

Meeting Stuttgart– 22 October, 2007

Dr. ir. Jan C.A. van der Lubbe

J.C.A.vanderLubbe@TUDelft.nl

M.Sc. Héctor Moreu Otal

H.MoreuOtal@TUDelft.nl

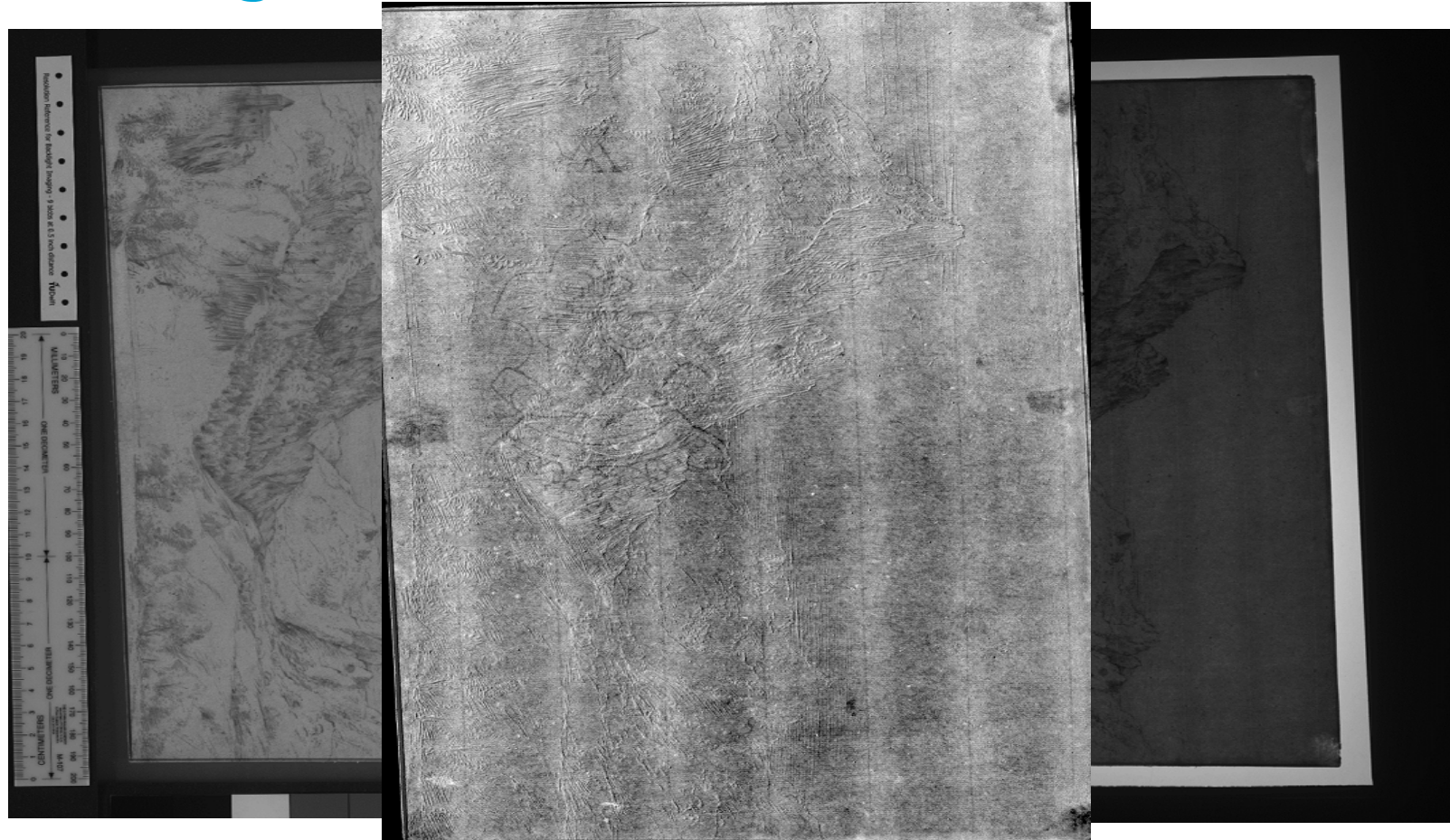
Have-Done List

1. Backlight subtraction
2. X-ray post-processing (resolution and mask)
3. Paper analysis (chain and laid lines)
4. Paper retrieval based on chain and laid lines descriptors

Current work

1. Automatic Watermark detection in gray-scale images (X-ray, Backlight)
2. Printed Piccard Vs. Piccard online. Watermark extraction and retrieval.

Backlight subtraction



Reflected Image

Transmitted image

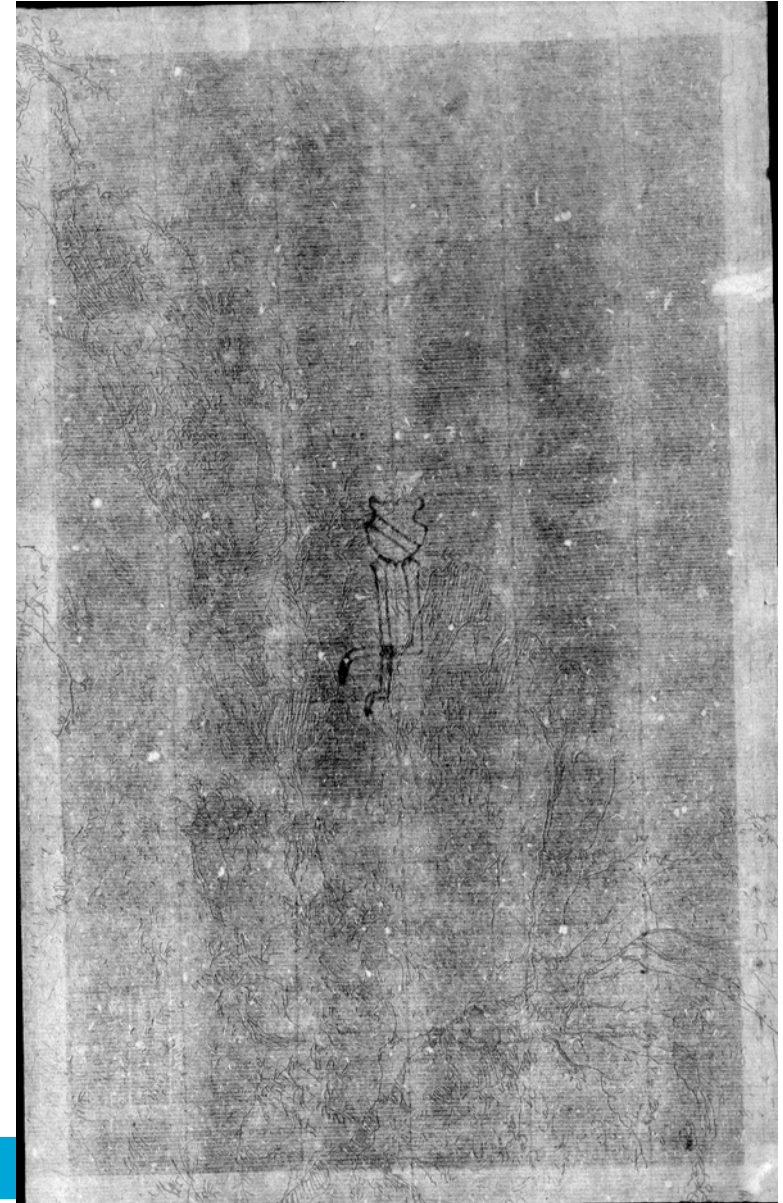
Backlight subtraction

1. Obtain normalized paper structure (~ Combination of reflected and transmitted image)
2. Estimate resolution based on the ruler. Compute paper structure with the right resolution (150 dpi)
3. Select mask in order to remove the paper parts which do not represent paper features

Matlab GUI

X-ray post-processing

1. Transform the original x-ray image to a 150 dpi resolution image
2. Compute a mask in order to remove the parts of the paper with no paper structure information



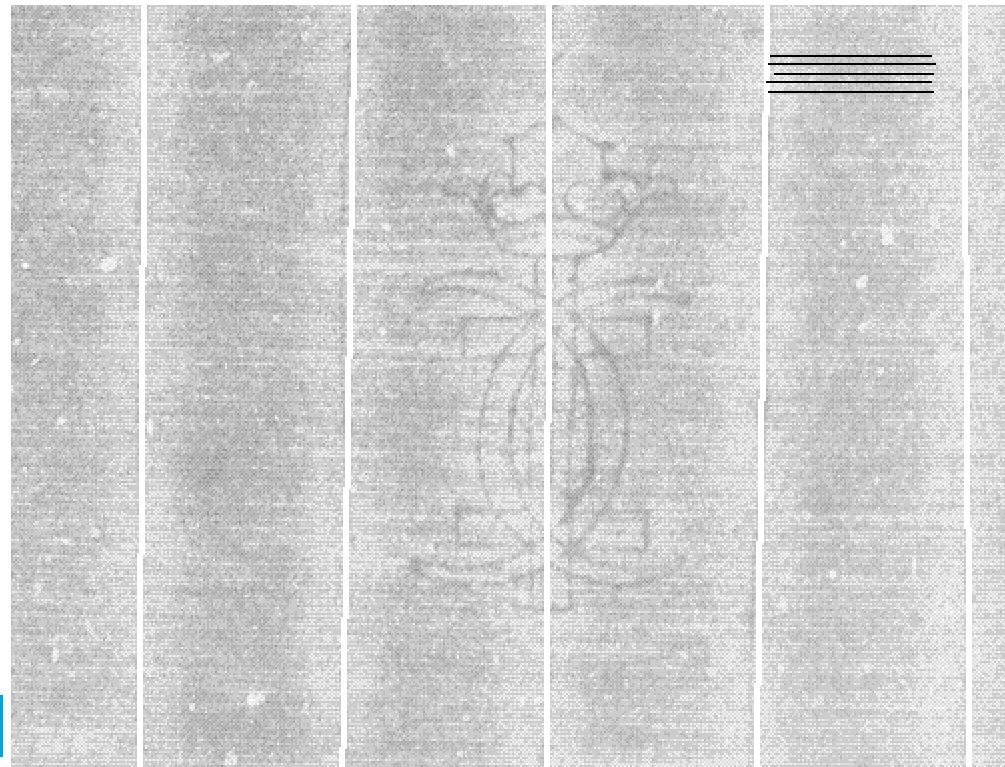
29 October, 2007

$[I, C]^T$

Paper Analysis (chain and laid lines)

1. Detect chain and laid lines
2. Compute chain and laid lines density

Matlab GUI



29 October, 2007

$[I, C]^T$

Paper retrieval based on chain and laid lines descriptors

1. Every piece of paper is characterized by its laid and chain lines descriptors
 - Laid lines descriptor: laid lines density, orientation and peak values
 - Chain lines descriptor: chain lines position, density and orientation.
2. Build database with the paper structure information

Paper retrieval based on chain and laid lines descriptors

1. Compute the distance between the database descriptors and query descriptors. The minimum distance corresponds to the most probable matching
2. Web user interface. The user confirms the matchings.

<http://rembrandt.ewi.tudelft.nl/>

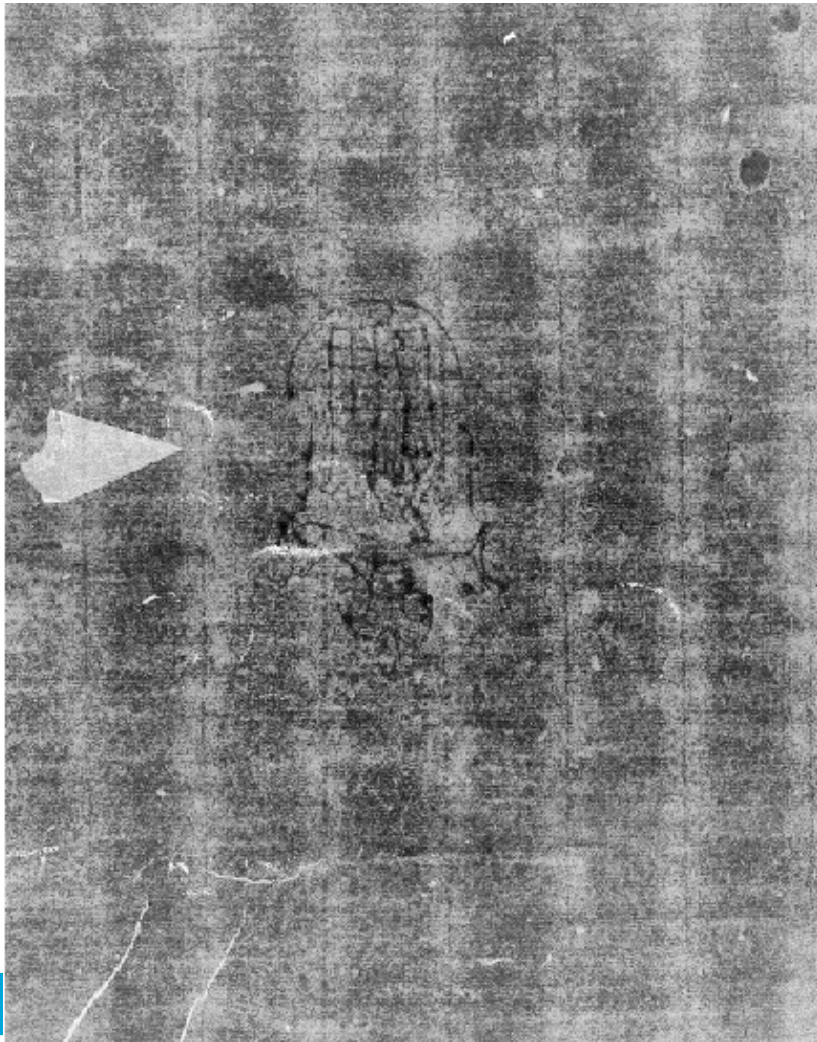
Automatic watermark detection in gray scale images

Exploit line properties: profile, contrast, length, spatial connectivity.

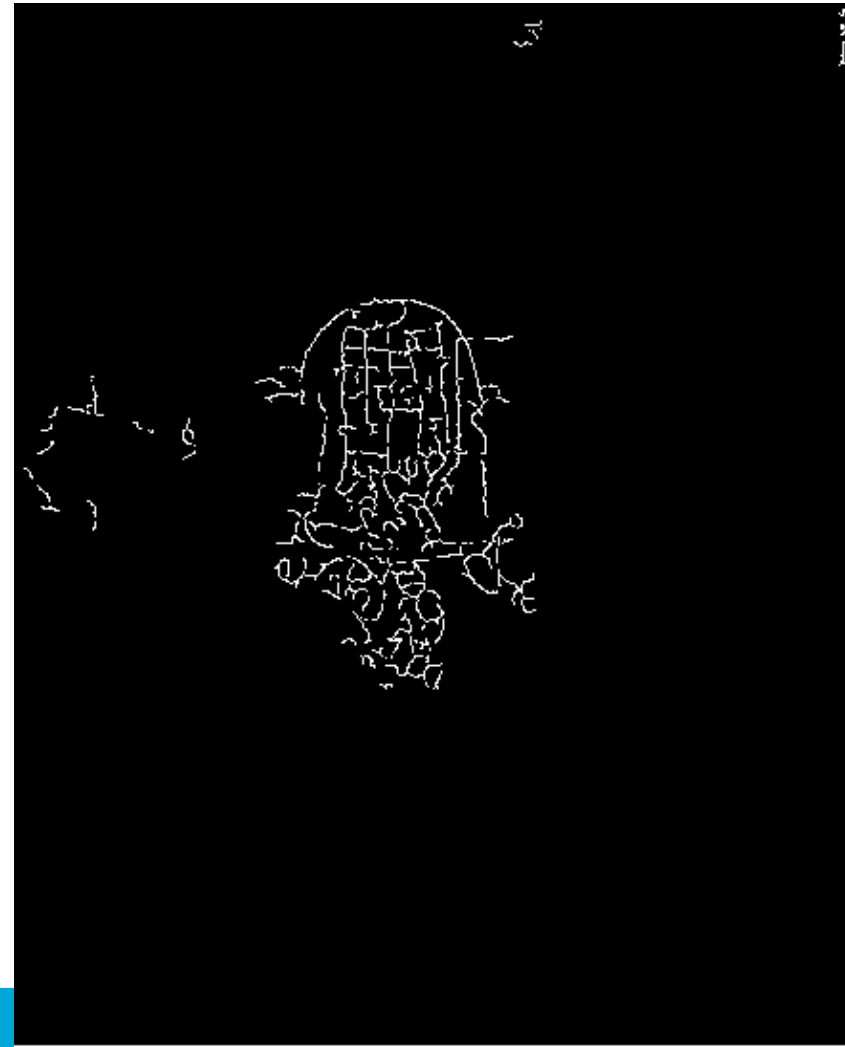
Optimal parameters are tuned automatically depending on the image to provide the best possible watermark detection

Matlab GUI

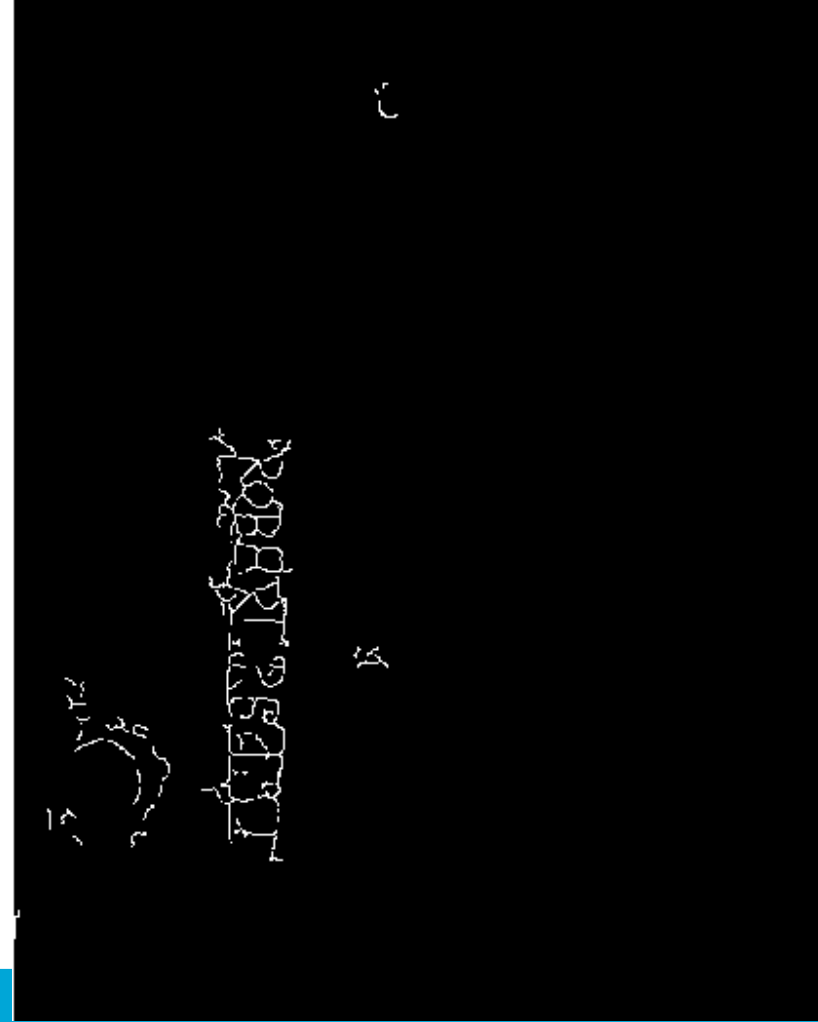
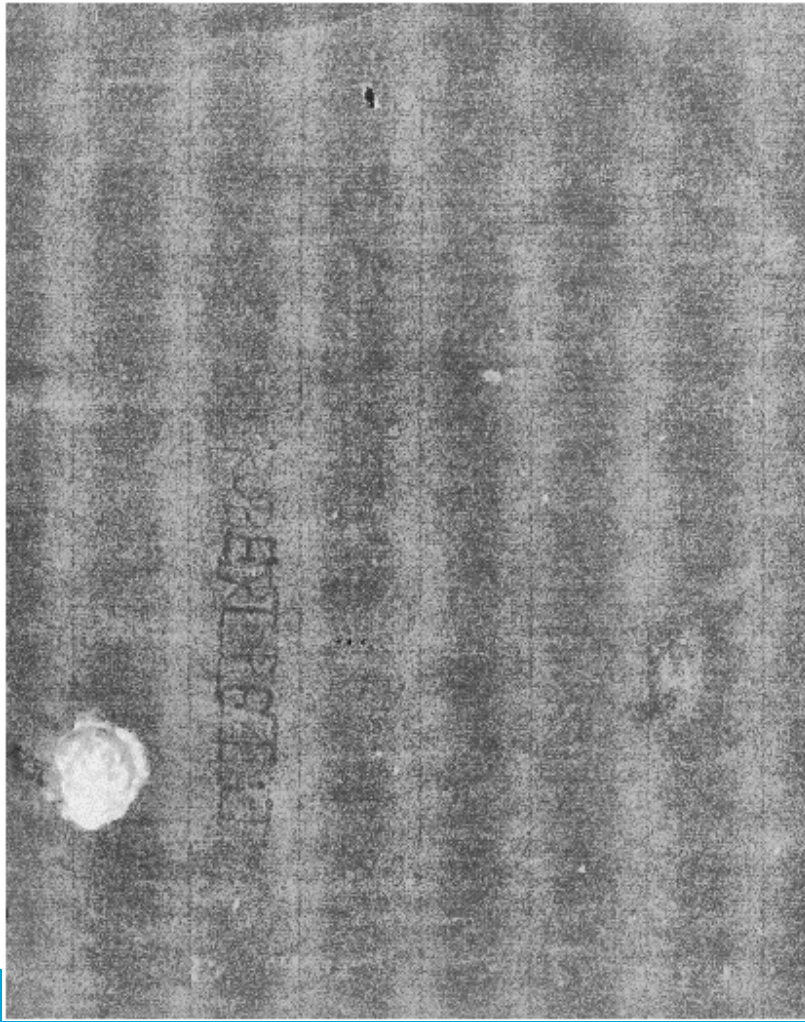
Automatic watermark detection in gray scale images



$[I, C]^T$

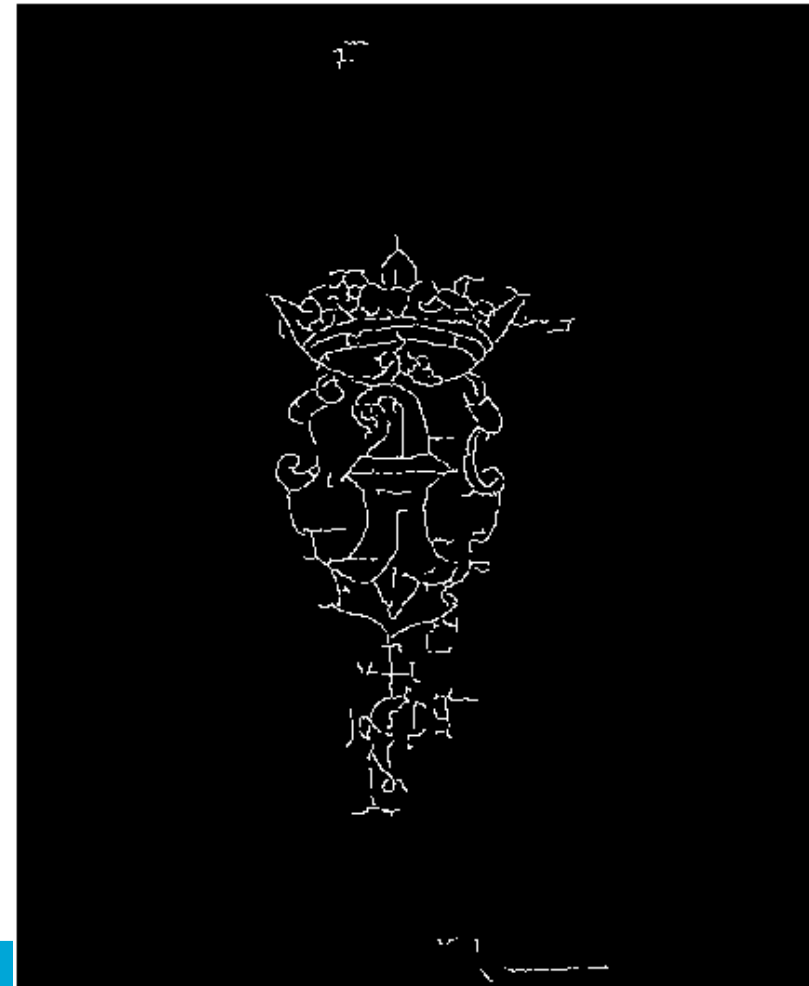
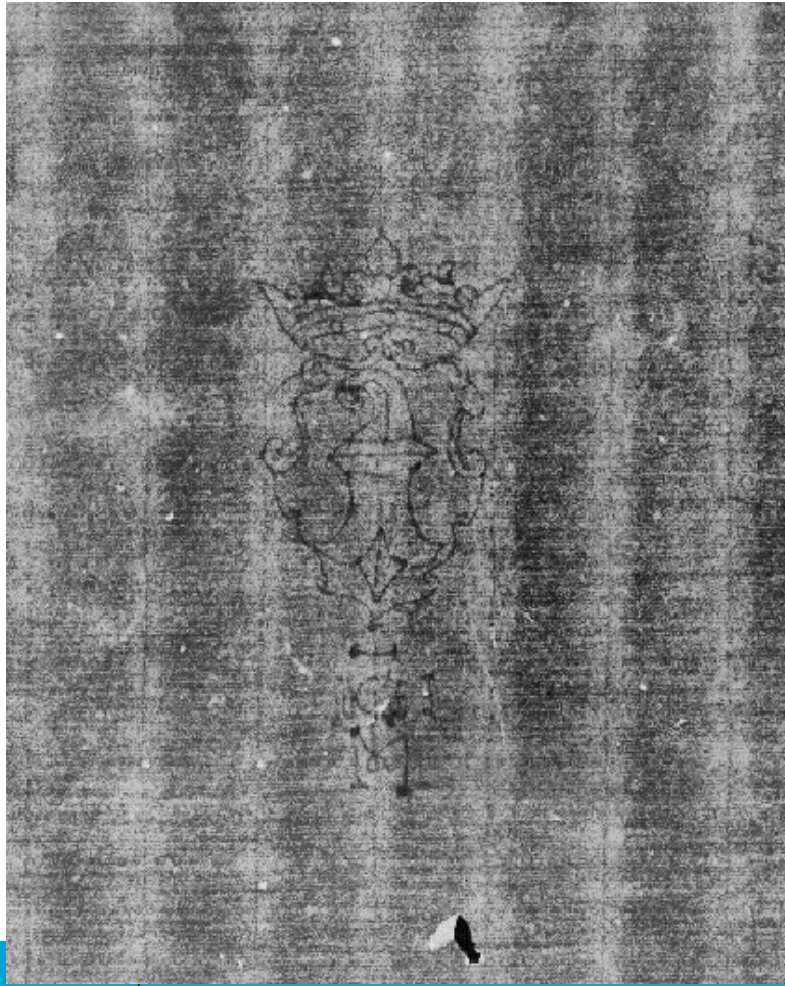


Automatic watermark detection in gray scale images



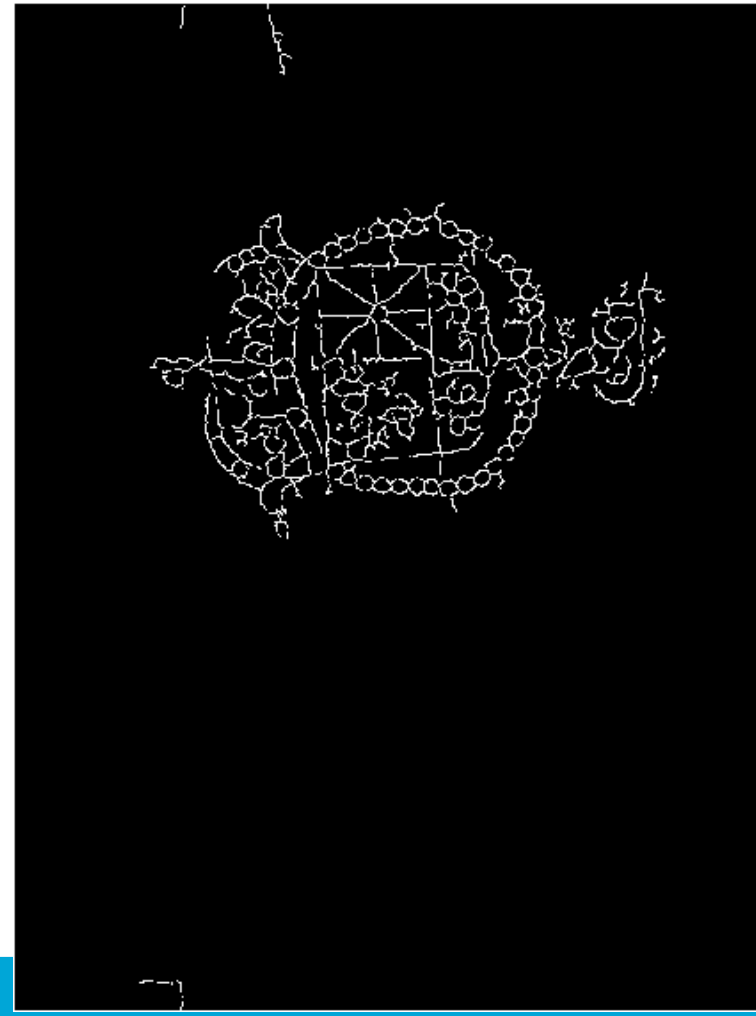
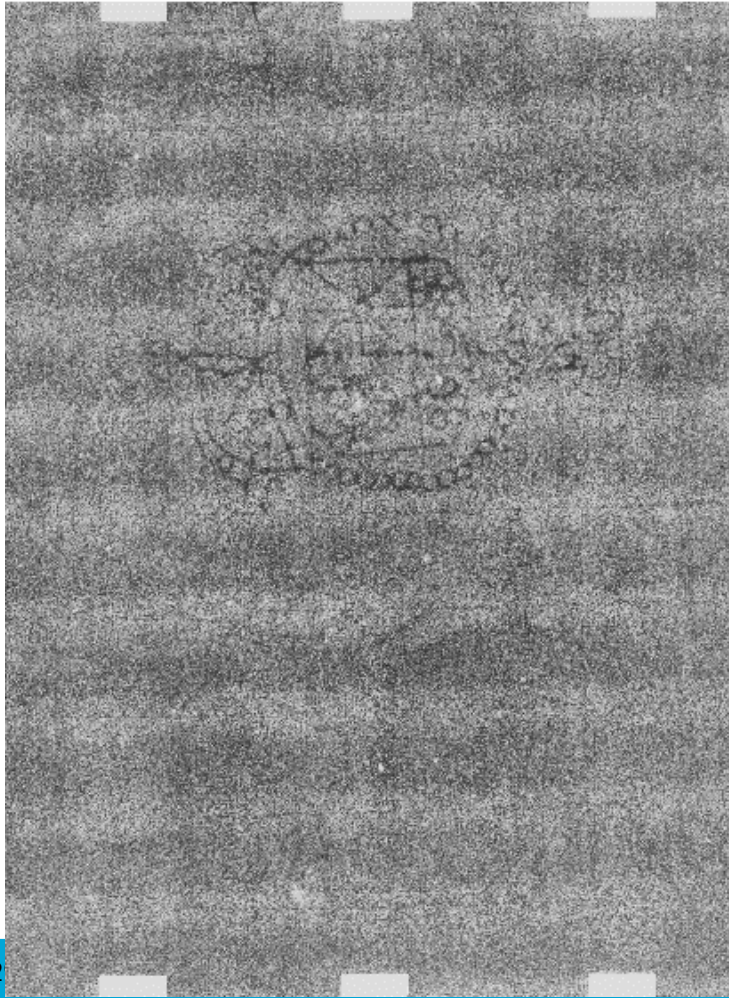
$[I, C]^T$

Automatic watermark detection in gray scale images



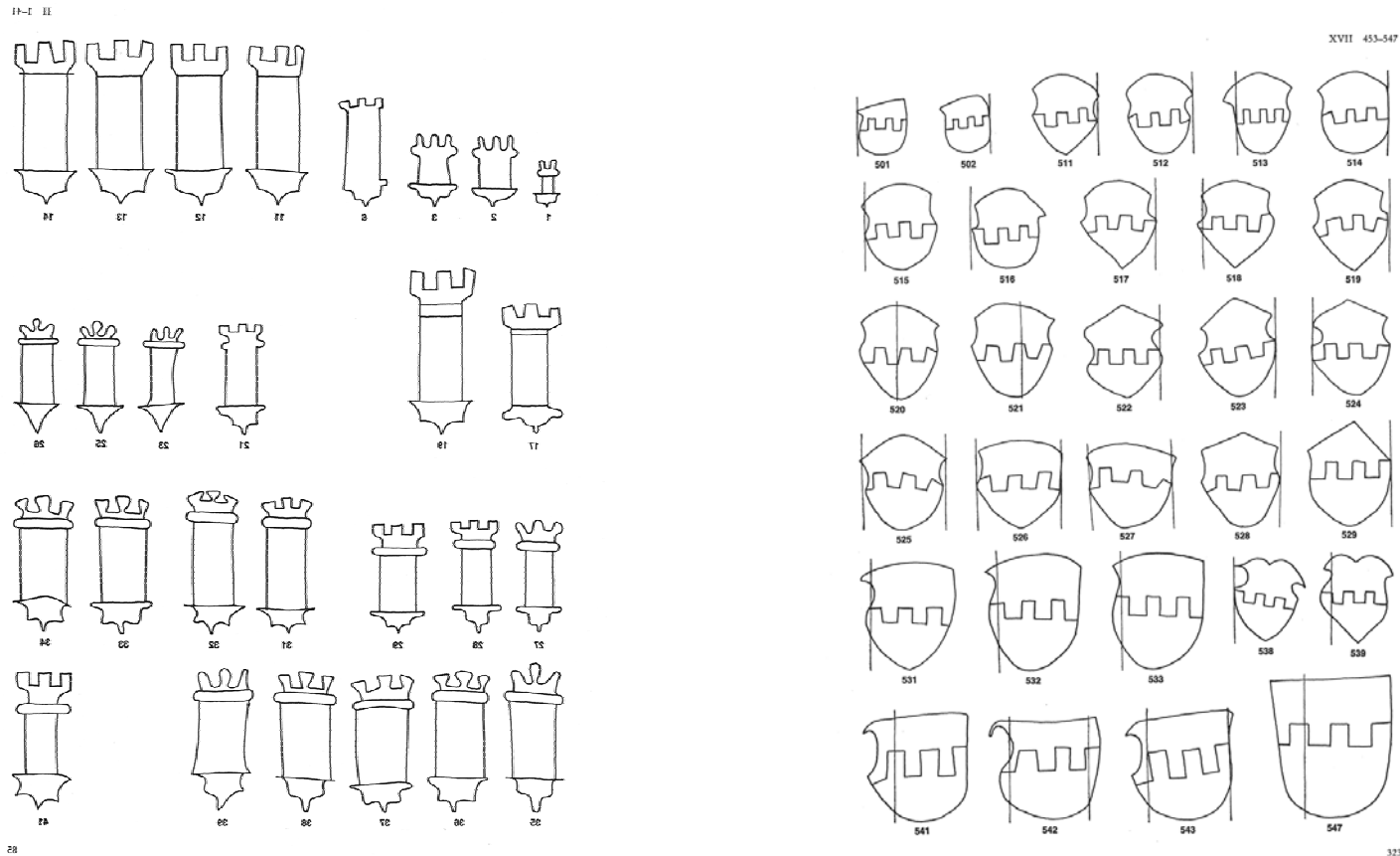
$[I, C]^T$

Automatic watermark detection in gray scale images

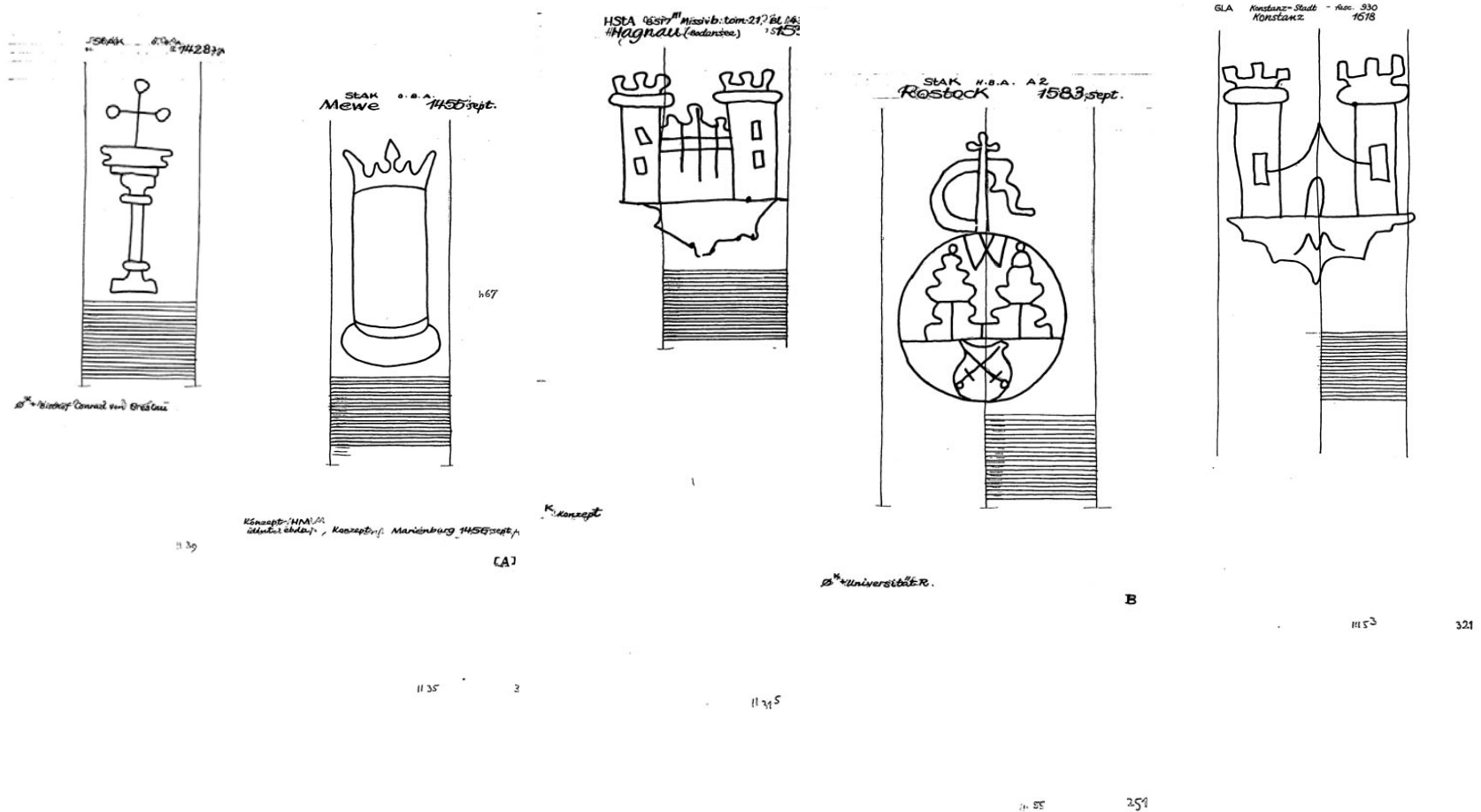


$[I, C]^T$

Printed Piccard Vs. Piccard Online (Retrieval)



Printed Piccard Vs. Piccard Online (Retrieval)

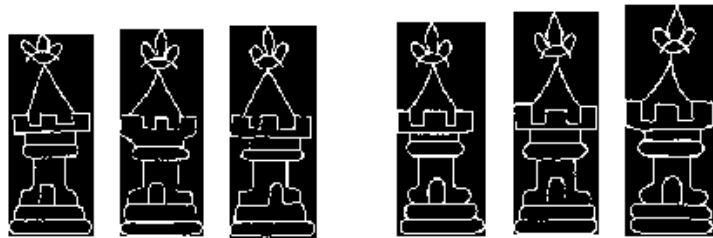
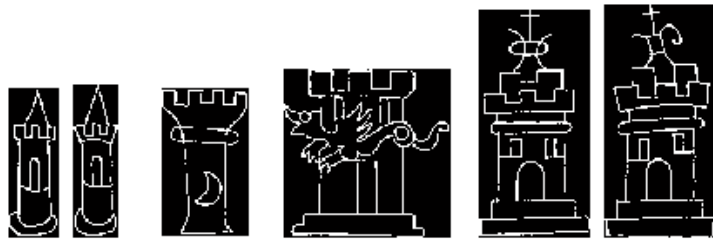


29 October, 2007

16

[I,C)^T

Printed Picard Vs. Picard Online (Retrieval)

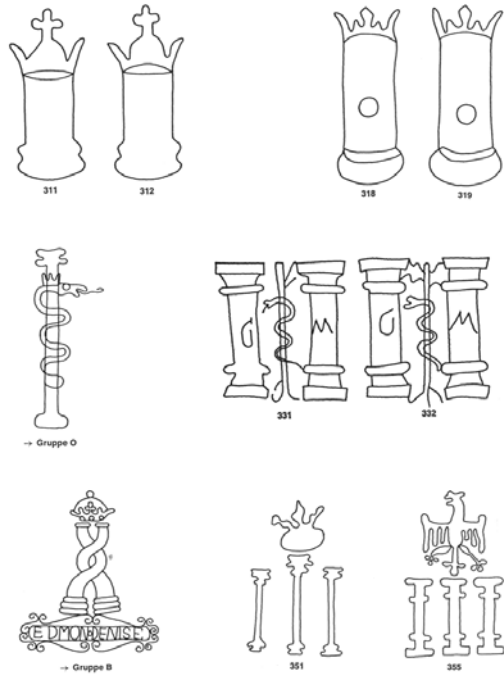


1. Count automatically the number of Watermarks taking into account the numbers.
Performance =93.85%
2. Extract automatically the Watermarks. Performance=92.2 %

Matlab GUI

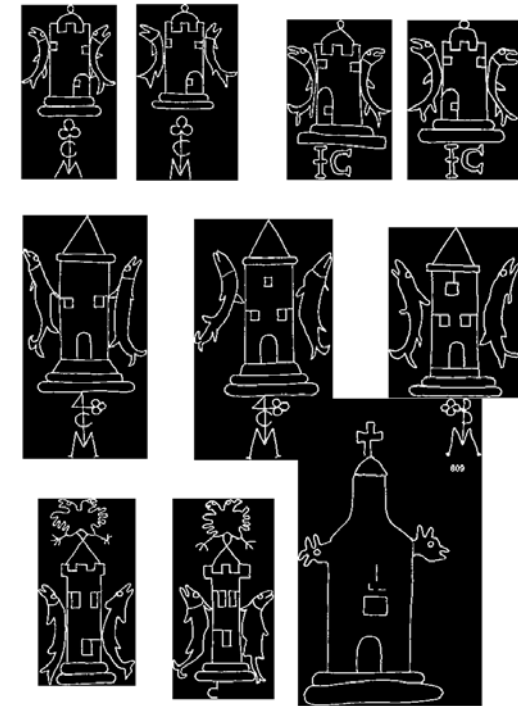
Test set: 244 pieces of paper coming from the printed Picard collection

Printed Picard Vs. Picard Online (Retrieval)

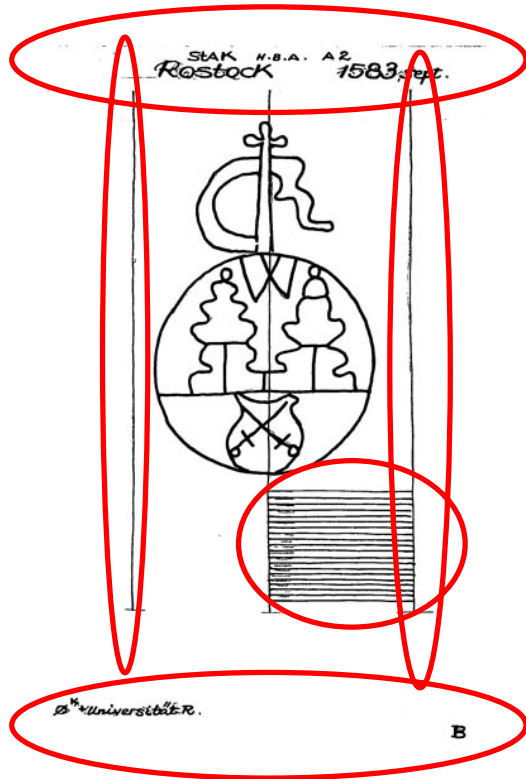


Errors:

1. Bad estimation of number of WM in the page
2. Small parts closer to other WM



Printed Piccard Vs. Piccard Online (Retrieval)



Detect automatically the bounding box of the watermark.

Remove noisy parts

Performance= 89%

Test set= 6000 images

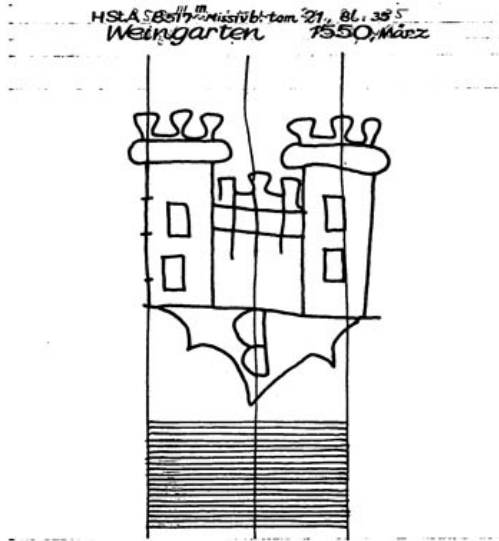
Matlab GUI

$[I, C]^T$

Printed Piccard Vs. Piccard Online (Retrieval)

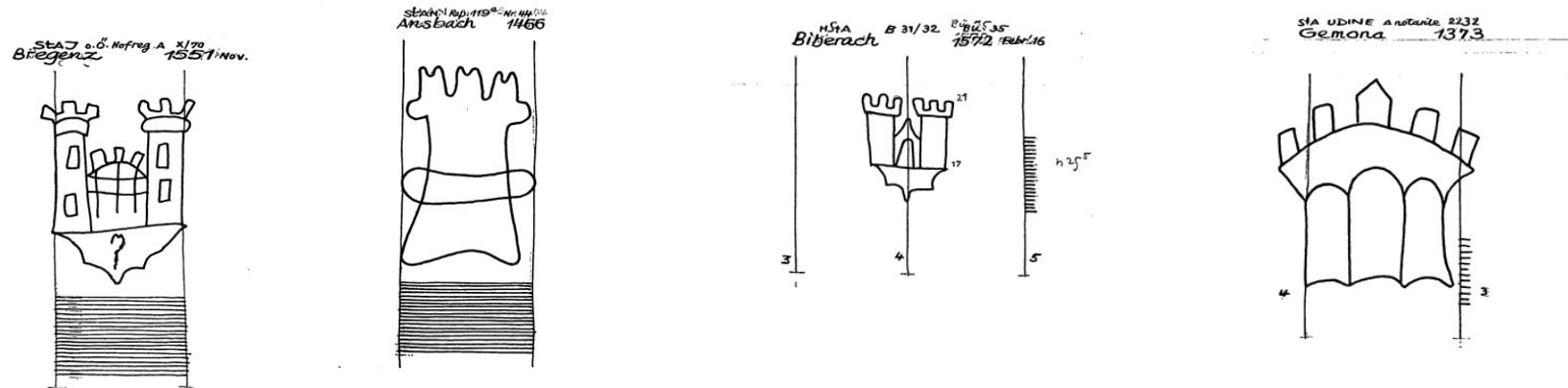
Error causes:

- Chain lines. Contact with watermark (65%)
- Watermark is removed partially during preprocessing (20%)
- 2-parts watermark very separated (10%)
- Laid lines (5%)



K

Printed Piccard Vs. Piccard Online (Retrieval)

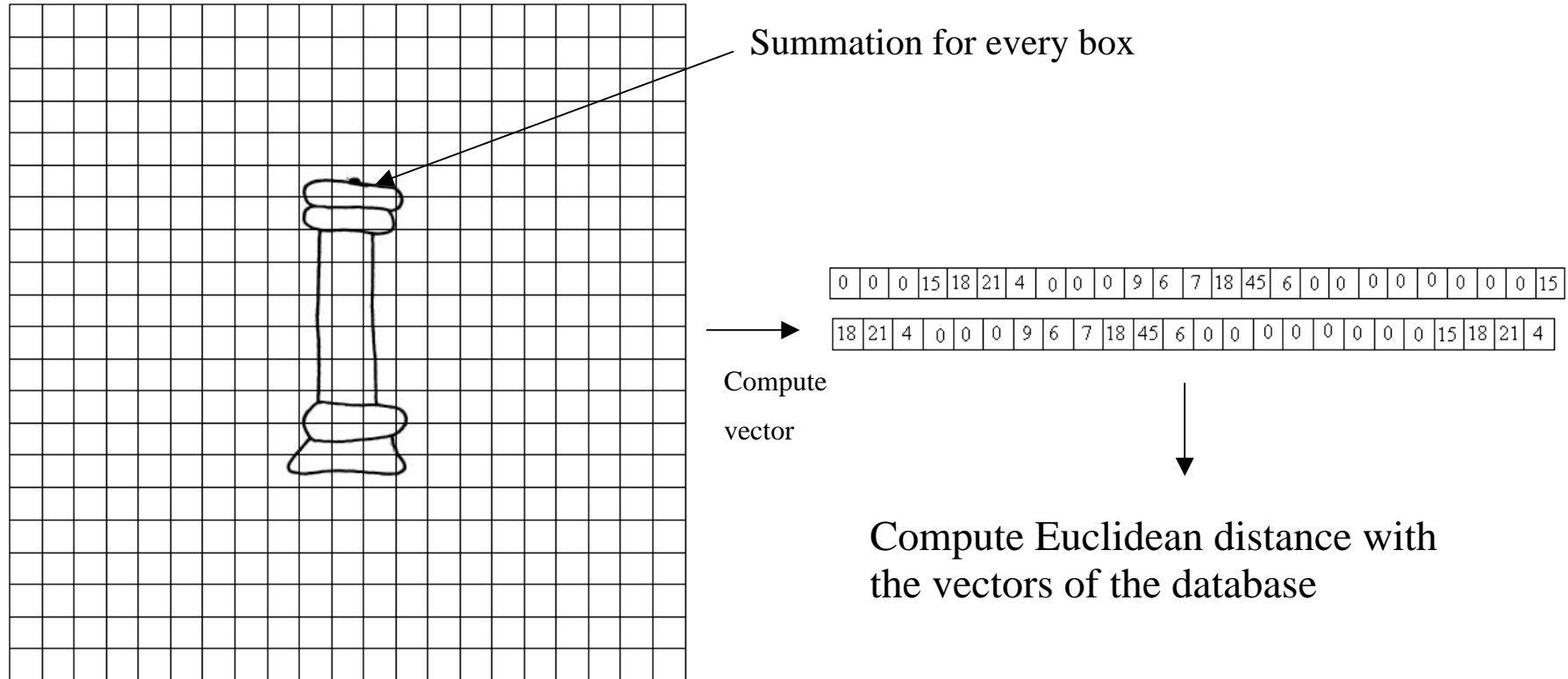


Chain lines in contact with Watermark

o. F. Michel Bauknecht

Semi-automatic method?

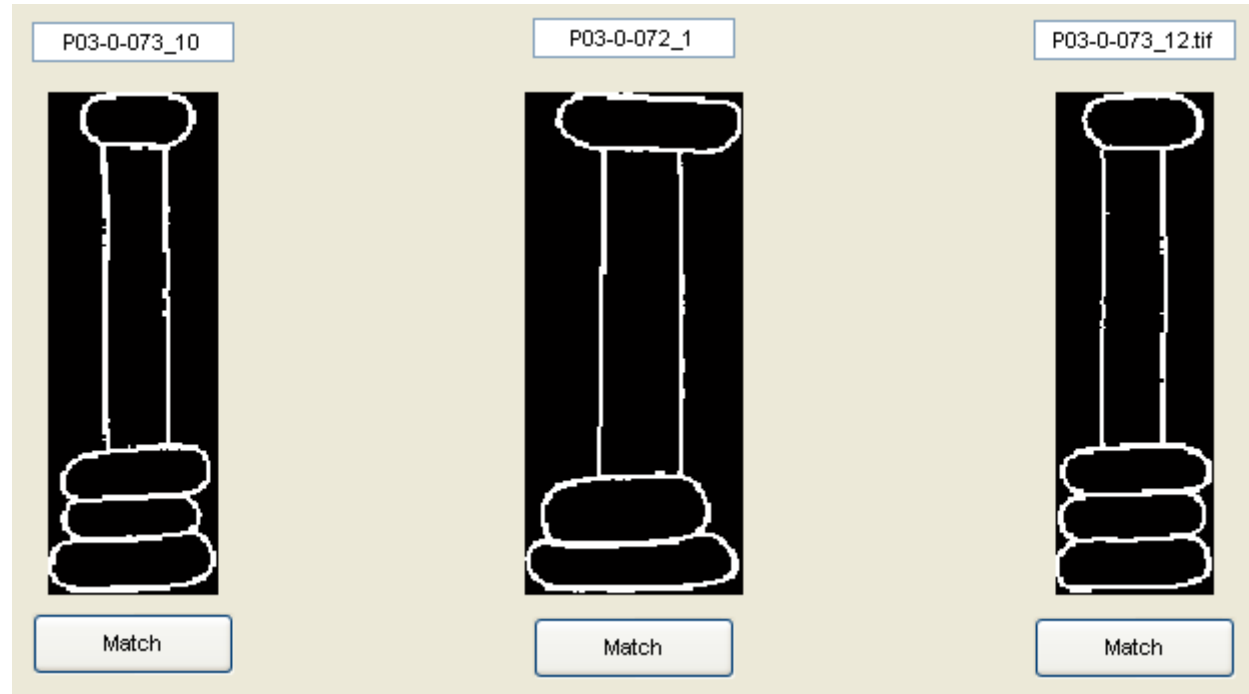
Printed Piccard Vs. Piccard Online (Retrieval)



Printed Piccard Vs. Piccard Online (Retrieval)



Query (Piccard online)

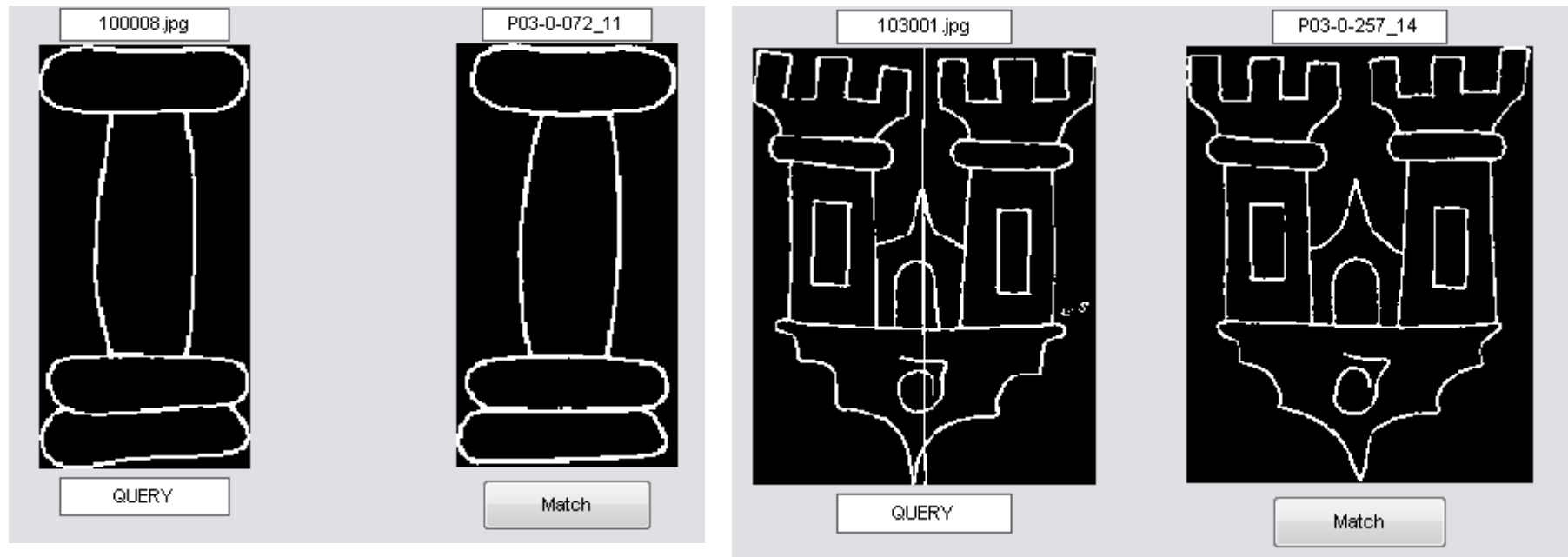


Database (Printed piccard)

Printed Piccard Vs. Piccard Online (Retrieval)

Problems:

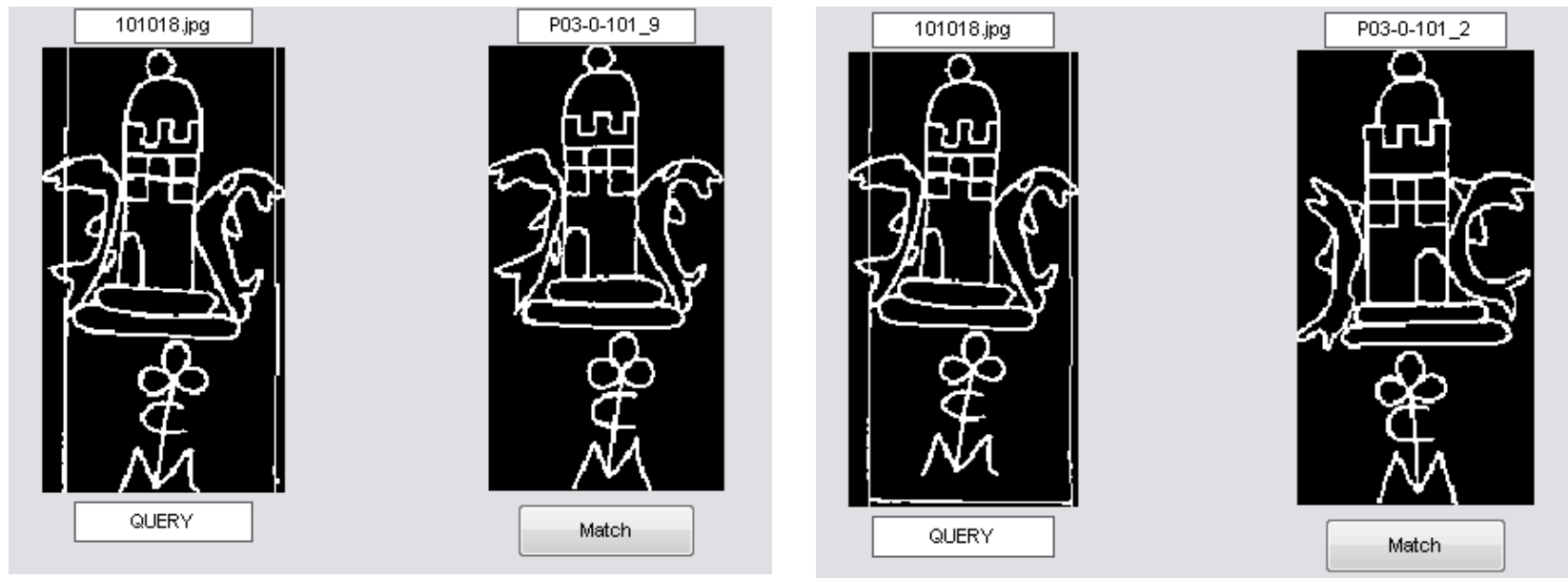
1. Matches are subjective. They cannot be detected automatically



Printed Piccard Vs. Piccard Online (Retrieval)

Problems:

2. Resolution, translation and rotation influence. Precision of boundary box. Future work: Invariance content-based features



Future work

- 1- Improve Automatic watermark detection in gray-scale images. Consider the line width. Noise filters. Use together with the semi-automatic detection method developed by Wenger. User interaction?
- 2- Watermark Scissors for Briquet collection? It is necessary to adapt the method for Piccard to Briquet since the distribution of the watermarks in the paper, noise and numbers are different.
- 3- Improve Automatic watermark detection in binary images (Piccard Online). Improve the performance of the already developed method. Remove the chain lines keeping the watermark pixels. Semi-automatic method ?

Future work

4- Matching. Develop other techniques which work properly when the boundary box of the watermarks is not perfectly detected. Confirm matching automatically. Make it useful for detection coming from gray-scale images.

5- Interfering papers.