ÖAW AUSTRIAN ACADEMY OF SCIENCES

ARI

From Thermographic Data to Watermark Images

Günther Koliander, Paul Gulewycz, Marlene Peterlechner,
Clemens Gubsch, Andrea Lindmayr-Brandl
and Katharina Loose-Einfalt

Acoustics Research Institute, Austrian Academy of Sciences

ÖAW AUSTRIAN ACADEMY OF SCIENCES

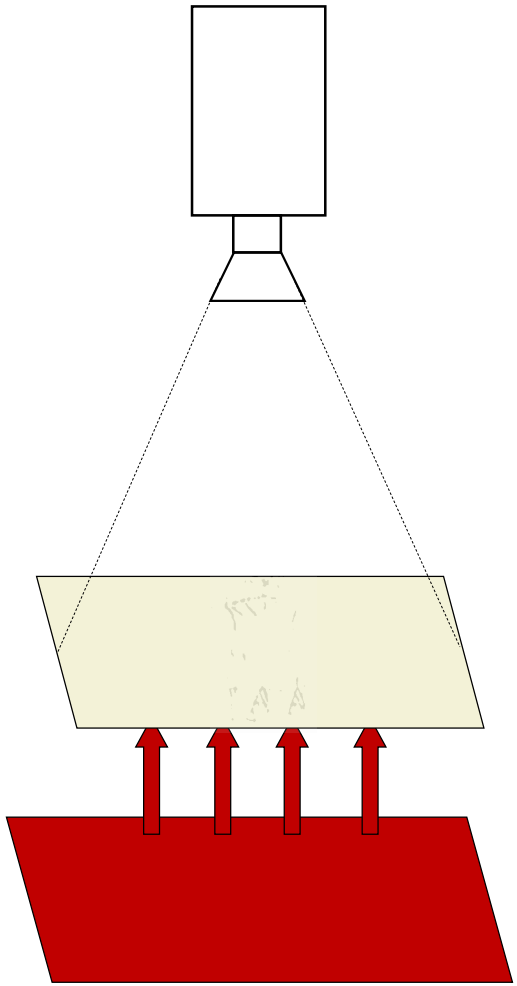
DRACMarkS Project

ARI

- Digitization, Recognition, and Automated Clustering of Watermarks in the Music Manuscripts of Franz Schubert
- Use a thermographic imaging device to digitize the music manuscripts
- Focus on capturing the whole page, not only the watermark
- Digitization of “raw data” as a goal
- As objective as possible

Thermographic Watermark Imaging

- IR-source (heating plate) emits IR-radiation
- IR-radiation traverses manuscript
- Intensity of IR-radiation is detected by IR-camera
- No direct conduction of heat



Benefits and Challenges

- Most writing materials are invisible in the IR spectrum
- No hazardous radiation
- Poor resolution (640 x 512 pixels)
- Challenging acquisition procedure
- Expensive components
- Difficult to capture watermarks close to binding
- No standardized procedure

ÖAW AUSTRIAN ACADEMY OF SCIENCES

Acquisition Procedure

ARI

- Gap between manuscript and IR source necessary
- Short exposure to IR radiation gives best contrast
- Difficult to locate watermarks in advance
- Only two to three seconds to place manuscript and record image

Data Pipeline

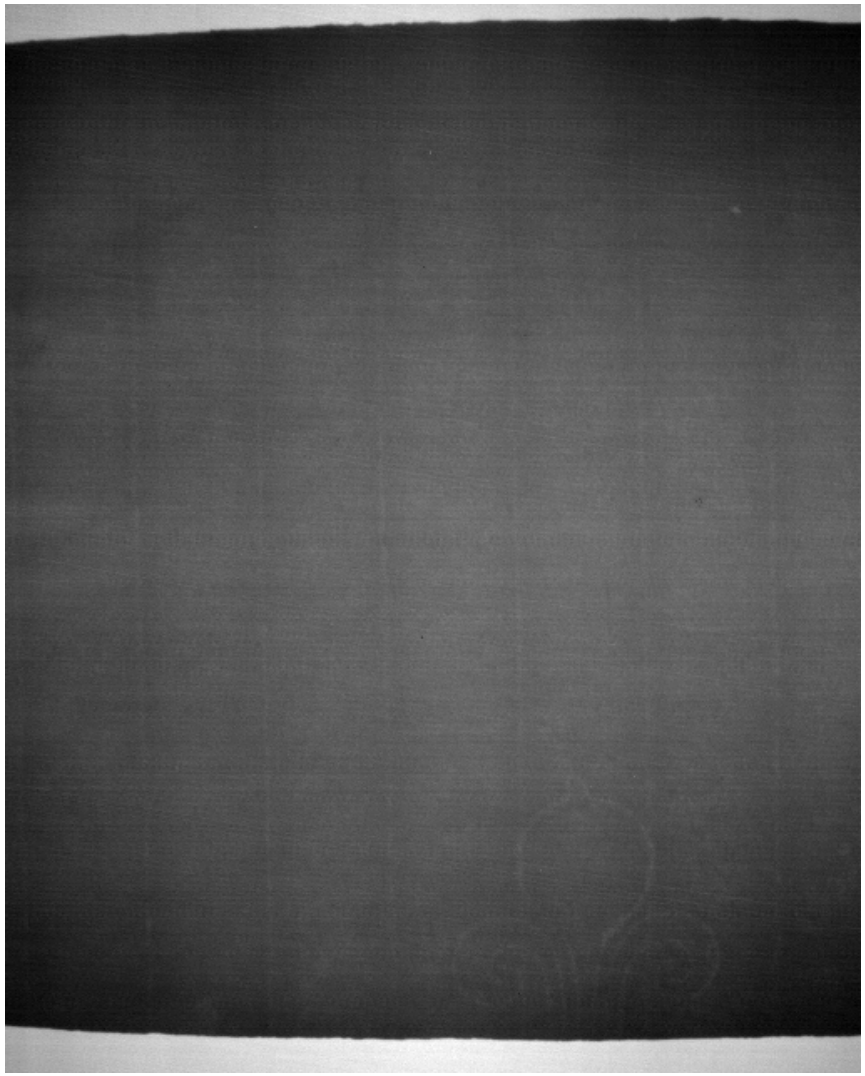
- Matrix of intensity values 0–65553
- Visualization as gray scale image
- White is the highest value, black the smallest

Data Pipeline

- Matrix of intensity values 0–65553
- Visualization as gray scale image
- White is the highest value, black the smallest
- Remove bad pixels (hardware problem)

Data Pipeline

- Matrix of intensity values 0–65553
- Visualization as gray scale image
- White is the highest value, black the smallest
- Remove bad pixels (hardware problem)

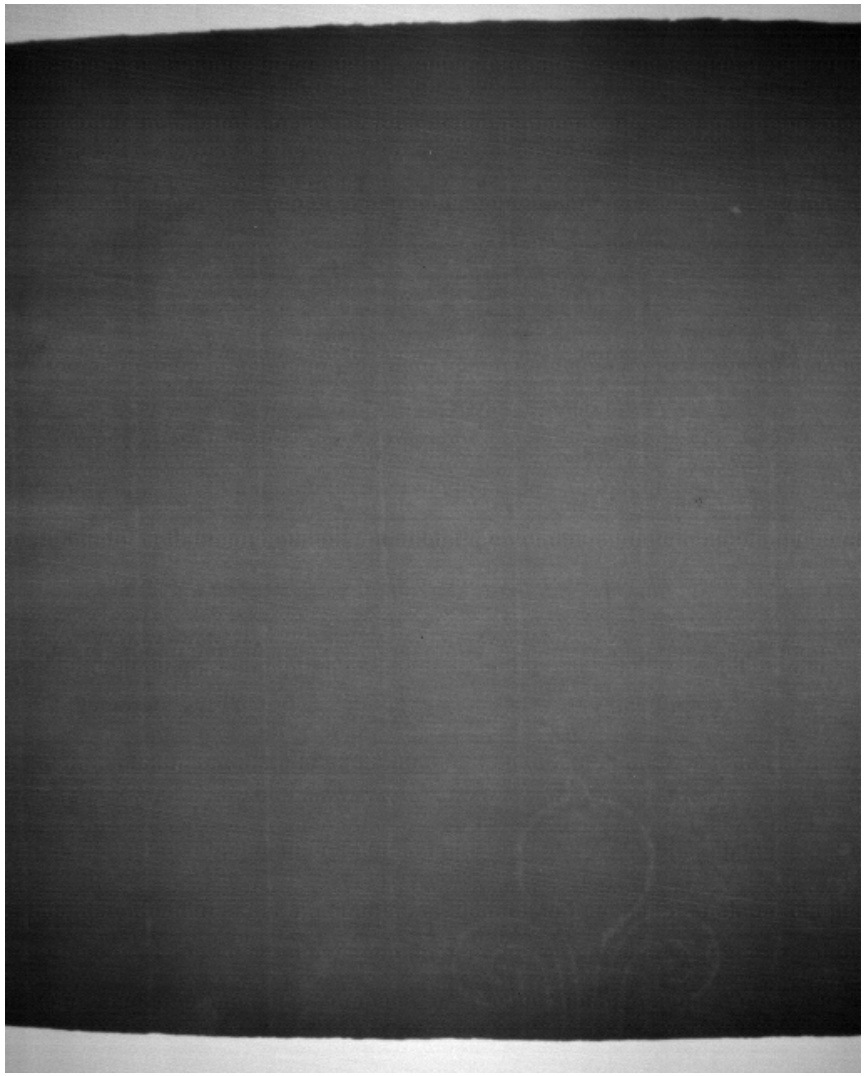


ÖAW AUSTRIAN ACADEMY OF SCIENCES

Data Pipeline

ARI

- Matrix of intensity values 0–65553
- Visualization as gray scale image
- White is the highest value, black the smallest
- Remove bad pixels (hardware problem)
- Radial brightness effect (lens)



Data Pipeline

- Subtract reference image

Data Pipeline

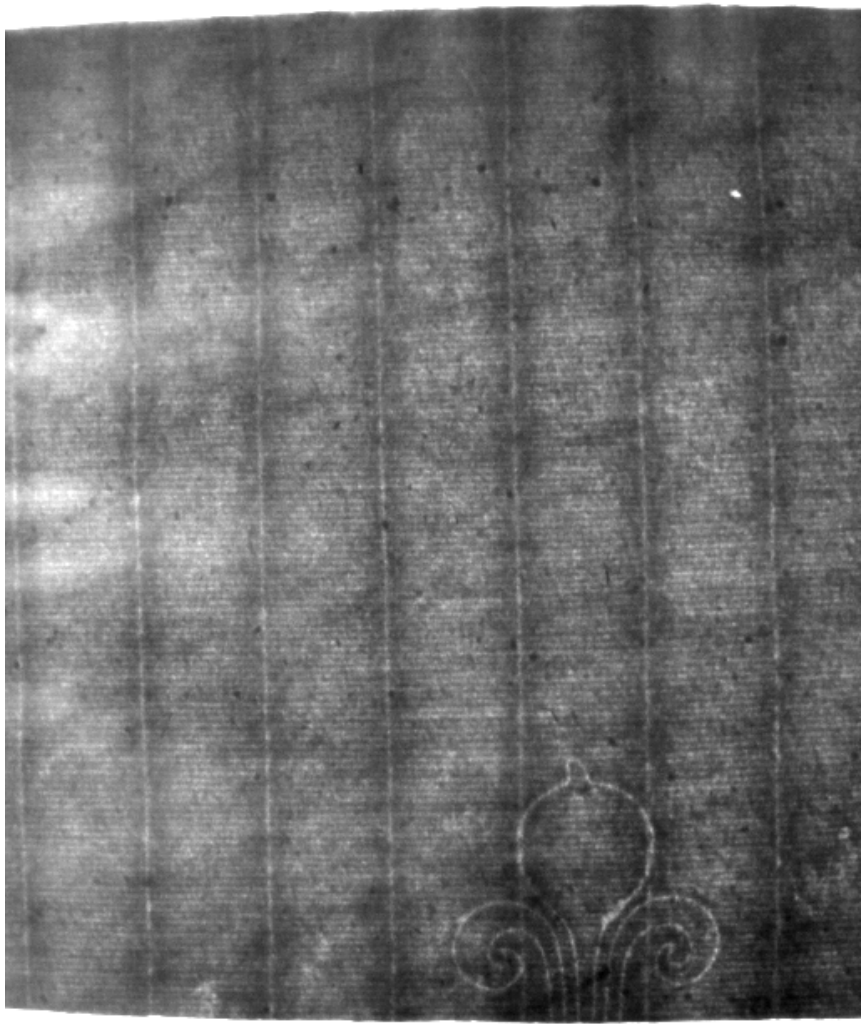
- Subtract reference image
- Huge gap between background and foreground

ÖAW AUSTRIAN ACADEMY OF SCIENCES

Data Pipeline

ARI

- Subtract reference image
- Huge gap between background and foreground
- Restrict to foreground values

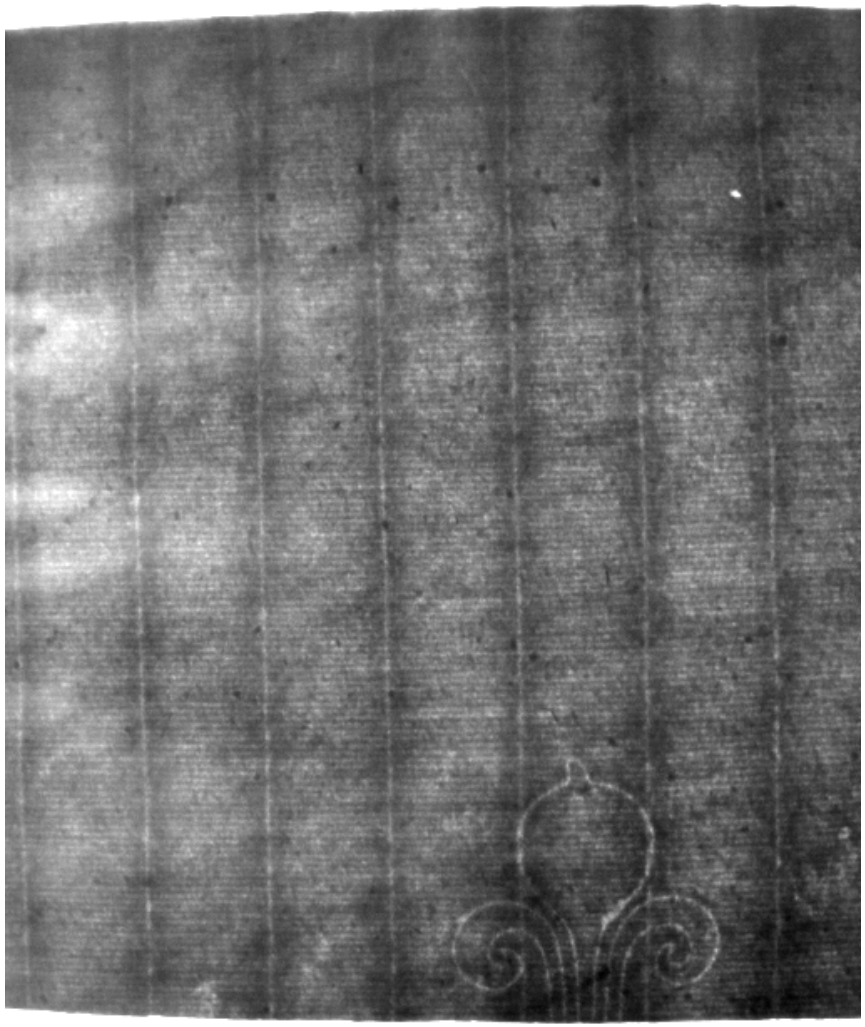


ÖAW AUSTRIAN ACADEMY OF SCIENCES

Data Pipeline

ARI

- Subtract reference image
- Huge gap between background and foreground
- Restrict to foreground values
- Barrel effect (lens distortion)

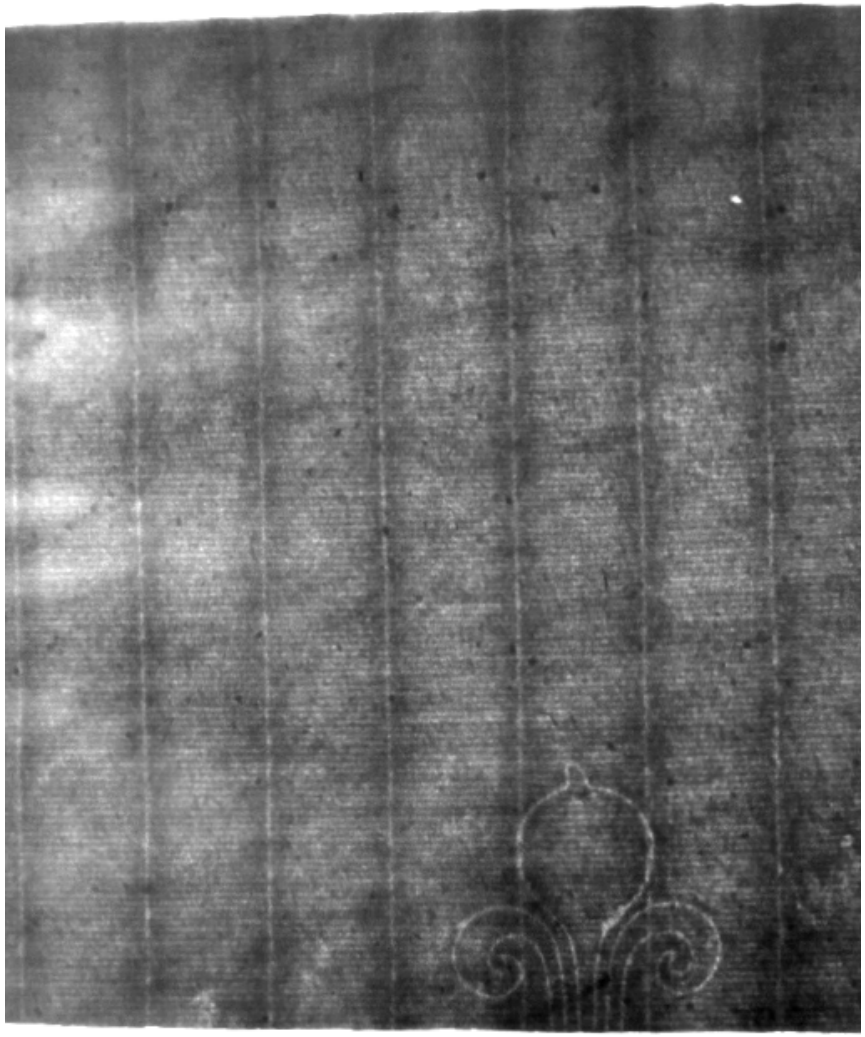


ÖAW AUSTRIAN ACADEMY OF SCIENCES

Data Pipeline

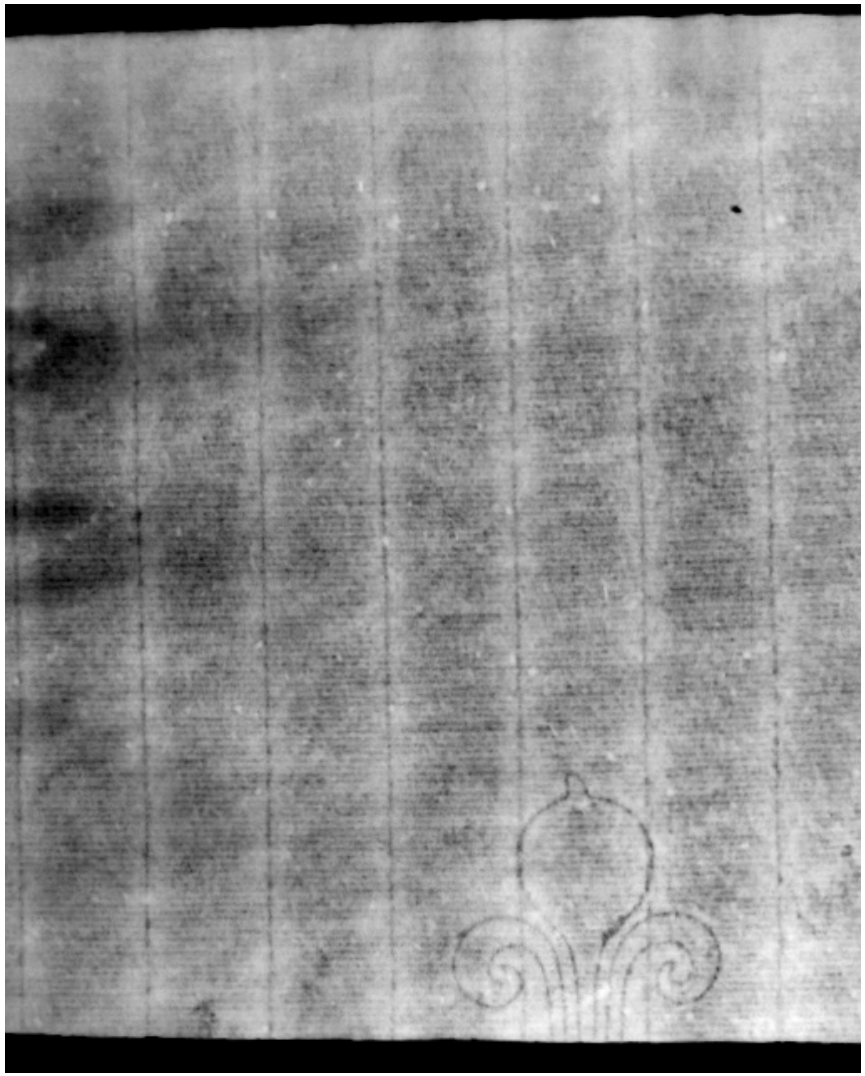
ARI

- Subtract reference image
- Huge gap between background and foreground
- Restrict to foreground values
- Barrel effect (lens distortion)
- Correct radial distortion



Data Pipeline

- Subtract reference image
- Huge gap between background and foreground
- Restrict to foreground values
- Barrel effect (lens distortion)
- Correct radial distortion
- Invert black-white

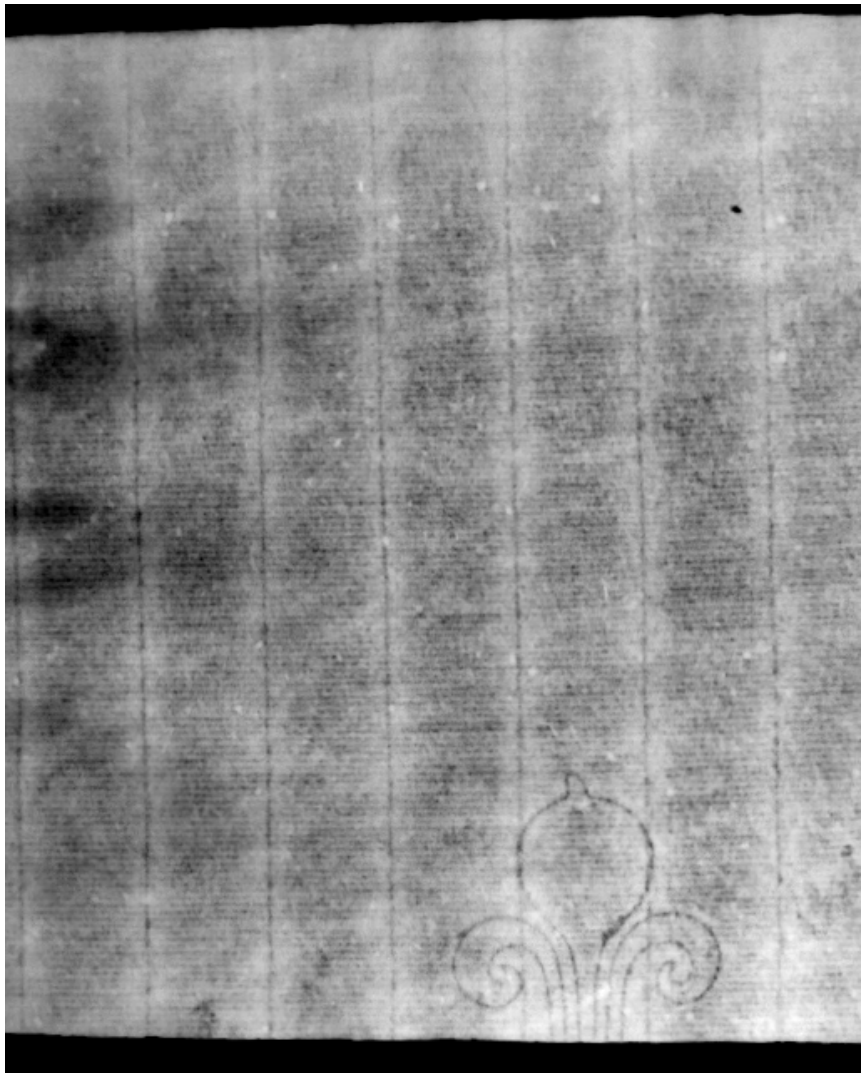


ÖAW AUSTRIAN ACADEMY OF SCIENCES

Data Pipeline

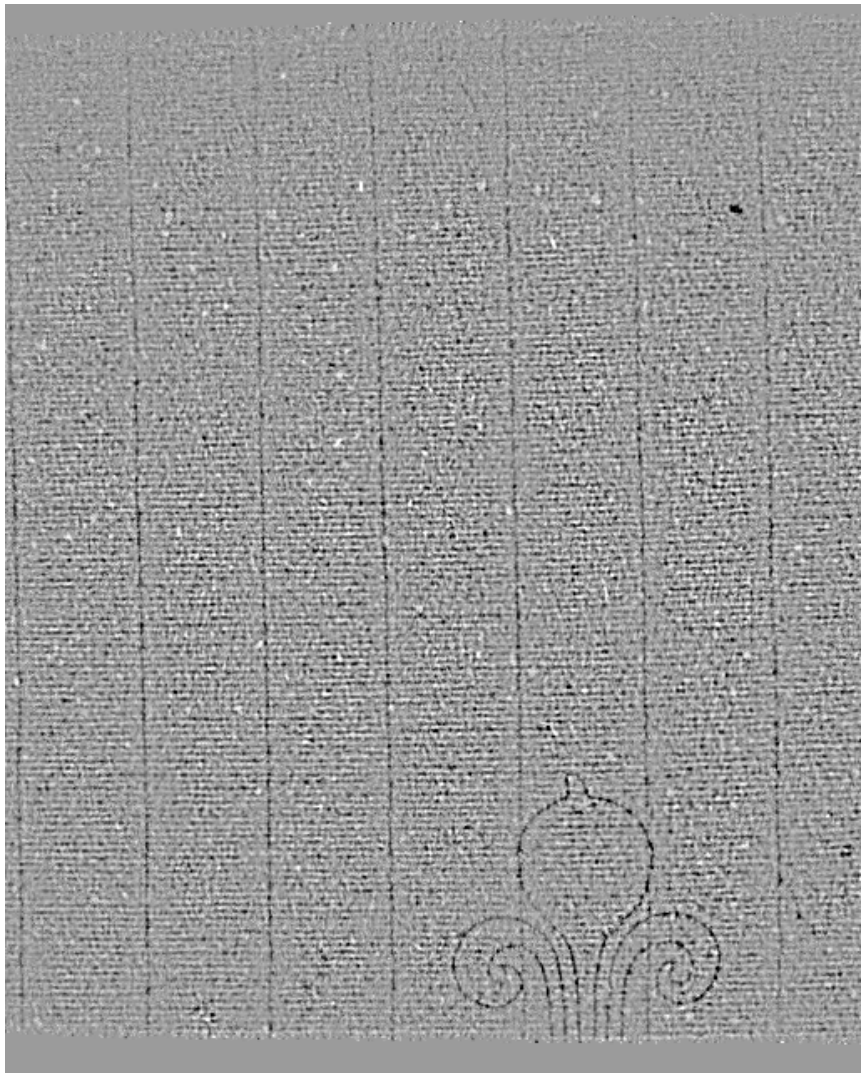
ARI

- Subtract reference image
- Huge gap between background and foreground
- Restrict to foreground values
- Barrel effect (lens distortion)
- Correct radial distortion
- Invert black-white
- Local brightness variation
- Correction warranted?



Data Pipeline

- Subtract reference image
- Huge gap between background and foreground
- Restrict to foreground values
- Barrel effect (lens distortion)
- Correct radial distortion
- Invert black-white
- Local brightness variation
- Correction warranted?



Half-page to Page

- So far illustration of a single recording
- Claimed that we are interested in whole page

ÖAW AUSTRIAN ACADEMY OF SCIENCES

Half-page to Page

ARI

- So far illustration of a single recording
- Claimed that we are interested in whole page
- Resolution not good enough for whole page at once
- We have to combine 2–4 recordings to get a single page

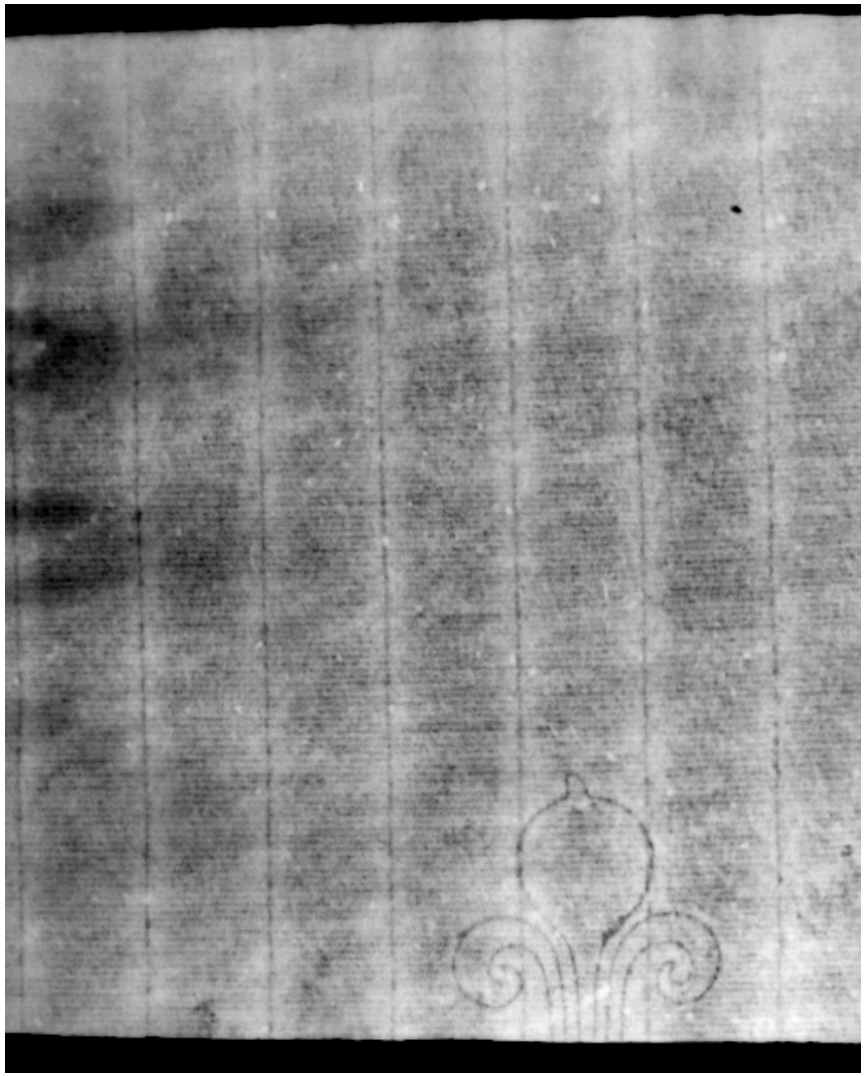
Half-page to Page

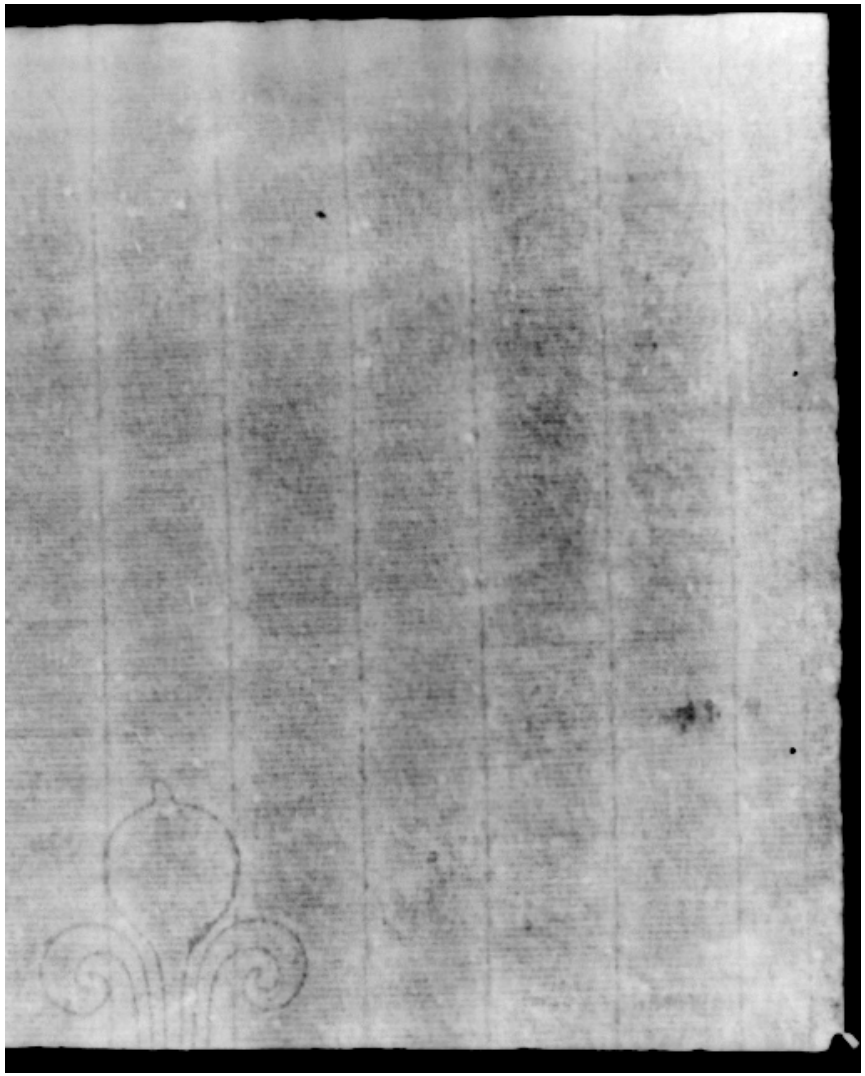
- So far illustration of a single recording
- Claimed that we are interested in whole page
- Resolution not good enough for whole page at once
- We have to combine 2–4 recordings to get a single page
- Standard image stitching methods are designed for photographs with sufficiently many distinctive features
- Paper structure does not provide these features

Half-page to Page

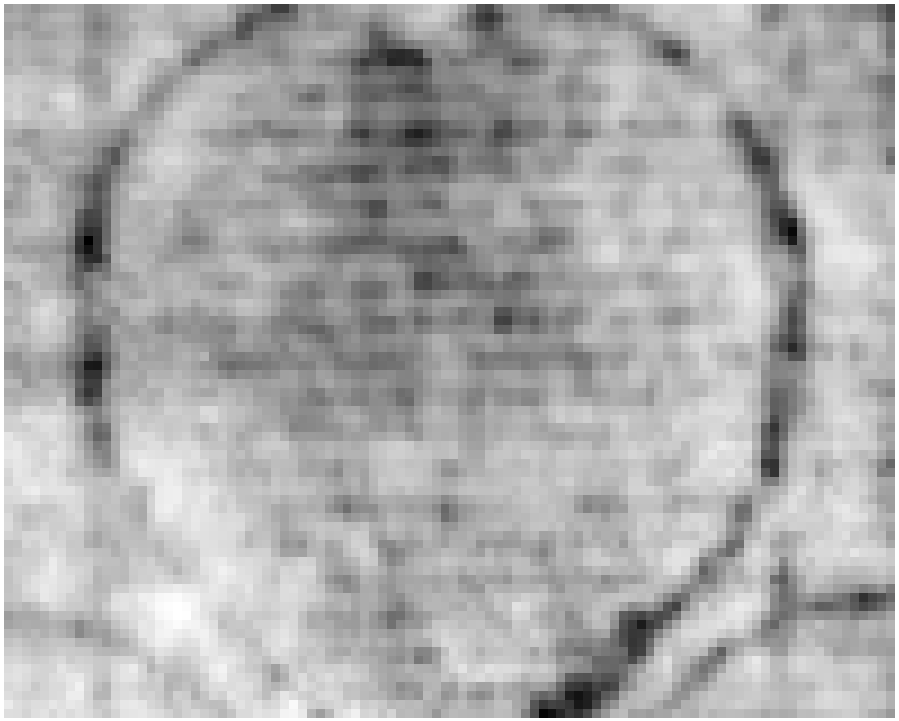
- So far illustration of a single recording
- Claimed that we are interested in whole page
- Resolution not good enough for whole page at once
- We have to combine 2–4 recordings to get a single page
- Standard image stitching methods are designed for photographs with sufficiently many distinctive features
- Paper structure does not provide these features
- Had to develop our own stitching algorithm

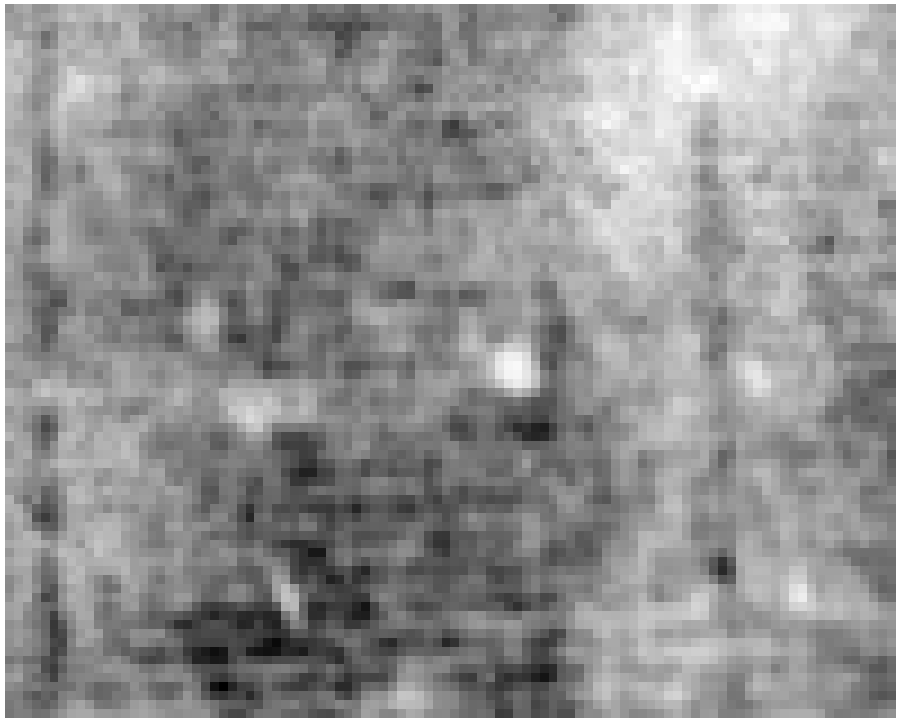
Stitching





Stitching

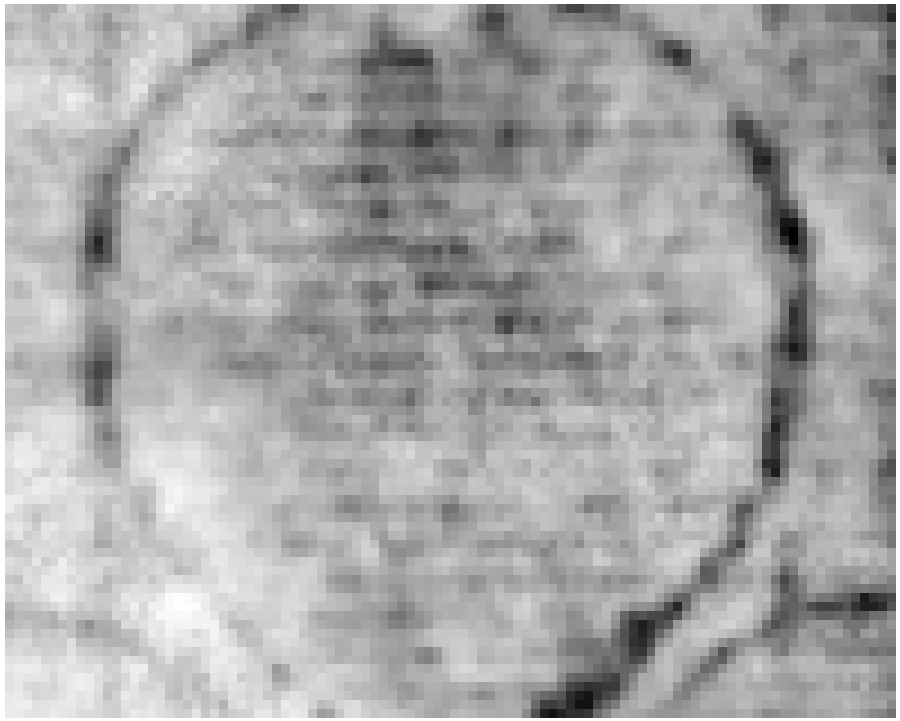


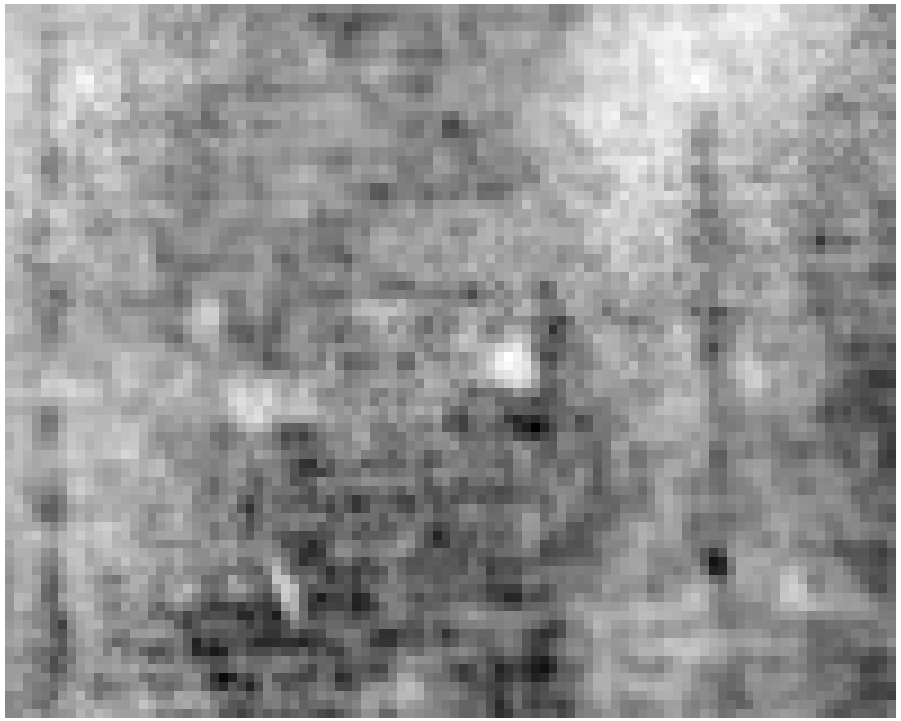


ÖAW AUSTRIAN ACADEMY OF SCIENCES

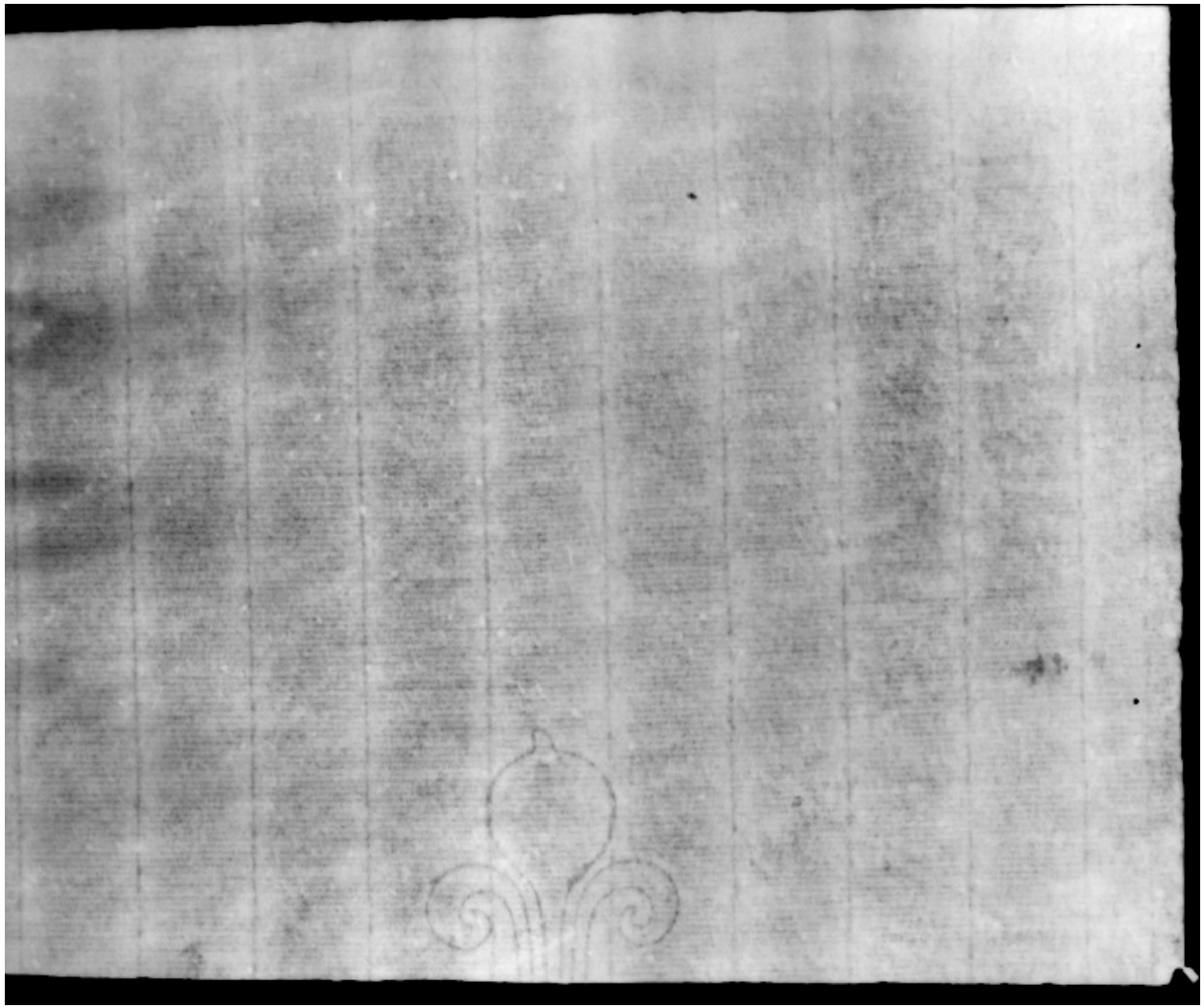
Stitching

ARI





Combined Page



Stitching

- Two reference points enable alignment of rotation and scaling

Stitching

- Two reference points enable alignment of rotation and scaling
- Eventually, more reference points can be used for differently scaling both axes

Stitching

- Two reference points enable alignment of rotation and scaling
- Eventually, more reference points can be used for differently scaling both axes
- Different (non-linear) registration methods will be tried to correct wrinkles

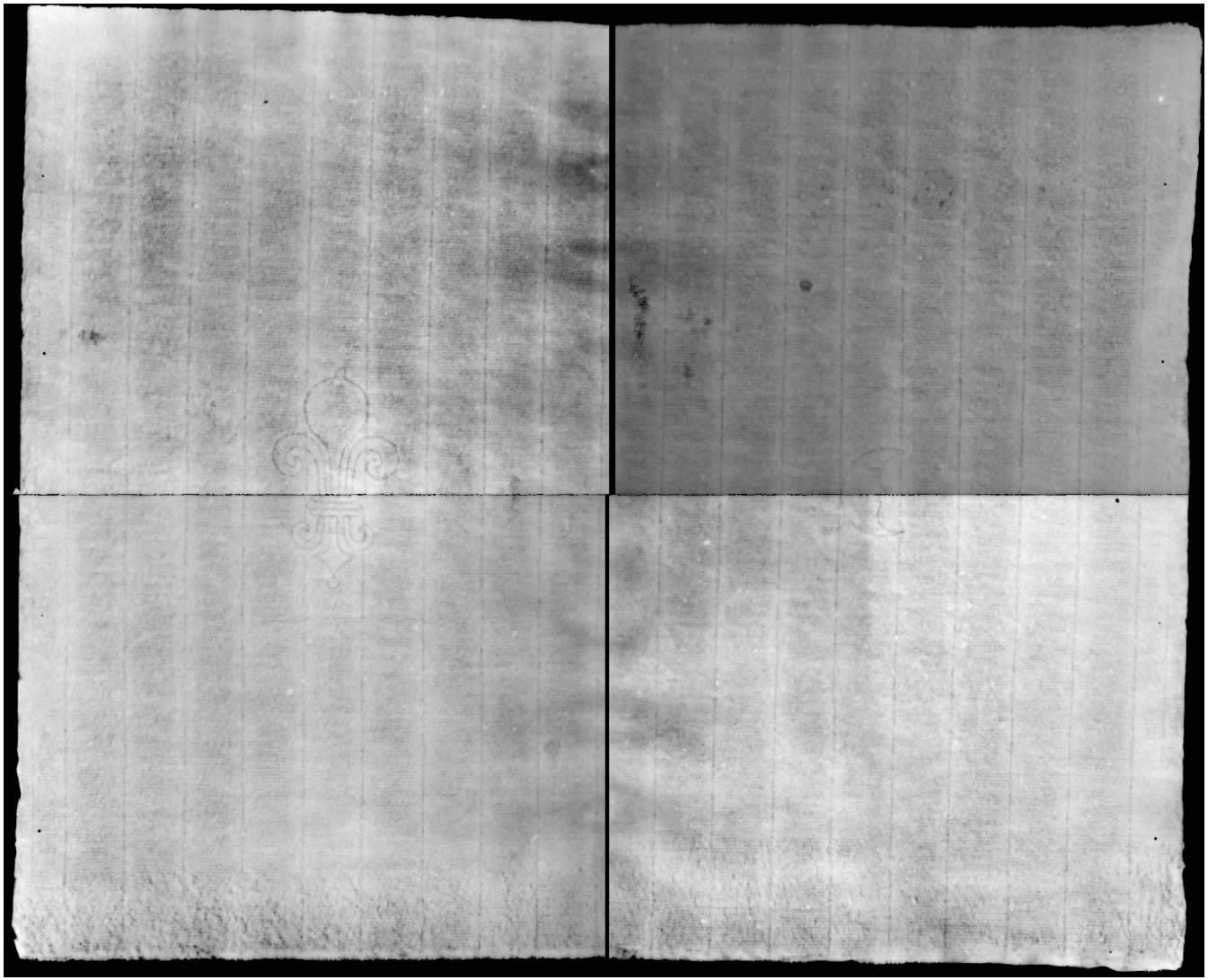
Stitching

- Two reference points enable alignment of rotation and scaling
- Eventually, more reference points can be used for differently scaling both axes
- Different (non-linear) registration methods will be tried to correct wrinkles
- Process is slow, optimization seems very difficult due to many local minima/maxima

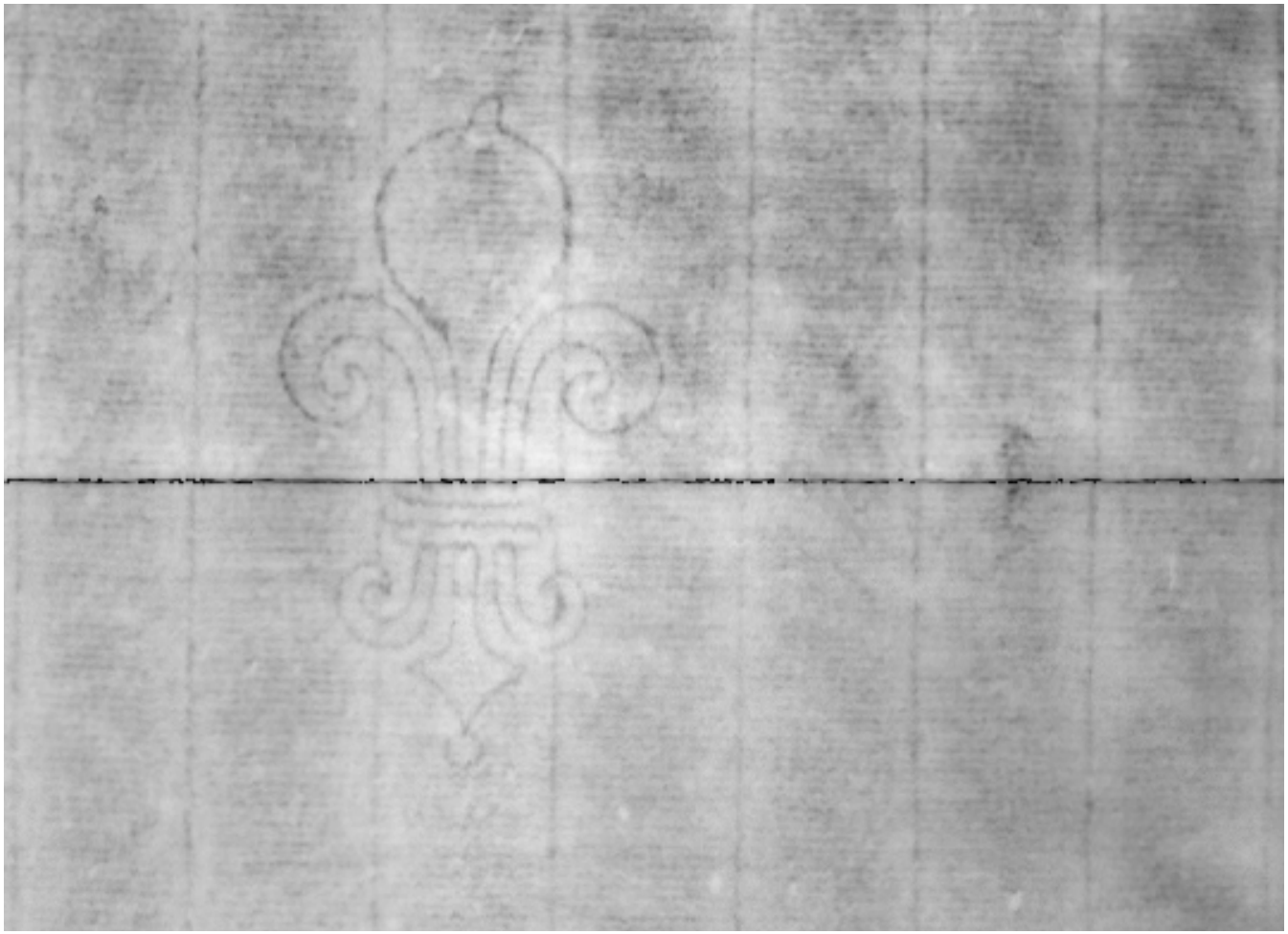
ÖAW AUSTRIAN ACADEMY OF SCIENCES

Why Full Page?

ARI



Why Full Page?



ÖAW AUSTRIAN ACADEMY OF SCIENCES

Outlook

ARI

- Standardize thermographic watermark digitization: DFG round table in October 2023

ÖAW AUSTRIAN ACADEMY OF SCIENCES

Outlook

ARI

- Standardize thermographic watermark digitization: DFG round table in October 2023
- Ongoing effort to binarize watermark from the “raw” images:
  - Use classical image processing techniques
  - Use machine learning methods

ÖAW AUSTRIAN ACADEMY OF SCIENCES

Outlook

ARI

- Standardize thermographic watermark digitization: DFG round table in October 2023
- Ongoing effort to binarize watermark from the “raw” images:
  - Use classical image processing techniques
  - Use machine learning methods
- Automatically combine sheets to find matching parts across manuscripts

ÖAW AUSTRIAN ACADEMY OF SCIENCES

Outlook

ARI

- Standardize thermographic watermark digitization: DFG round table in October 2023
- Ongoing effort to binarize watermark from the “raw” images:
  - Use classical image processing techniques
  - Use machine learning methods
- Automatically combine sheets to find matching parts across manuscripts
- Cluster watermarks based on mathematical distance measures

ÖAW AUSTRIAN ACADEMY OF SCIENCES

Outlook

ARI

- Standardize thermographic watermark digitization: DFG round table in October 2023
- Ongoing effort to binarize watermark from the “raw” images:
  - Use classical image processing techniques
  - Use machine learning methods
- Automatically combine sheets to find matching parts across manuscripts
- Cluster watermarks based on mathematical distance measures
- What does the community want? Which data should we provide?