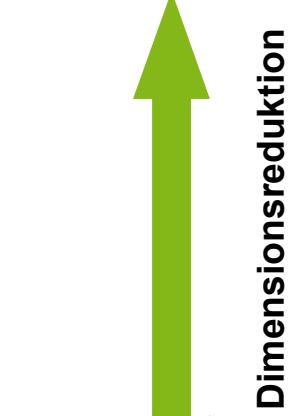
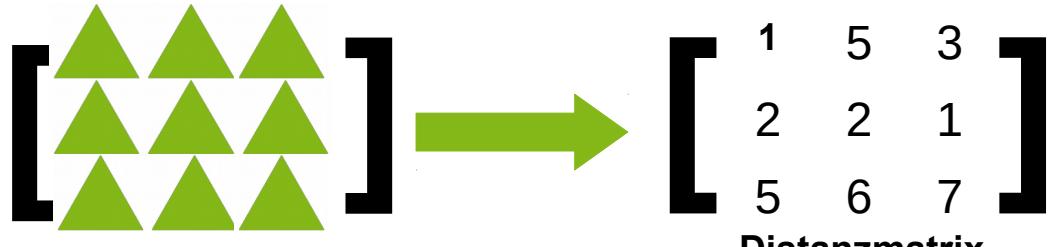
- **Query: Bildausschnitt aus einem Dokument**
- **Ziel: Alle Vorkommen des Querys finden**
- **Patchbasiert**



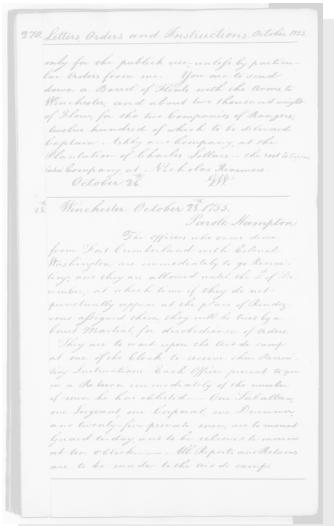
Unser Vorgehen



Document Level



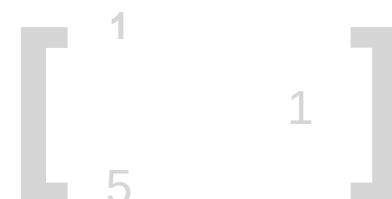
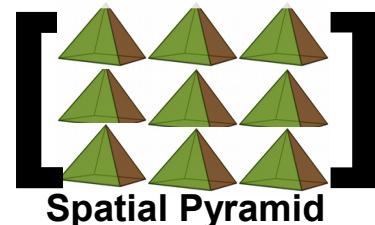
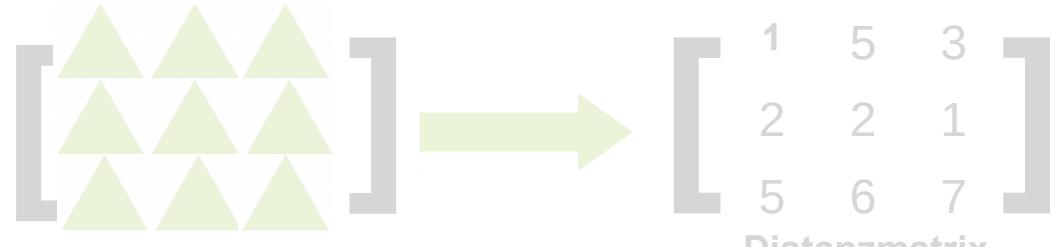
Unser Vorgehen



Document Level

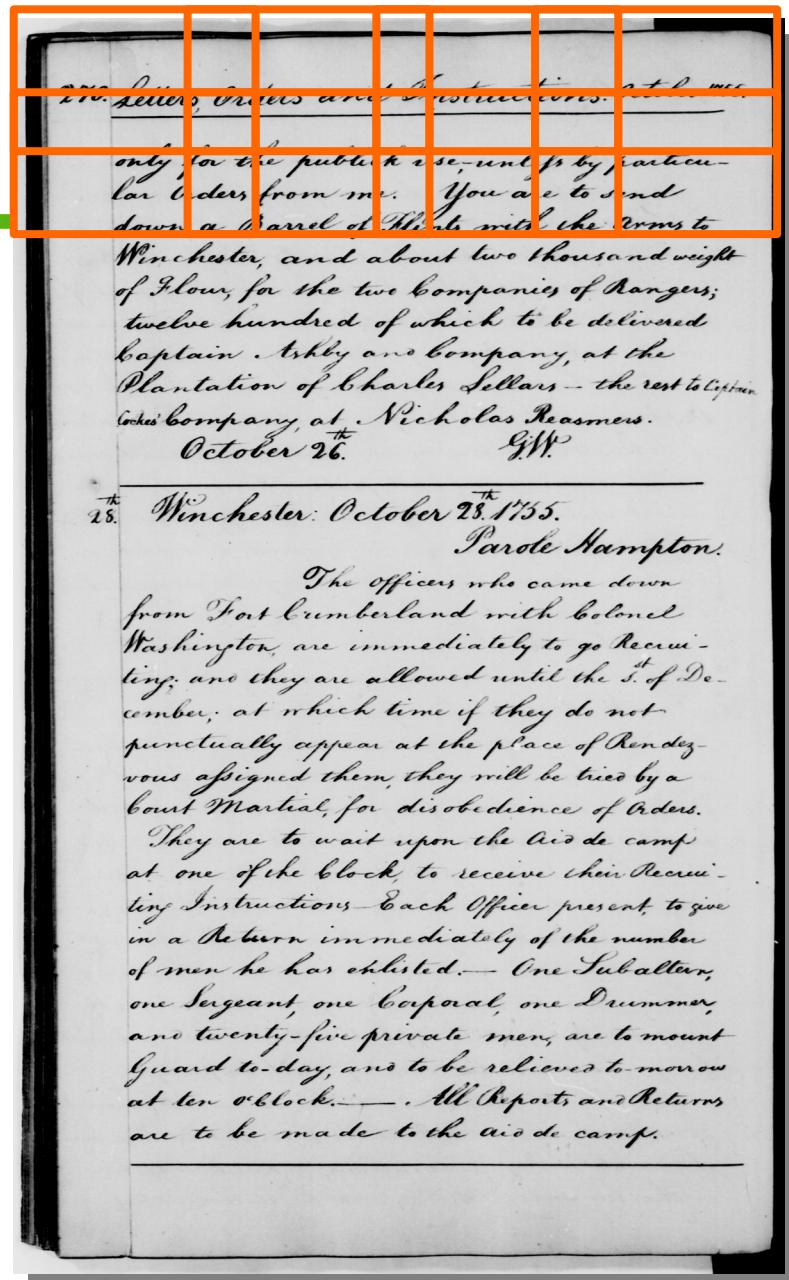


SIFT



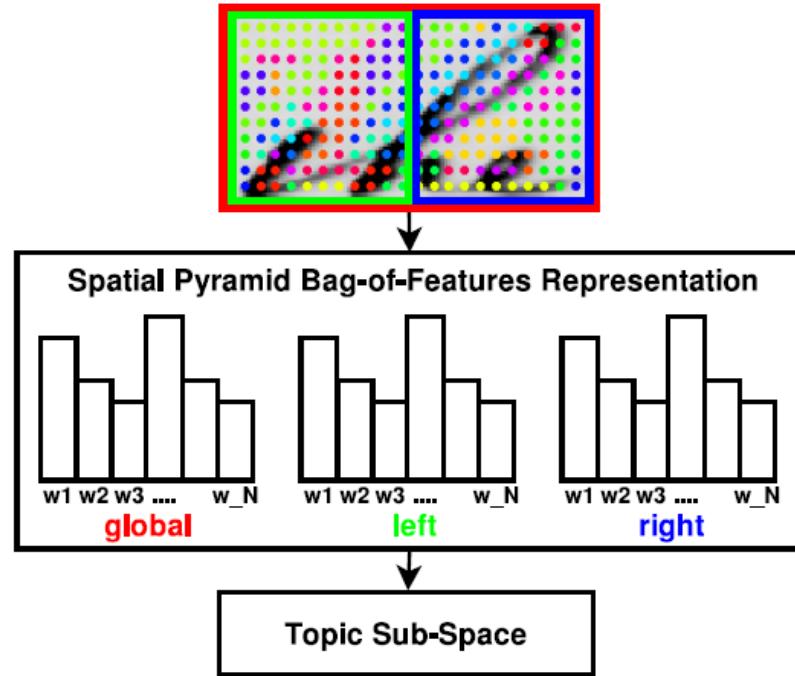
Patches extrahieren

- Einteilung des Bildes in Bereiche
- Überlappung
- Dynamische Patchbreite nach Querybreite
- Patches beinhalten zu Visual Words quantisierte Sift Diskreptoren



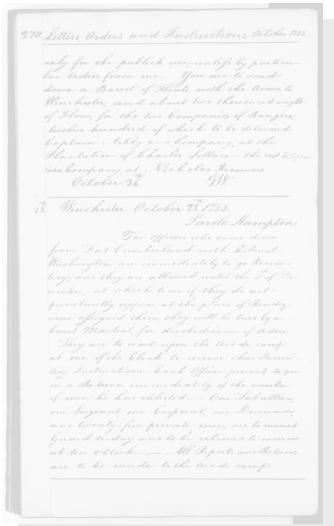
Spatial Pyramid Berechnung

- Drei Histogramme
 - Gesamter Ausschnitt
 - Linker Ausschnitt
 - Rechter Ausschnitt
- Gesamte Spatial Pyramid aus den Histogrammen konkateniert.
- Histogrammgröße: Größe des Codebooks
- Gesamtgröße: $3 \times$ Größe des Codebooks



Rothacker, Leonard; Fink, Gernot: „Dokumentenanalyse 2016“, unter: <http://patrec.cs.tu-dortmund.de/lectures/SS16/dokumentenanalyse/document-analysis.pdf> (abgerufen am 05.07.2016)

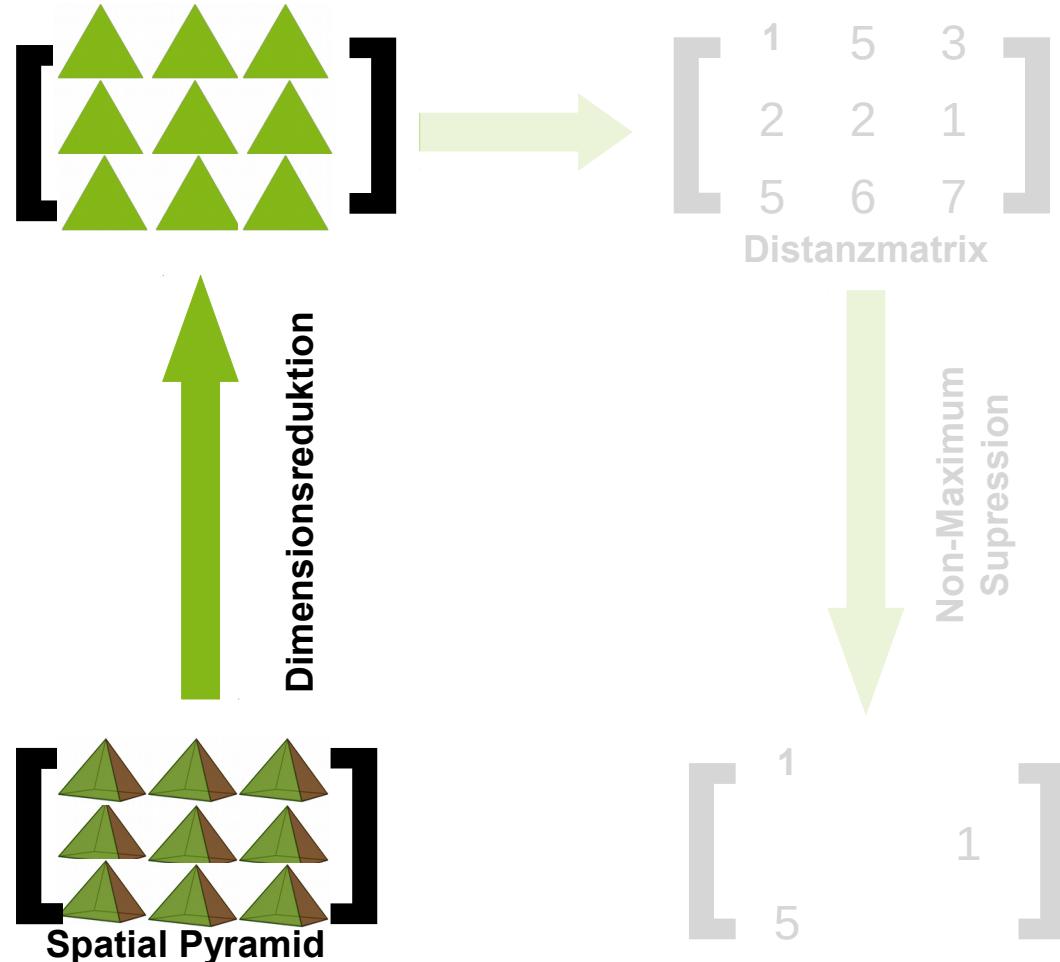
Unser Vorgehen



Document Level



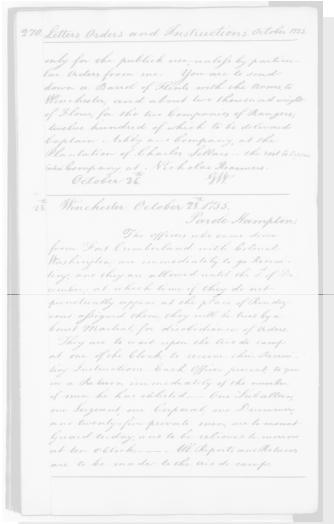
SIFT



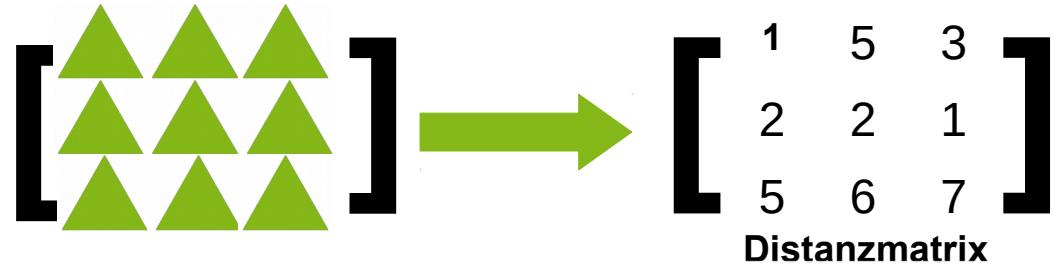
Dimensionsreduktion

- Latent Semantic Indexing
- Reduziert Anzahl der Dimensionen
- Erleichtert Distanzberechnung im nächsten Schritt
- Topic Feature Transform
- Query muss in den gleichen Unterraum transformiert werden

Unser Vorgehen



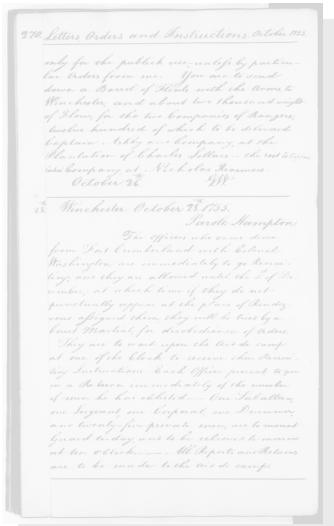
Document Level



SIFT



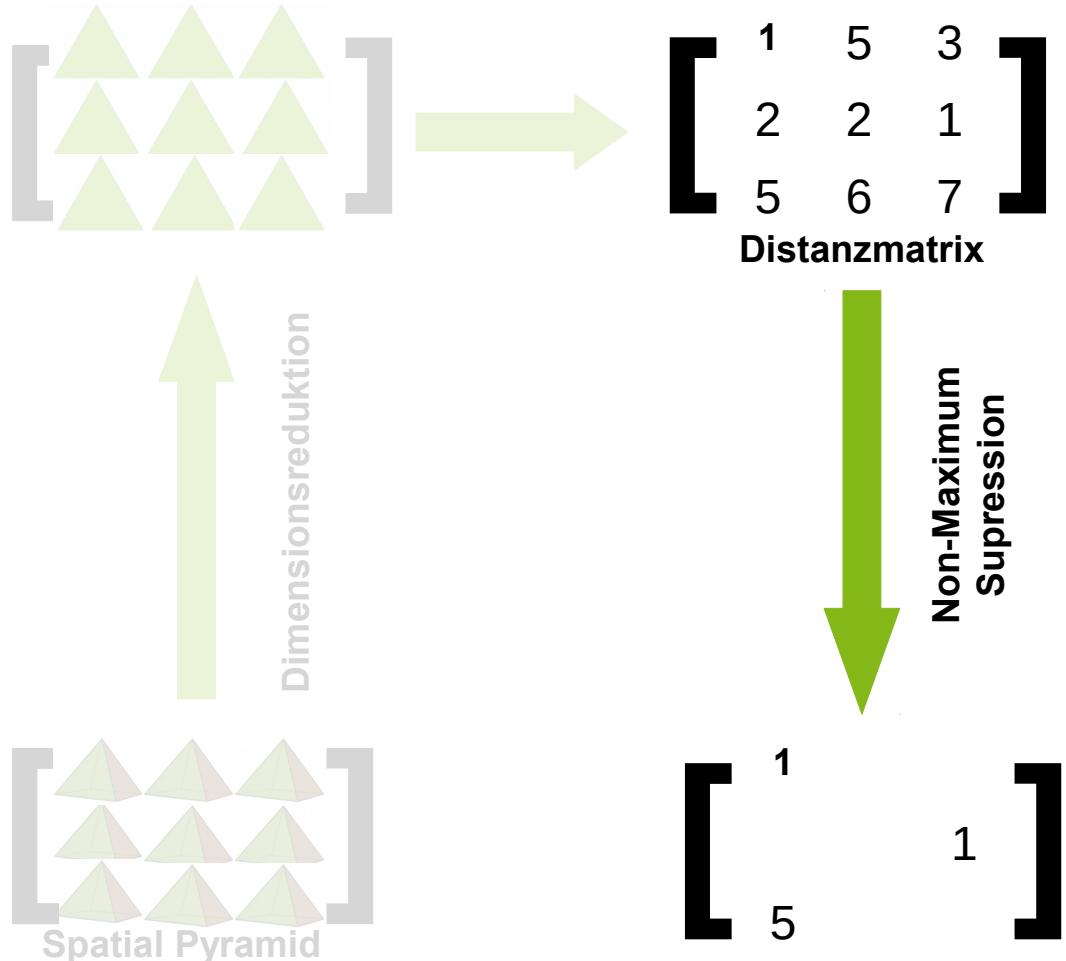
Unser Vorgehen



Document Level

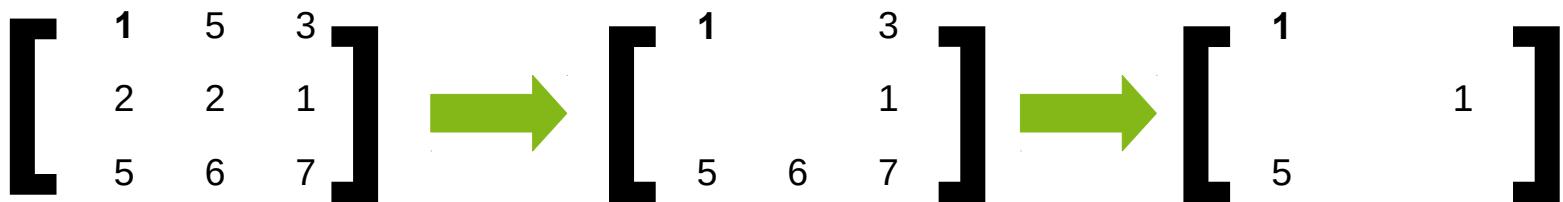


SIFT



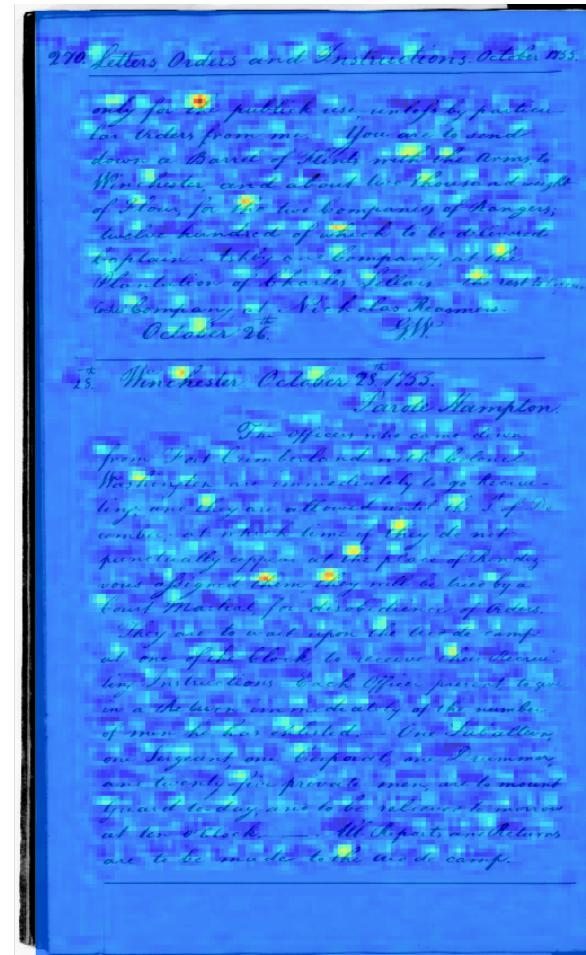
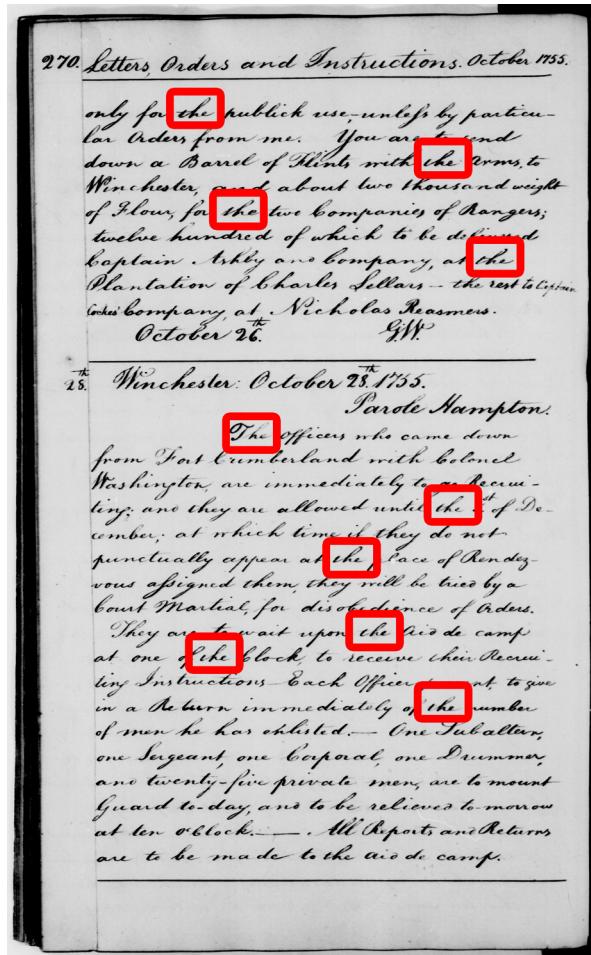
Non-Maximum Suppression

- Patches überlappen sich
- Durch Non-Maximum Suppression wird das beste Patch genutzt

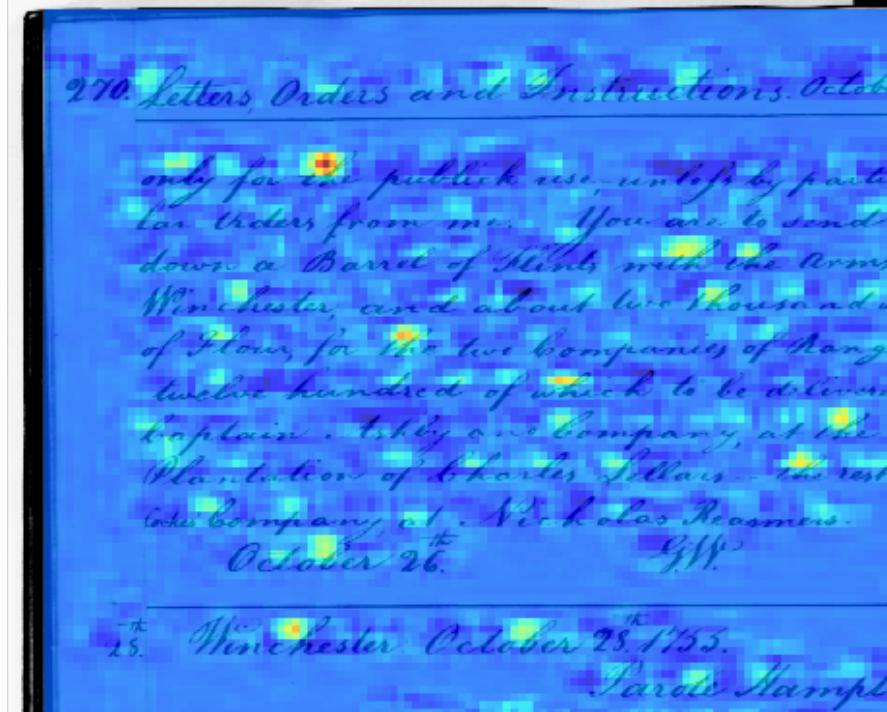
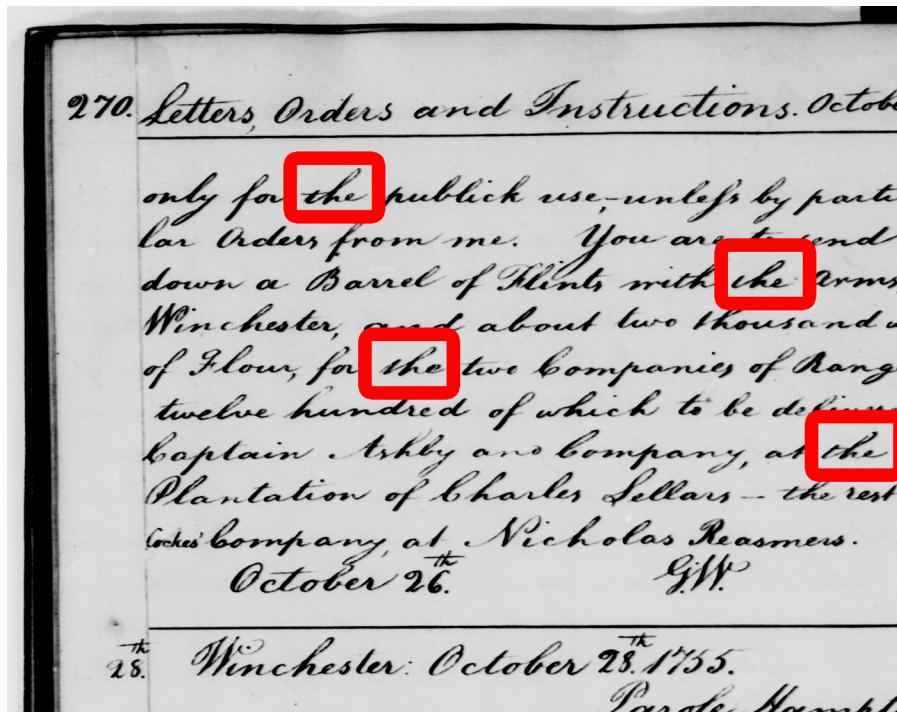


- Danach: Patches nach Distanz sortieren und ausgeben

Vorkommen von „the‘ auf Seite 1

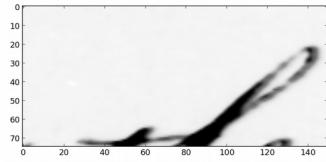


Vorkommen von ‚the‘ auf Seite 1

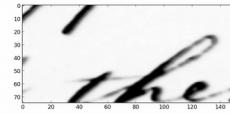


Ergebnisse

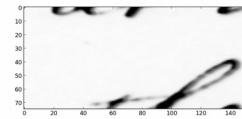
1.



2.



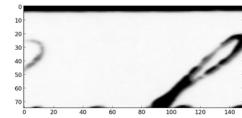
3.



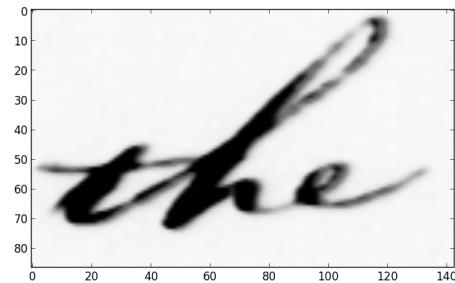
4.



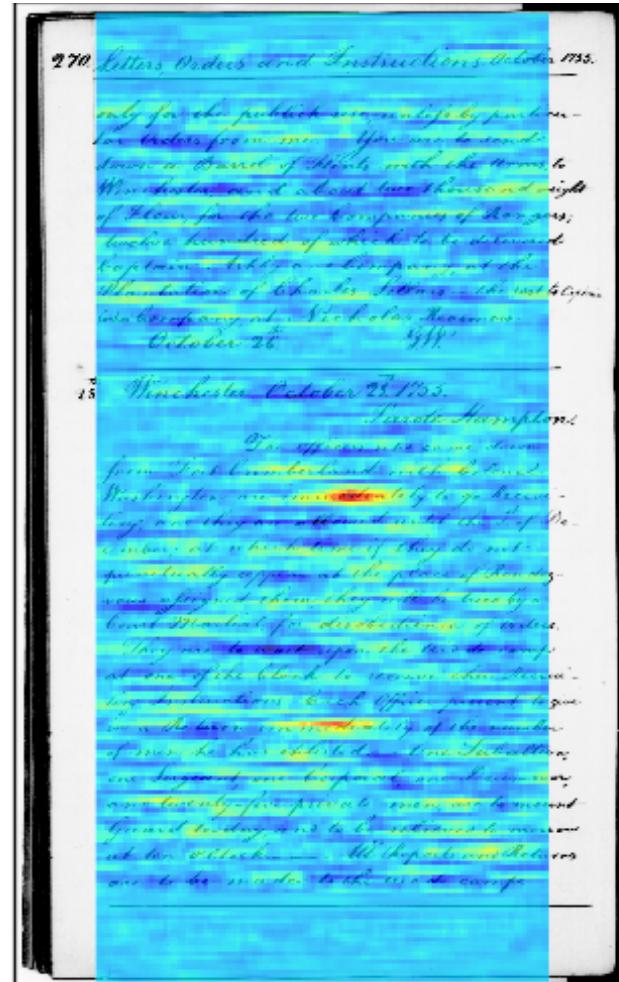
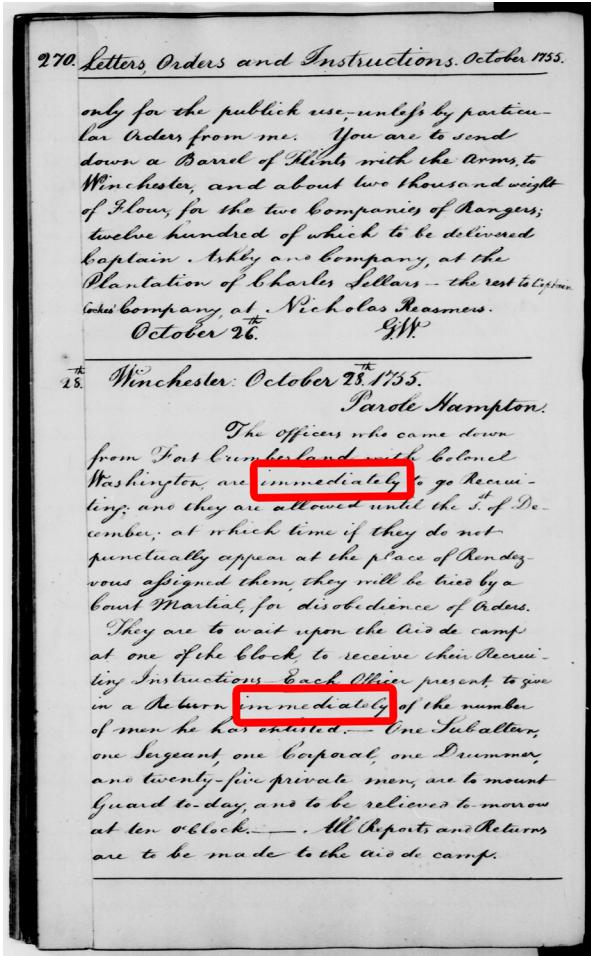
5.



Eingabe Query

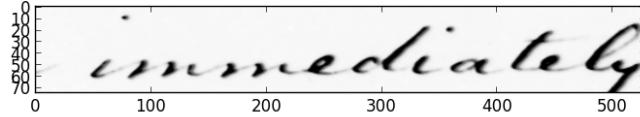


Vorkommen von „immediatly“ auf Seite 1

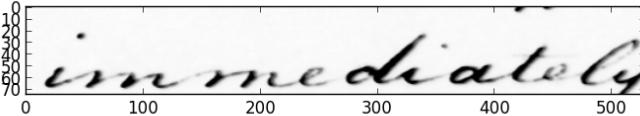


Ergebnisse

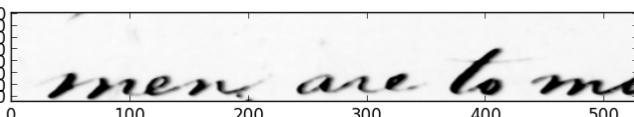
1.



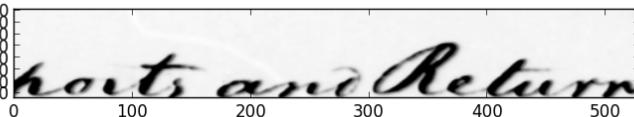
2.



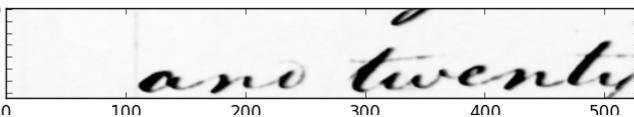
3.



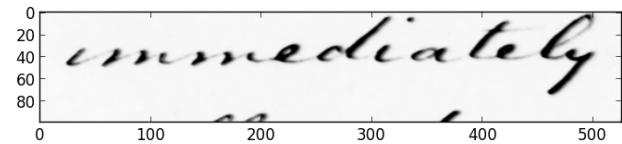
4.



5.



Eingabe Query



Quantitative Evaluation: patch_hop_size variieren

Konfiguration

Sift hop_size	5
Sift cell_size	15
Sift n_visual_words	200
Dimensionen Topic Raum	200
patch_hop_size	10, 20, 25

Resultate

patch_hop_size	mean_recall	mean_avg_precision
10	20 %	12 %
20	22 %	11,9 %
25	19,4 %	10,3 %