

Block Based Hough Transform Mapping for Offline Handwriting Recognition

手書き文字認識に向けた ブロックベースハフ変換マッピング

東北大学大学院工学研究科
電子工学専攻川又研究室
Perttu Pitkänen

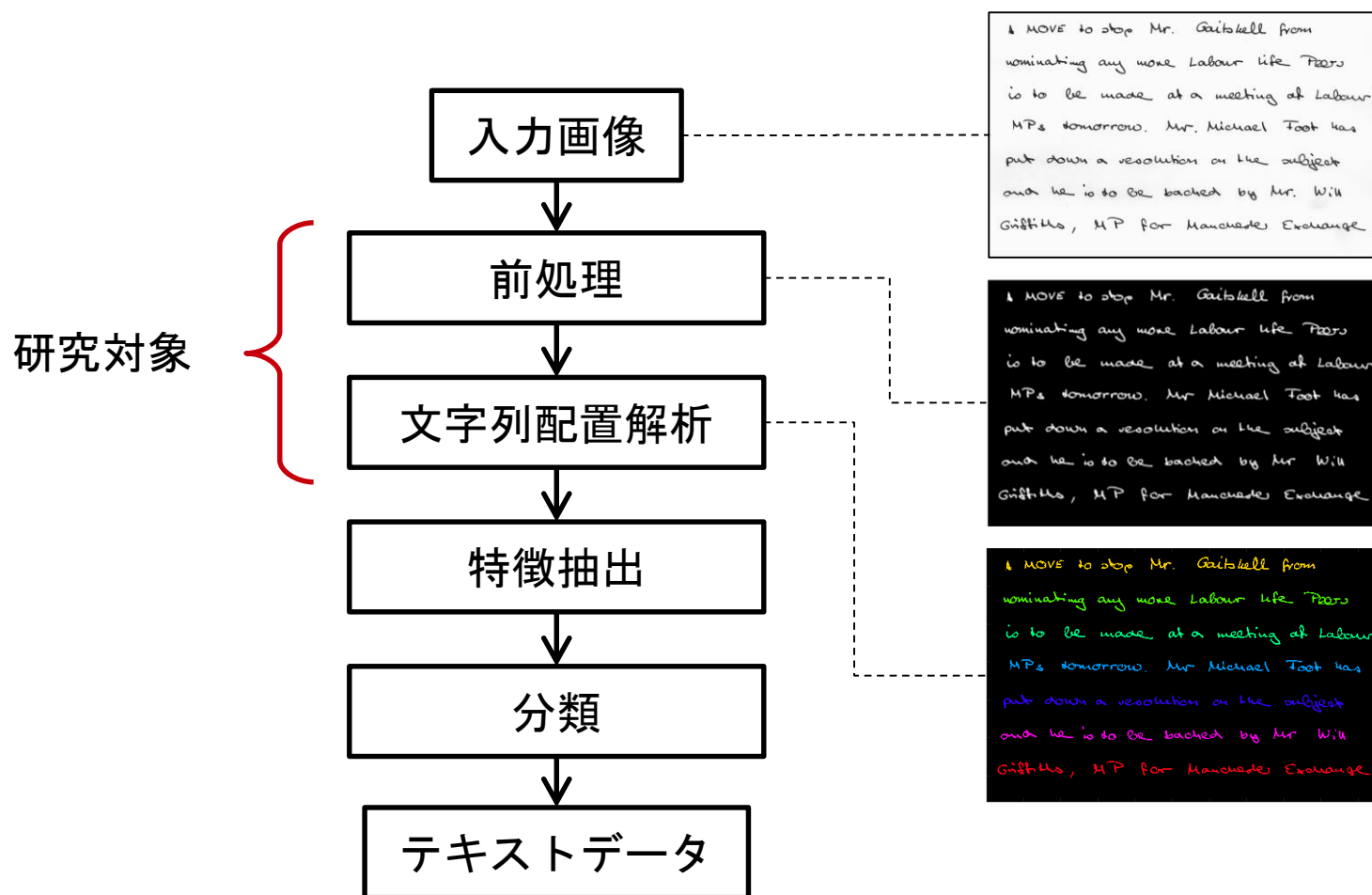
2016年7月14日

アウトライン

- オフライン手書き文字認識について
- ブロックベースハフ変換マッピング
 - 部分集合
 - ハフ変換マッピング
 - 部分集合 2 & 3
 - 追加の制約と手法について
- 実験
- 実験の結果
- まとめ

オフライン手書き文字認識

- オフライン手書き文字認識とは、手書き文字の画像からコンピュータが識別可能なテキストデータを抽出する処理である。



Offline Handwriting Recognition

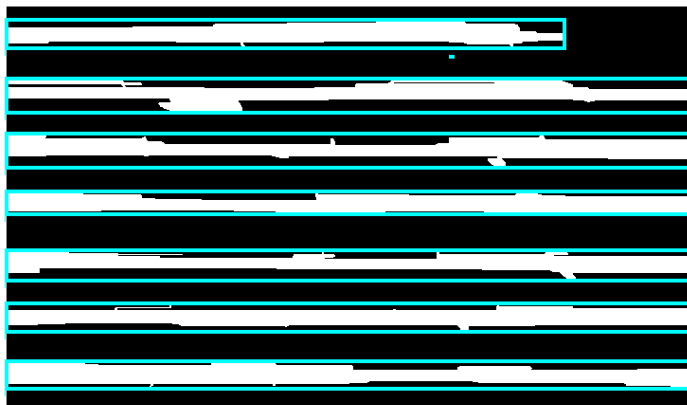
- The goal of *pre-processing* is to provide high quality binarized image by applying filters and binarization to it.
- *Layout analysis* aims to locate the text from image excluding images and figures.
- *Feature extraction* gathers numerical data of the word or character shape.
- *Classification* utilizes feature data and uses machine learning approaches to recognize which word or character current input is.

- Previously implemented features:
 - Pre-Processing:
 - Noise removal
 - Binarization
 - Layout analysis
 - Stroke width analysis
 - RLSA (Run Length Smearing Algorithm)
- Current pre-processing methods and stroke width analysis works well.
- Line detection needed better method than RLSA.
- Most important problem was to implement a good way to detect handwritten lines even when text lines overlap.

Run Length Smearing Algorithm

- Zero pixels in between two one pixels are set to one if their distance smaller than chosen threshold

A MOVE to stop Mr. Gaiskell from
nominating any more Labour life Peers
is to be made at a meeting of Labour
MPs tomorrow. Mr. Michael Foot has
put down a resolution on the subject
and he is to be backed by Mr. Will
Giffiths, MP for Manchester Exchange.

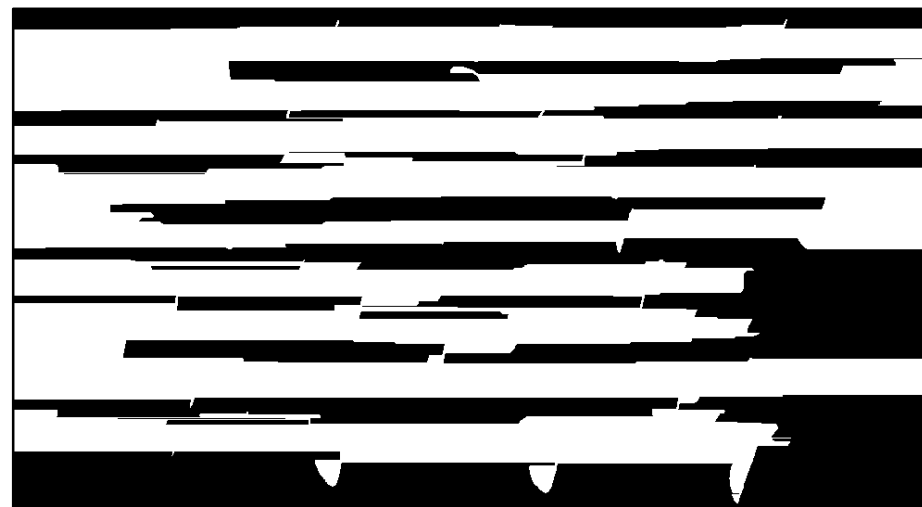


A MOVE to stop Mr. Gaiskell from
nominating any more Labour life Peers
is to be made at a meeting of Labour
MPs tomorrow. Mr. Michael Foot has
put down a resolution on the subject
and he is to be backed by Mr. Will
Giffiths, MP for Manchester Exchange.

Problem with RLSA

- Run Length Smearing Algorithm was very fast and could detect lines and words if they were completely separated and the handwriting had very clear layout.
- With overlapping lines or lines with skew the RLSA could not detect lines correctly
- More sophisticated approach was needed.

Delegates from Mr. Kenneth Kaunda's United National Independence Party (280'000 members) and Mr. Harry Nkumbula's African National Congress (400'000) will meet in London today to discuss a common course of action. Sir Roy is violently opposed to Africans getting an elected majority in Northern Rhodesia, but the Colonial Secretary, Mr. Iain Macleod, is insisting on a policy of change.



Only 2 "lines" detected

Block Based Hough Transform Mapping

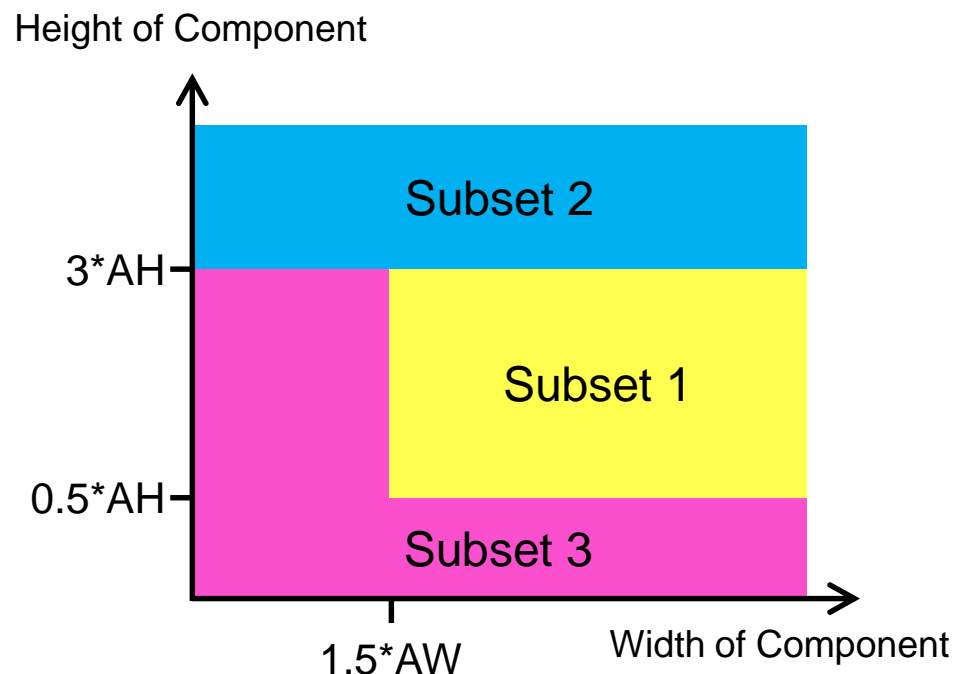
- Method proposed by Louloudis et.al. To detect handwritten text lines and words from images.
 - + Especially designed for handwritten text.
 - + Can detect lines with some variation in skew.
 - + Can segment overlapping characters.
 - + Outperforms other line detection methods.
 - Can natively detect only single column data.
 - No open source implementation available.
 - Complex methods → Relatively slow execution time.
- The method was implemented in MATLAB for this research.

Subsets

- Objects were categorized into three subset according to their relative size.
- Subset 1: Majority of characters
($0.5 * AH \leq H < 3 * AH$)
and ($0.5 * AW \leq W$)
- Subset 2: Large characters, often characters overlapping multiple lines.

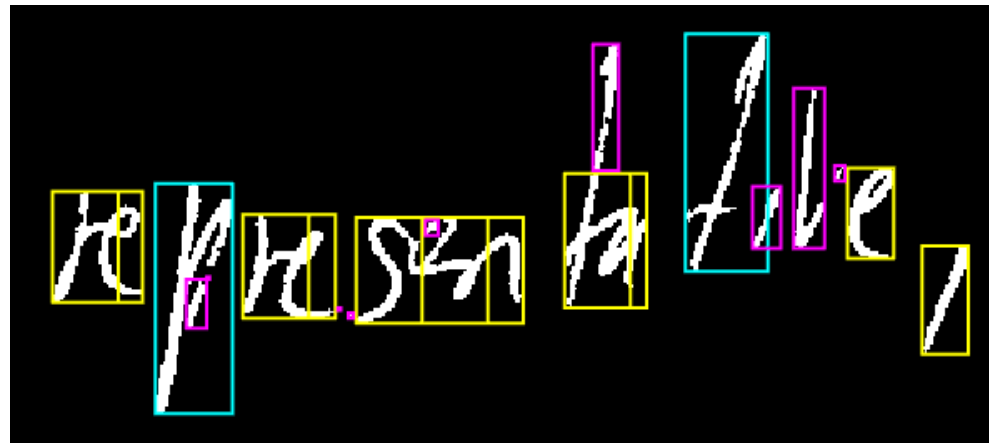
$$H \geq 3 * AH$$


- Subset 3: Small characters and accents.
($(H < 3 * AH) \text{ and } (0.5 * AW < W)$) or
($(H < 0.5 * AH) \text{ and } (0.5 * AW < W)$)

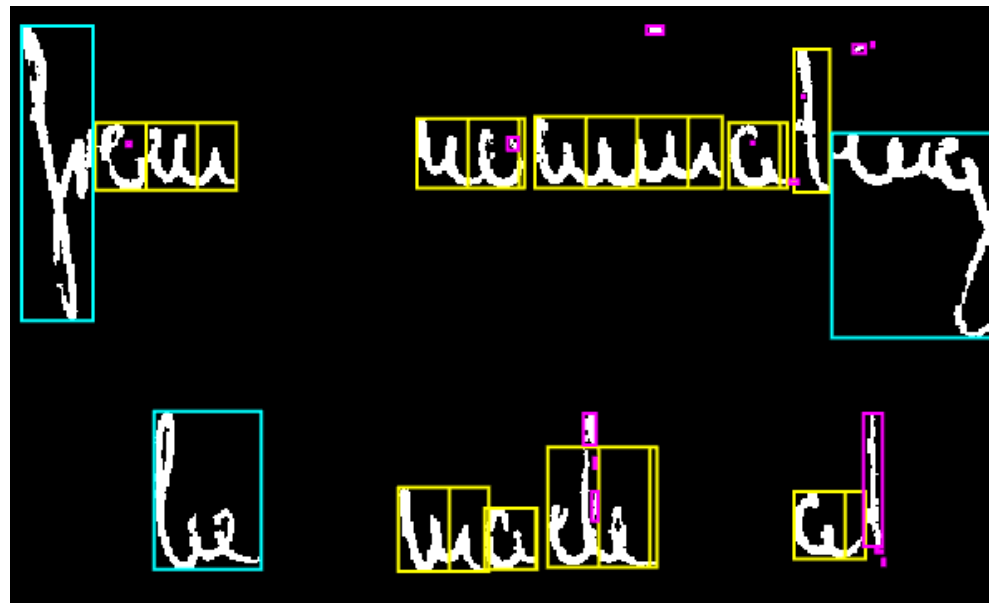


- H: Height of Component
- W: Width of Component
- AH: Average Height
- AW: Average Width

Examples of Subset Characters

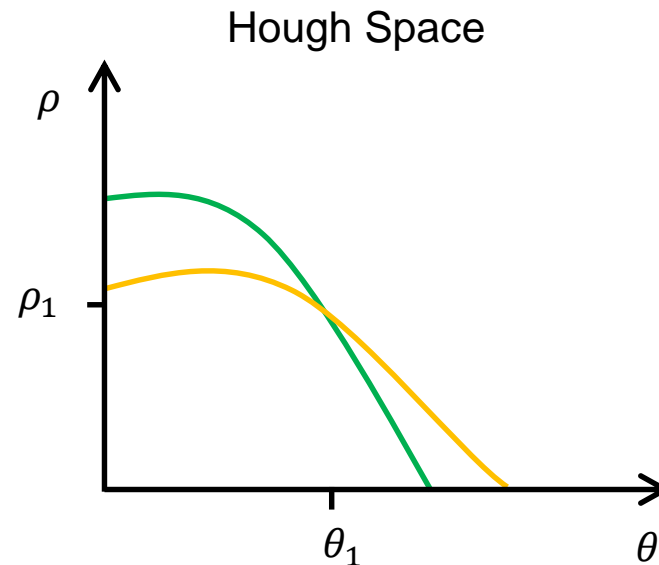
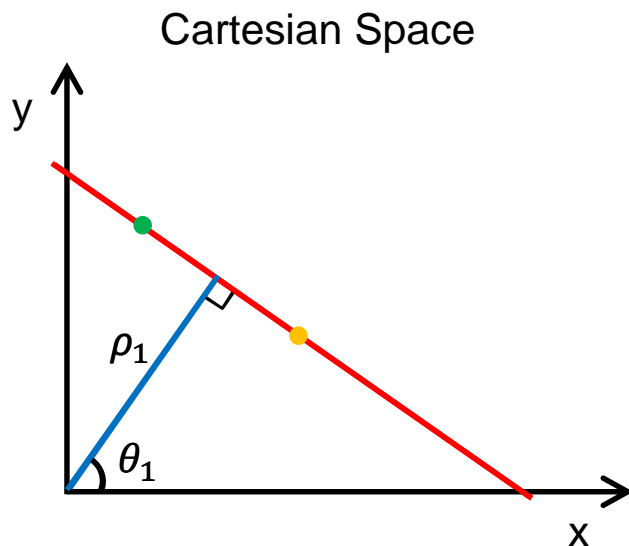


-  Subset 1
-  Subset 2
-  Subset 3



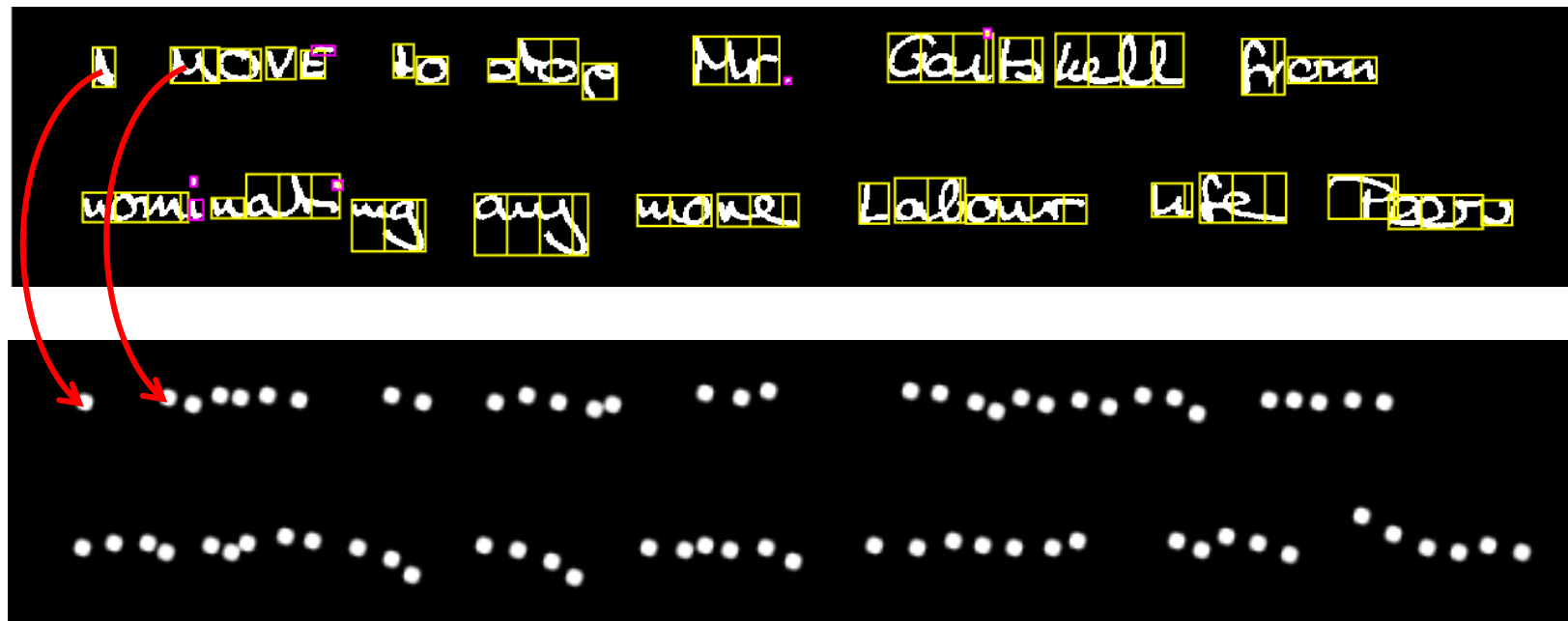
Hough transform

- Main functionality of the method is based on the Hough Transform
- Hough transform is algorithm designed to find lines in Cartesian space.
- Any line can be described with ρ and θ using form:
$$\rho = x \cos \theta + y \sin \theta$$

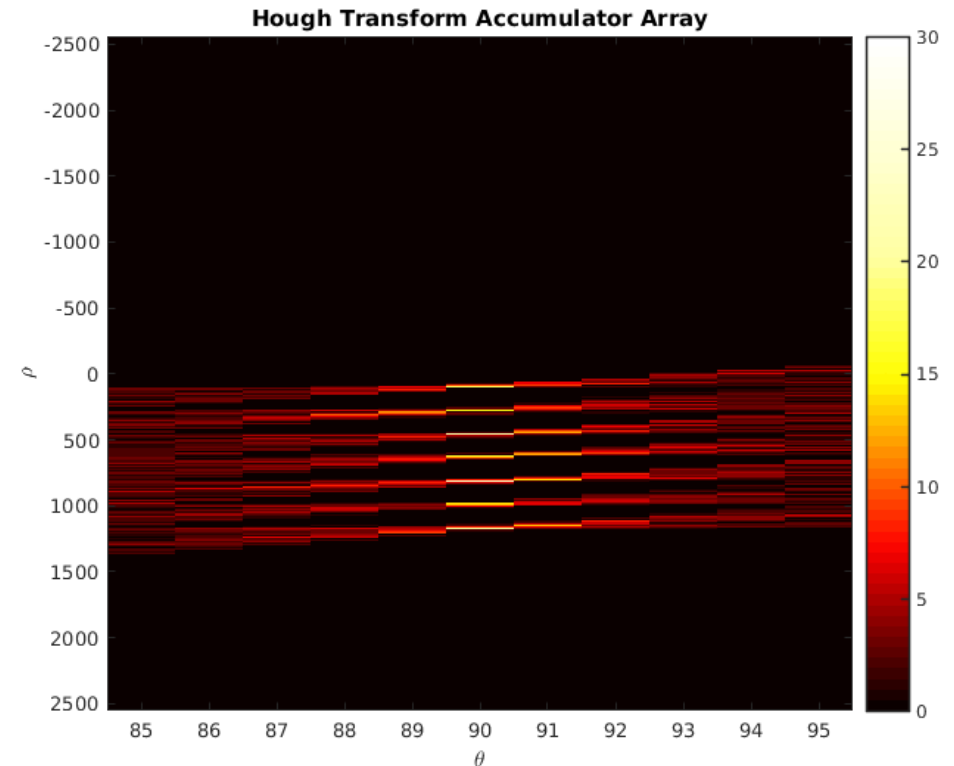
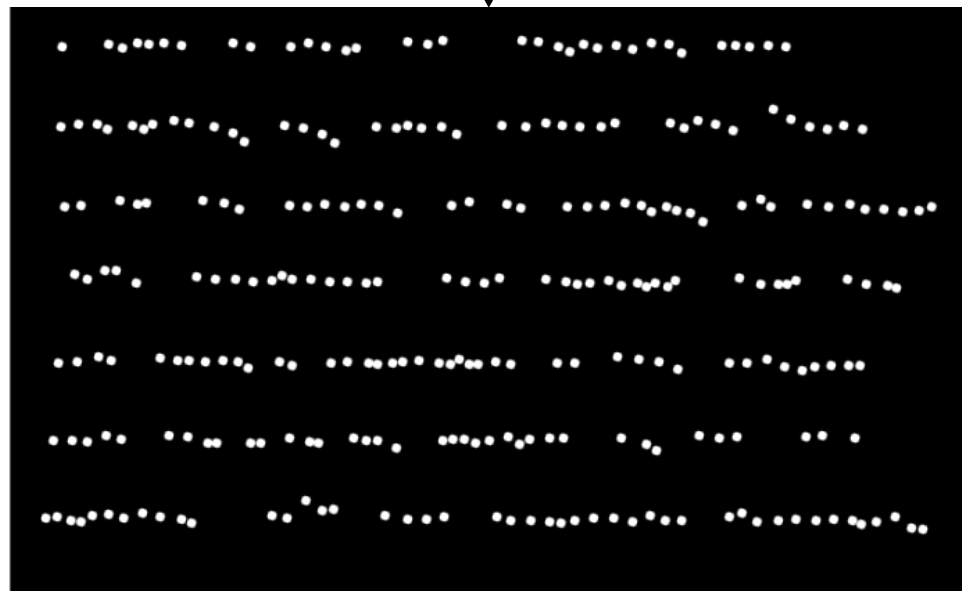
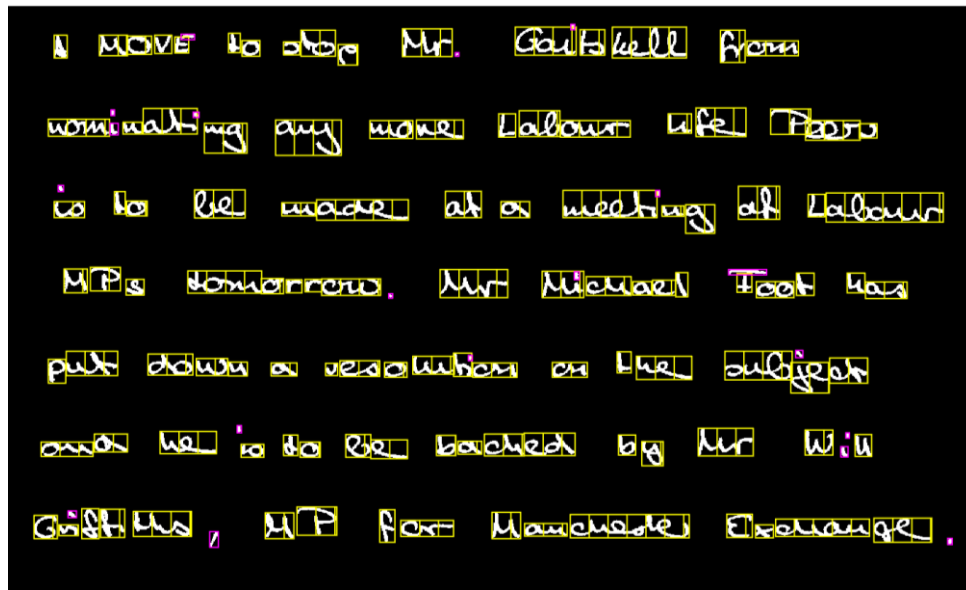


Hough Transform

- Subset 1 characters are partitioned into average width sized blocks.
- The centroid of each of these blocks is used as datapoints for Hough Transform.

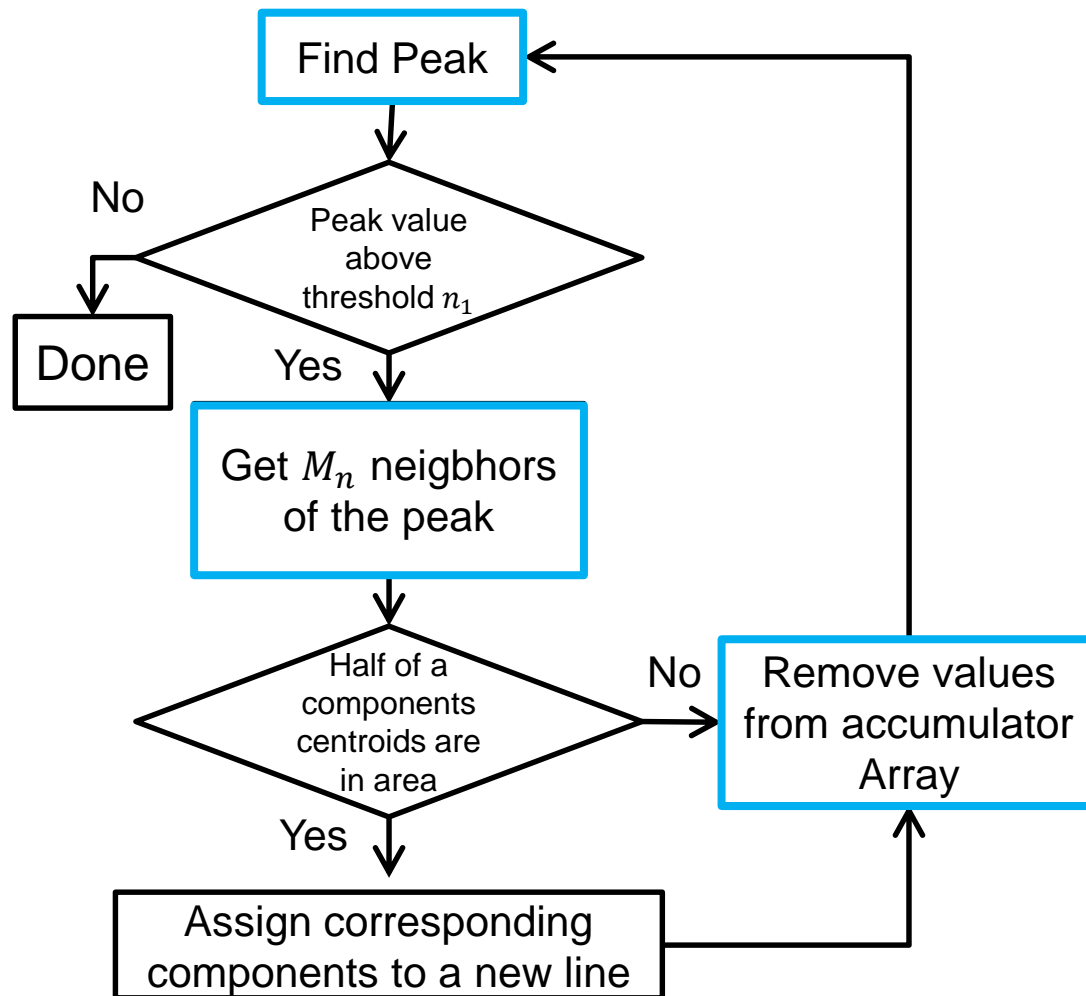


Data Points and Corresponding Accumulator Array



Hough Transform

Line Extraction from Accumulator Array



Accumulator array

0	0	14
0	0	6
0	0	8
0	0	2
2	0	3
11	0	8
10	0	13
17	0	4

ρ bin

θ bin

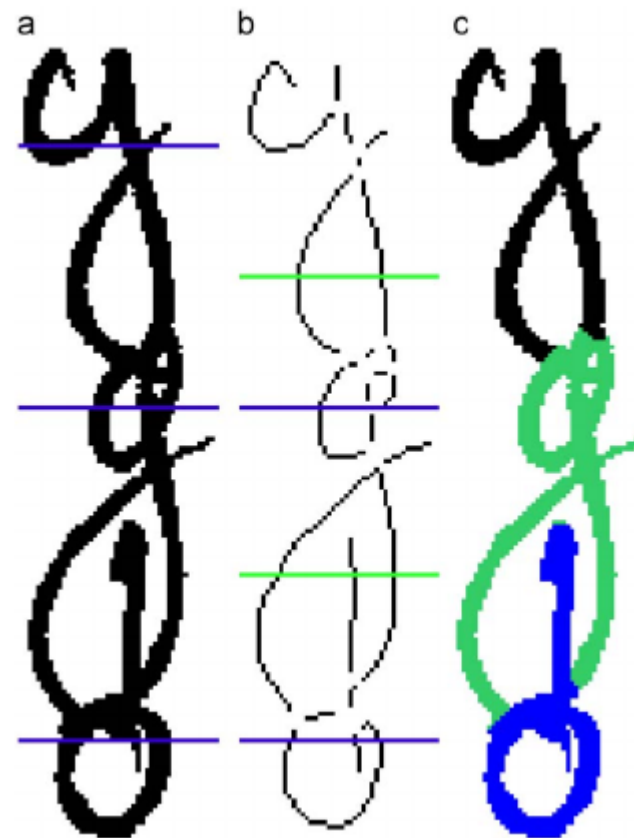
$M_n = 3$

n_1 Defines the minimum peak of Hough accumulator Array

M_n How many neighboring components in ρ direction are assigned to same line.

Subset 2 & 3

- Subset 3 objects are assigned to closest line if they are less than average distance from it.
- Splitting procedure is executed for Subset 2.
 - If two or more lines intersect Subset 2 component it can be splitted.
 - Splitting is done in the objects sekeletons intersection points above intersecting lines except the highest line.



a: Subset 2 component

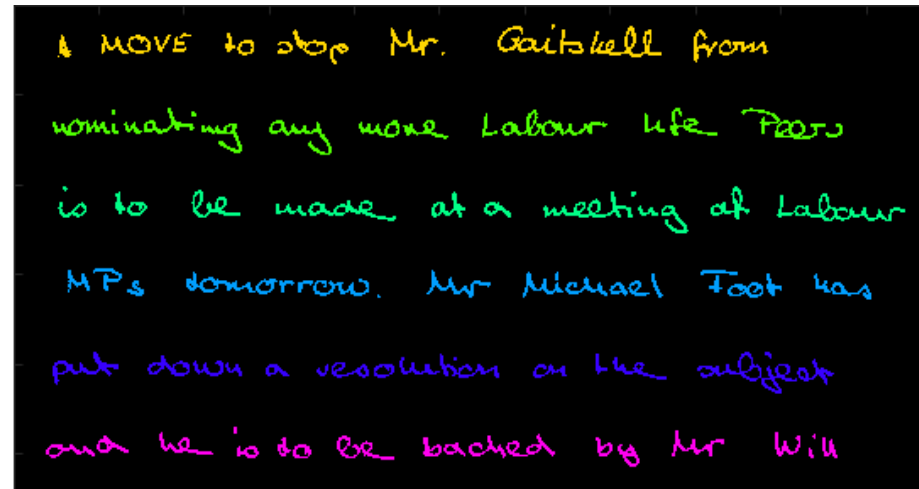
b: Skeleton image of the component

c: Successfully splitted component

Additional Constraints and Techniques

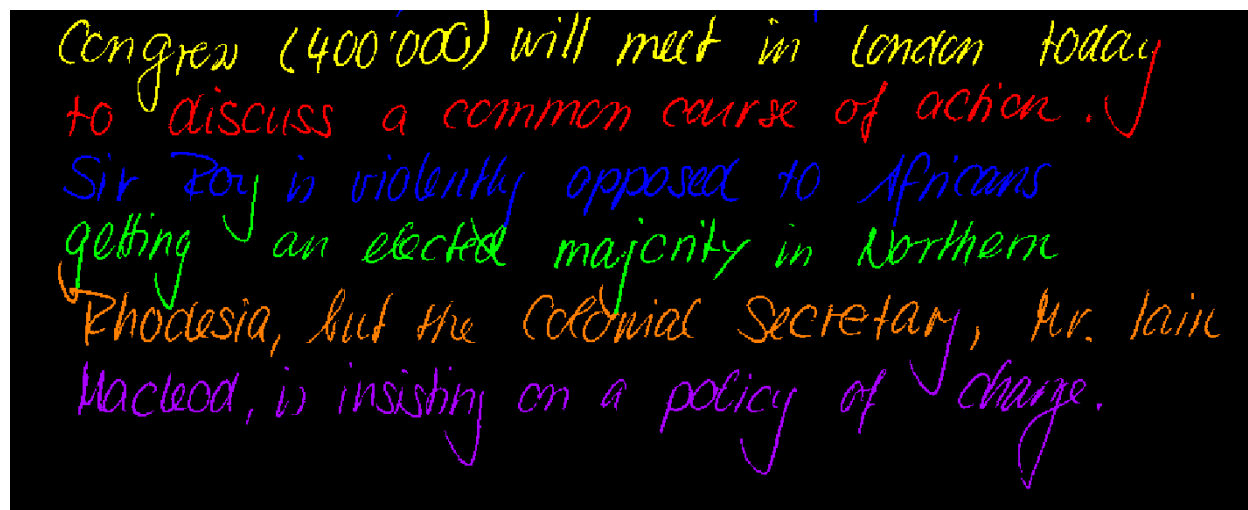
- Components that weren't assigned to any line must be assigned to the nearest line if they are close to it.
- If a line has only small amount of block centroids and its skew is larger than threshold, the line is discarded.
- Word detection was proposed in the article by Louloudis et.al. but it was not implemented due to lack of time.

Output



A MOVE to stop Mr. Gaitskell from
nominating any more Labour life Peers
is to be made at a meeting of Labour
MPs tomorrow. Mr Michael Foot has
put down a resolution on the subject
and he is to be backed by Mr Will

Successful line detection



Congress (400'000) will meet in London today
to discuss a common course of action. ✓
Sir Roy is violently opposed to Africans
getting ✓ an elected majority in Northern
Rhodesia, but the Colonial Secretary, Mr. Iain
Macleod, is insisting on a policy of ✓ change.

Line detection which shows some remaining problems

Evaluation

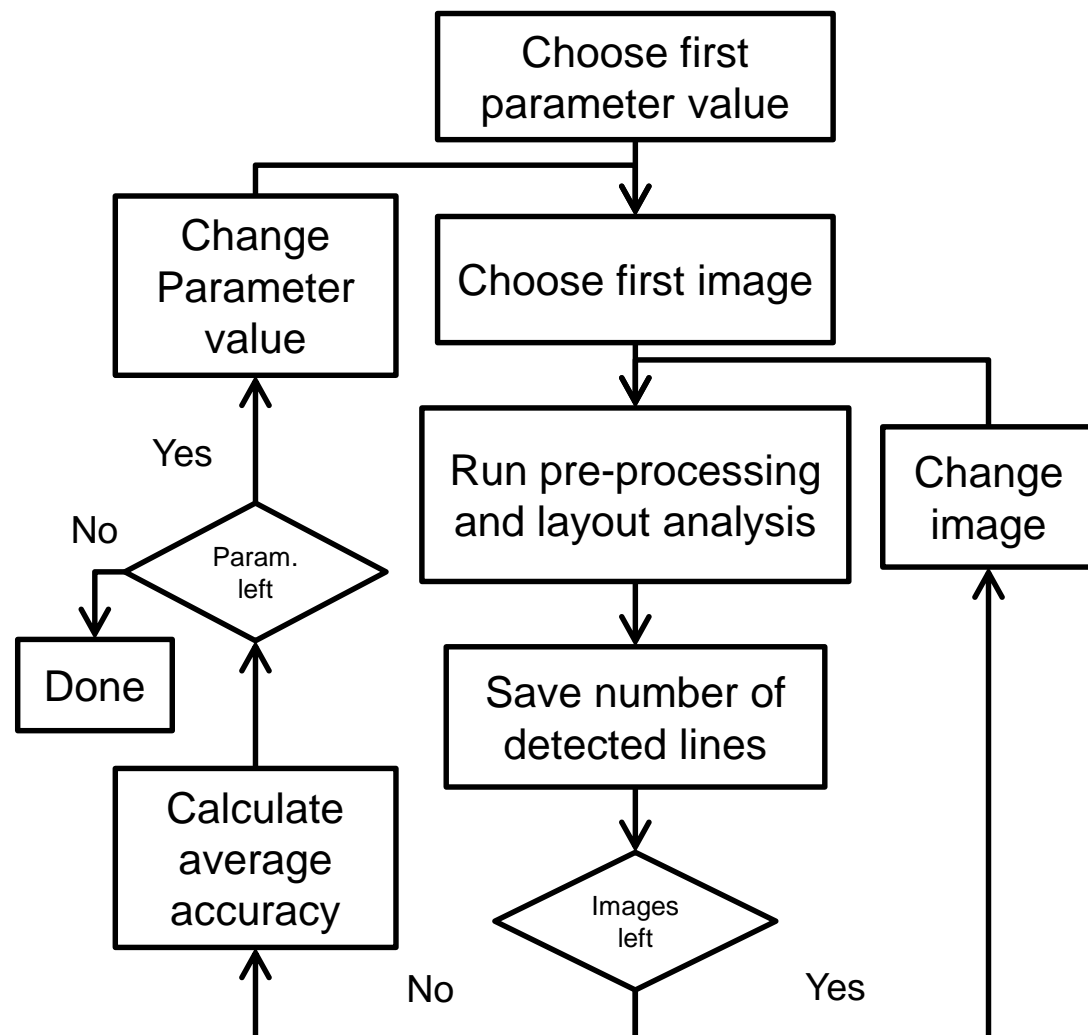
- Current method depends on multiple parameter values.
- Their effect was studied and optimal values were chosen according to these test results.
- Word detection was not implemented at this point.
- Only the number of lines could be used as a metric to measure the current implementations accuracy.
- The IAM Handwriting Database provided the image files and metadata suitable for these tests.

Evaluation Procedure

- Two nested loops:
 - Outer iterates through parameter values (usually 10-20 values)
 - Inner iterates through 100 random images from handwriting database.
- The average accuracy among these 100 images was calculated and saved to determine the accuracy of one parameter value.
- I^P iterations was needed for each tested parameter. I is the number of images and P is the number of parameters.
- Accuracy was calculated for each image with formula:

$$A = 1 - \frac{|L_r - L_d|}{L_r}$$

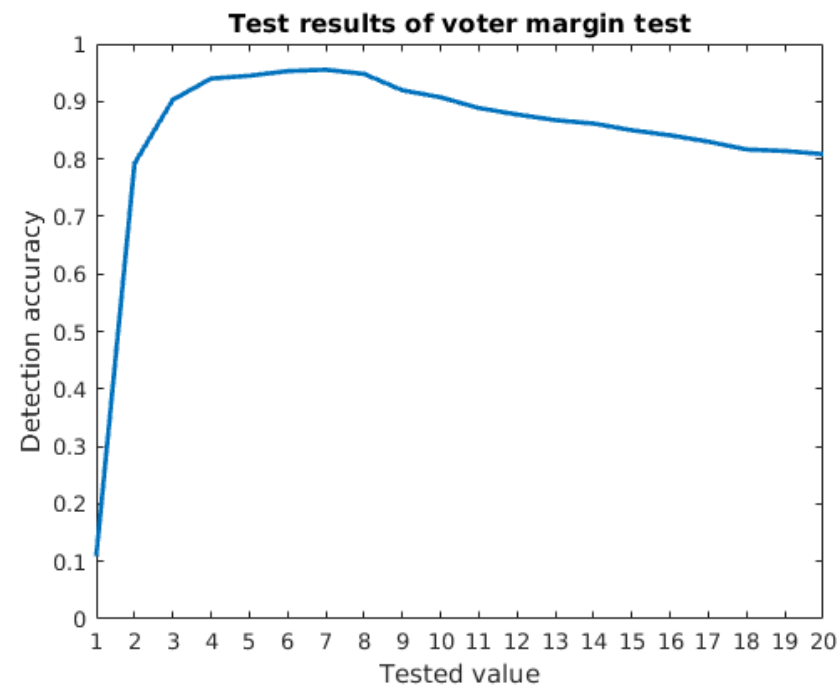
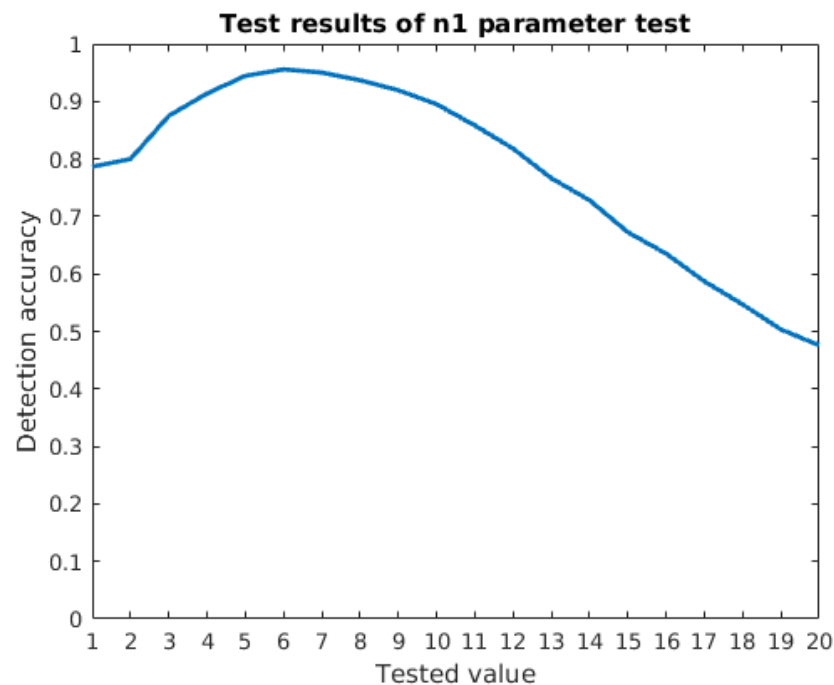
Where L_r is the number of real lines and L_d is the number of detected lines.



Tests

- All parameters were tested for IAM Handwriting database in mind.
- Parameters from pre-processing and layout analysis methods were evaluated.
- Best parameters were chosen according to tests.
- If test didn't show a noticeably better parameter value the parameter was chosen by visually inspecting the output or according to corresponding scientific article.

Result Samples



- These two parameters had the most effect on the line detection accuracy.
- n_1 parameter defines the minimum peak value from Hough accumulator array.
- Voter margin M_n determines How many neighboring components in ρ direction are assigned to same line.
- Often results weren't this prominent. More precise tests are needed.

Conclusions

- The method gained around 97 % accuracy regarding number of lines.
- Many parameter values didn't have prominent effect on the amount of lines.
- Remaining problems:
 - Line number can be right even if some components are assigned to wrong lines.
 - Given method can only detect single column text.
 - Word detection would provide more accurate tests and results.

Summary

- Block based Hough transform mapping promised good performance when segmentating handwritten lines.
- Implementation for the algorithm was completed to detect lines.
- When tested, the method achieved 97% accuracy in regard of number of lines.
- Implementation of the word detection feature and more accurate tests are needed.