# Introduction

## Historical Document Analysis by Non Destructive Testing

For many historical documents, the standard procedure of using a book scanner is not applicable, because the document would get damaged by the page-turning process. A lot of those old manuscripts where written with iron gall ink, which is used since the 5th century until present [1]. This kind of ink consists of metallic particles. When performing a 3-D X-ray CT scan, those metallic particles allow us to see the writings as their X-ray absorption differs from the paper’s [2].

## Materials and Methods

### Book Model
- 22 pages of handmade paper, 150 µm thickness
- Iron gall ink writings
- Leather cover

### 3-D X-ray CT Scan and Reconstruction
- Cone beam geometry, book upright
- 3-D CT scan with 800 projections over 360°
- X-ray energy 50 kVp, detector pixel size 44 µm²
- FDK reconstruction, 44 µm³ voxel size

## Page Extraction Algorithm

![Fig.2. 3-D Volume Processing Pipeline demonstrating the process of the 2-D page mapping.](image)

**Fig.2.** shows that a single page can not be easily investigated due to its waviness in the volume. A segmentation by hand is very time consuming.  
- Need of an automatic algorithm for extraction of pages [3].

## Results and Discussion

- Fig.5. shows three exemplary original pages and the 3-D reconstructed and extracted 2-D mapped pages.  
- All 22 pages of the volume were extracted properly. Also tests on other scanned books showed similar results.  
- The algorithm converged after 24 iterations resulting in a total computation time of around 5.5 minutes.  
- A post-processing step for image denoising (guided filtering) and unsharp masking can improve quality.

## Conclusions and Outlook

- Manuscripts can be scanned by 3-D X-ray CT without opening or page-turning.  
- Proposed algorithm extracts and 2-D maps all pages.  
- Tested also for extreme cases (wavy and thin pages).  
- No user interaction needed and low computation time.

Future work:  
- Find the best scan parameters, best reconstruction algorithm and take real historical documents.

## References