

Fast, Robust and Efficient Extraction of Book Pages from a 3-D X-ray CT Volume

Daniel Stromer, Vincent Christlein and Andreas Maier
Pattern Recognition Lab,
Friedrich-Alexander-University Erlangen-Nuremberg, Germany
Email: daniel.stromer@fau.de, vincent.christlein@fau.de,
andreas.maier@fau.de

Tobias Schoen and Wolfgang Holub
Fraunhofer Dev. Center for X-ray Technology,
Fuerth, Germany
Email: tobias.schoen@iis.fraunhofer.de,
wolfgang.holub@iis.fraunhofer.de

Abstract—Investigating historical documents makes it necessary to use special imaging systems. We already showed in an earlier work that it is possible to use common 3-D X-ray CT scanners for the reconstruction of historical documents written with iron gall ink. Our tests were based on a self-made book which was scanned and investigated without opening or page-turning. However, when analyzing the reconstruction results, we faced the problem of a proper automatic extraction of single pages within the volume in an acceptable time without losing information of the writings. In this paper, we present a robust and efficient algorithm for the extraction of book pages from the original 3-D volume. This step is a necessary prerequisite for a possible identification of the writing. Our method delivers high quality results for our book model and can be easily adapted to other imaging modalities. We show that it performs well even for an extreme case with low resolution input data and wavy pages.

Index Terms—Iron Gall Ink Reconstruction; Historical Document Analysis; Page Extraction; 3-D X-ray CT Reconstruction

I. INTRODUCTION

For many historical documents, the standard procedure of using a book scanner is not applicable, because the document would get damaged by the page-turning process, even when the page-turning is done by an air flow. Only very few methods exist for imaging and reconstructing such documents. One approach is to use X-ray phase contrast imaging [1] [2] with the downside of being complex, expensive and immobile. Another approach is based on Terahertz radiation [3] which has the drawback of a limited depth resolution meaning that only a few pages can be imaged at once. Our former work showed that using a common C-arm X-ray CT system, such as it is used for medical imaging or material testing, is also capable of imaging such documents [4]. We took advantage of the fact that iron gall ink [5], an historic ink that has been used since the 5th century until present [6] [7], consists of metallic particles. These particles absorb more radiation than the cellulose-based book pages allowing to differentiate the ink from the paper giving the ability to read the writings without opening the book itself. The scanners are mobile and cheap compared to the Synchrotron system as well as the field-of-view is only limited to the detector size. A major challenge when visualizing such documents by using 3-D imaging techniques is the extraction of single pages to inspect

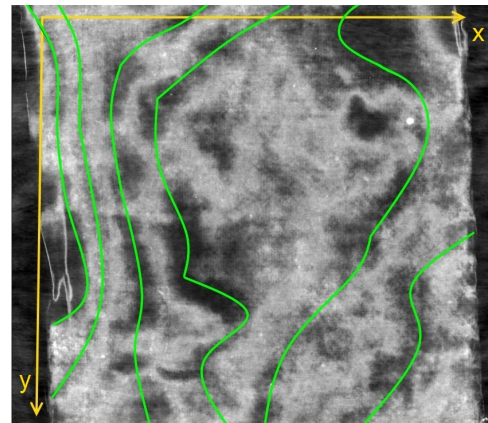


Fig. 1. Original 3-D volume center slice acquired by a micro-CT scan. The green lines denote the air gaps.

the writings on the paper. Not only that the rather thin pages are stacked together in a very small area (sometimes more than 100 pages per *cm*) and squeezed together at many points, but they are also wavy and the writings have a higher absorption than the paper. It is hardly possible to identify writings or single pages in the output volume of a reconstructed X-ray 3-D scan. This can be seen in Fig. 1 showing the center slice of our generated volume where the green lines denote air gaps.

Initial works based on the voxel intensities and B-Spline fitting produced poor results such that rectification had to be done. Seals et al. showed that using the structure tensor performs good for scrolls [8], but it requires high computational time and has not been applied to books yet. This work shows a simple and efficient algorithm with low computational time capable of extracting a single page by only selecting two points on the page-of-interest. The algorithm also has a good performance with small paper thicknesses down to 2 – 3 pixel per page, wavy pages, overlaps as well as fully and partially ink-saturated paper.

II. MATERIALS AND METHODS

A. The Book Model

We built a book model consisting of ten different pages. The paper is handmade and has a thickness of about 200 μm .

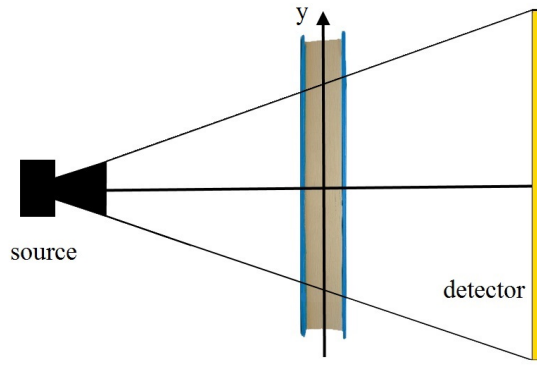


Fig. 2. Experimental setup for the 3-D CT scan. The rotation axis is the y -axis of the book.

A pressboard holder is used to keep the pages close together during the X-ray scan with the side effect of lowering the air gaps between the pages. With a quill, letters of different font sizes, line thicknesses and paper saturations have been written on the pages. The height of the letters range from 1 cm to 2.5 cm . Fig. 7(a), Fig. 7(b) and Fig. 7(c) show three selected pages and its writings. For the following investigations, we exemplarily show the restoration of these pages by the use of our proposed extraction algorithm.

B. CT Scan Trajectory and 3-D Reconstruction

For this work, we performed a 360° circular scan with 2400 projections using cone-beam geometry with a micro-CT system built for non-destructive material testing. The detector size is $2864 \times 2864\text{ px}$ with a pixel shift of $80\text{ }\mu\text{m}$. We set up an energy of 70 kVp which is enough to see larger letters. The book model was positioned upright as shown in Fig. 2 and the y -axis denotes the axis of rotation.

The common 3-D reconstruction algorithm for cone beam geometry is the FDK algorithm [9] consisting of three major steps – cosine weighting, ramp filtering and back projection. With the large number of projections used, the algorithm should be accurate enough. We defined a voxel size of $68\text{ }\mu\text{m}^3$ so that a single page has a thickness of about $3 - 4\text{ px}$ which is enough for the presented extraction algorithm. The energy, the pixel shift, and the voxel size can be further decreased, but as the aim of this work is the page extraction, we wanted to keep the acquisition and reconstruction time relatively low (about 1.5 hours in total).

C. Page Extraction Algorithm

All calculations were performed with the CONRAD framework for cone beam geometries [10]. The algorithm consists of three major steps:

1) *Mask Creation*: Looking at a snippet of the original 3-D volume \mathbf{V}_O in Fig. 3 shows that the writings on the page can not be clearly identified with the human eye due to the waviness and overlaps of the pages. The black spots are the air gaps while the paper appears bright. Every slice of \mathbf{V}_O only



Fig. 3. Original 3-D volume slice snippet showing page 6, 7 and 8 and the orange arrow points on letter 'L'.

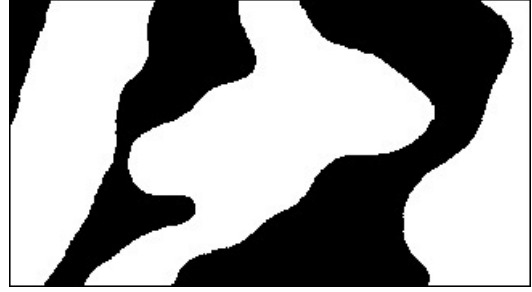


Fig. 4. Created 3-D mask slice at the position of Fig. 3

covers a small area of the page that should be inspected. One can see parts of page 6, 7, and 8. A closer look also reveals parts of the capital letter 'L' of page 7 (Fig. 7(c)) denoted by the orange arrow.

At first, a binary mask volume \mathbf{V}_M is created where the air gaps of \mathbf{V}_O should be set to 0 and the pages to 1. At this stage, we do not care about the writings and assume that they are part of the page. We apply a median filter in xy -direction on each slice with a kernel size of 20×20 to smooth and denoise the volume [11]. At next, we binarize the volume. Due to the differing exposure of X-ray imaging, a global thresholding, such as Otsu's method [12], is no option. Adaptive Gaussian Thresholding [13] is a method where the threshold value is calculated by the neighborhood's Gaussian weighted average, solving the problem in a proper manner. It is applied to the xz -layer with a radial kernel size of 8 and $\sigma = 2$ which has emerged as the best fitting kernel size for our given problem derived from double the page thickness (approximately $3 - 4\text{ px}$). Fig. 4 shows the resulting mask snippet of the original Fig. 3.

2) *Page Segmentation*: The second step is considered to be the key procedure of the algorithm. Briefly summarized, the algorithm extracts the page from every slice s of \mathbf{V}_M . The resulting output \mathbf{R} should be the position of the page in the volume for each slice. The segmentation of one slice works as follows:

Initially, the user has to select two points p_1 and p_2 on the page-of-interest. For each starting point, the algorithm walks up and down along the page separately. One step in a direction includes searching the left and right border b_l and b_r of the

page, calculating its thickness $t = b_r - b_l$ and the center point p_c by applying

$$p_c = \begin{cases} b_l + \frac{t}{2}, & \text{if } t < 2 \cdot \bar{t} \text{ or Right Overlap} \\ b_r - \frac{t}{2}, & \text{else} \end{cases}, \quad (1)$$

where \bar{t} is the mean page thickness. If another pages touches the current, the page overlap direction is detected by comparing the new center point p_c with the previous center point $p_{c_{old}}$ ($p_c > p_{c_{old}}$: Right Overlap, else: Left Overlap). The up and down walking procedure can easily be implemented in parallel minimizing the computation time. As a result, the images I_{p_1} and I_{p_2} containing the segmented pages for each point are generated. The next step is merging those images to a result image I_R . Three scenarios can occur during the merge process:

- A The centerlines of I_{p_1} and I_{p_2} match completely, delivering a perfect solution. One image is considered to be the result I_R at slice s .
- B The centerlines of I_{p_1} and I_{p_2} do not match completely but have a connection point $p_m(x, y)$. Considering that p_1 is the upper and p_2 is the lower point on the page, I_R is formed by copying the upper image part from I_{p_1} (0 to $p_m(y)$) and the lower part from I_{p_2} ($p_m(y) + 1$ to y_{max}).
- C If no connection point is found, the centerlines of I_{p_1} and I_{p_2} are compared to the stored centerline of the previous slice with regard to their Euclidean distance. For every column y , the image with the lower distance is saved as result image for this slice.

Afterwards, the new center line is stored and the next mask slice can be processed. The standard procedure is to take the old starting points again, which is reasonable because a small voxel size implies only small movements for each slice. However, if a new starting point is 0 in the mask, the nearest neighbor is searched and considered. The resulting volume contains the segmented page which gets median filtered in xy -direction again to suppress outliers and smooth the overlaps resulting in the final mask V_S .

3) *2-D Mapping*: The last step is to multiply the final binary mask V_S with the original volume V_O . A 2-D image is created by only considering the maximum value of the page in z -direction. Another option would be to take the average, but since the ink has a higher intensity than the paper, the maximum shows better results regarding the visualization of the writings. Additionally, we applied a unsharp mask filter ($r = 100, \sigma = 0.4$) [14] for highlighting the writings.

III. RESULTS

At first, we tested the segmentation process. A test image has been created combining two central problems – the center page is wavy and touches its neighboring pages at different points. After selecting two starting points p_1 and p_2 the segmentation algorithm has been applied. Fig. 5 shows the results. Please note that in terms of visualization the image

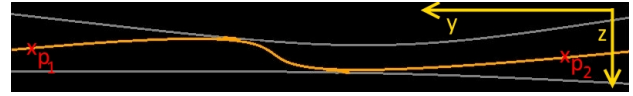


Fig. 5. Test image of a wavy center page touching two other pages. After applying the segmentation, the wavy page is correctly segmented (orange) while the overlapping pages (gray) are not considered. p_1 and p_2 denote the selected start points.

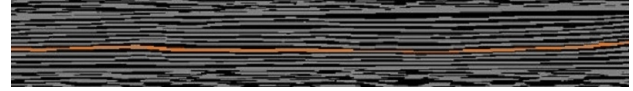


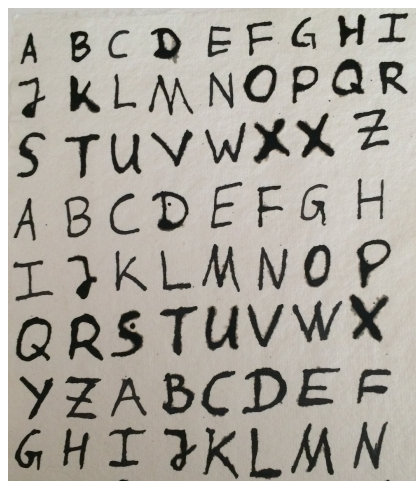
Fig. 6. Segmented page 7 of the original volume in a center slice showing that the algorithm also works on real data.

has been rotated by 90° . The segmented, orange denoted page has been extracted properly while cropping the overlapping, gray denoted pages. Fig. 6 demonstrates that our segmentation algorithm segments the page robustly also in the real case. At next, the algorithm was applied on the entire X-ray volume and all pages were extracted separately. Fig. 8(a), Fig. 8(b) and Fig. 8(c) show the 2-D mapping results of three different pages of our book model (Fig. 7). In comparison to Fig. 1 each page's writings can be identified and assigned to the respective page with the naked eye showing that pages were extracted properly. Missing letters are caused by the relatively high radiation energy of 70 keV . Tests revealed that using lower energies improves the visibility of the ink, however, this it was out of the scope of this work to display all writings perfectly. The algorithm was tested for all ten pages of the phantom and always performed well. For a volume with the size of $1100 \times 1000 \times 110$ the extraction of one page took about 6 minutes. The segmentation itself took about 5 minutes. The calculations were performed on a standard notebook with a NVIDIA Quadro M1000M graphic card and 32 GB RAM. However, the algorithm does not work well on pages where the thresholding detected ink as air and split the page on a large range.

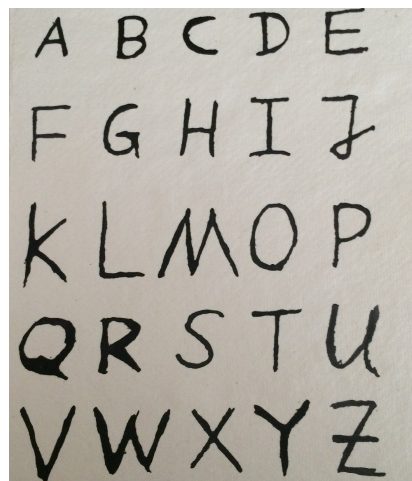
IV. CONCLUSION AND OUTLOOK

We developed an algorithm to extract a single book page, acquired with a micro-CT system, used for material testing. The algorithm works robust and efficient such that all pages of our book model could be extracted and investigated. This kind of filtering also suppresses spots where the wrong page has been selected due to multiple overlaps at the same position. The algorithm can easily be adapted to other imaging modalities, as it is an intensity based, region growing segmentation algorithm.

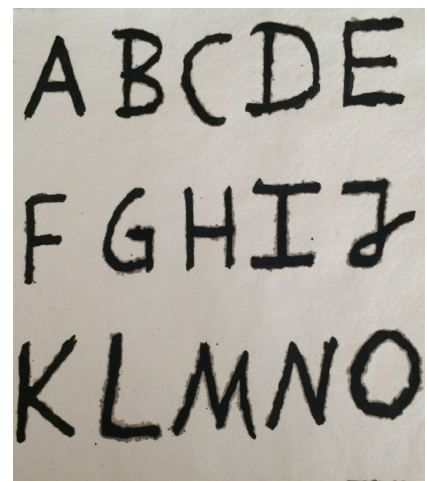
In Future, we want to test different filtering methods for preprocessing as well as other local thresholding methods. We also want to test our algorithm on real historical documents, where not only books are the focus but also scrolls. Another aim is testing the method with different radiation energies, trajectories, detector pixel shifts and voxel sizes.



(a) Page 1



(b) Page 2

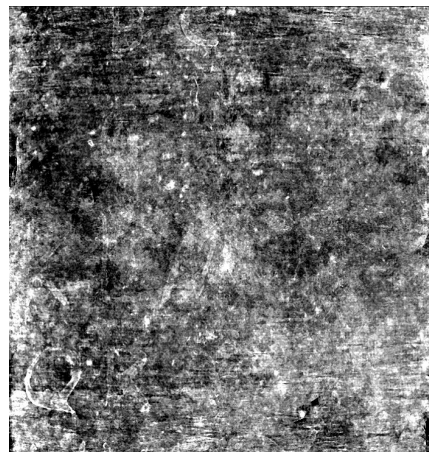


(c) Page 7

Fig. 7. Original Pages of the book model notated with letters written with iron gall ink and a quill.



(a) Page 1



(b) Page 2



(c) Page 7

Fig. 8. Extracted and 2-D mapped pages of the 3-D reconstructed X-ray volume after applying the presented algorithm. While the writing is hardly readable (which is out of the scope of this paper), it is clear that the characters are from the respective page.

REFERENCES

- [1] V. Mocella, E. Brun, C. Ferrero, and D. Delattre, "Revealing letters in rolled herculaneum papyri by x-ray phase-contrast imaging," *Nature communications*, vol. 6, 2015.
- [2] L. Glaser and D. Deckers, "The basics of fast-scanning xrf element mapping for iron-gall ink palimpsests," *Manuscript cultures*, vol. 7, no. PUBDB-2015-06320, pp. 104–112, 2014.
- [3] A. Redo-Sanchez, B. Heshmat, A. Aghasi, S. Naqvi, M. Zhang, J. Romberg, and R. Raskar, "Terahertz time-gated spectral imaging for content extraction through layered structures," *Nature Communications*, vol. 7, p. 12665, 2016.
- [4] D. Stromer, V. Christlein, G. Anton, P. Kugler, and A. Maier, "3-D Reconstruction of Historical Documents using an X-Ray C-Arm CT System," in *Proceedings of the 31th conference on Image and Vision Computing New Zealand 2016*, M. University, Ed., 2016.
- [5] A. Stijnman, "Historical iron-gall ink recipes: art technological source research for inkcor," *Papier Restaurierung*, vol. 5, no. 3, pp. 14–17, 2004.
- [6] O. Hahn, W. Malzer, B. Kanngiesser, and B. Beckhoff, "Characterization of iron-gall inks in historical manuscripts and music compositions using x-ray fluorescence spectrometry," *X-Ray Spectrometry*, vol. 33, no. 4, pp. 234–239, 2004.
- [7] Staatshochbauverwaltung and G.-A. Universität des Landes Niedersachsen, *Niedersächsische Staats- und Universitätsbibliothek Göttingen: Dokumentation des Neubaus zur Eröffnung a 30.4. 1993*. Niedersächsische Staats- und Universitätsbibliothek, 1993.
- [8] W. B. Seales, C. S. Parker, M. Segal, E. Tov, P. Shor, and Y. Porath, "From damage to discovery via virtual unwrapping: Reading the scroll from en-gedi," *Science Advances*, vol. 2, no. 9, p. e1601247, 2016.
- [9] L. Feldkamp, L. Davis, and J. Kress, "Practical cone-beam algorithm," *JOSA A*, vol. 1, no. 6, pp. 612–619, 1984.
- [10] A. Maier, H. Hofmann, M. Berger, P. Fischer, C. Schwemmer, H. Wu, K. Müller, J. Hornegger, J.-H. Choi, C. Riess, A. Keil, and R. Fahrig, "CONRAD - A software framework for cone-beam imaging in radiology," *Medical Physics*, vol. 40, no. 11, 2013.
- [11] A. Maier and R. Fahrig, *GPU Denoising for Computed Tomography*, 1st ed., Boca Raton, Florida, USA, 2015, vol. 1.
- [12] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, no. 285–296, pp. 23–27, 1975.
- [13] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (2nd Ed)*. Prentice Hall, 2002.
- [14] W. K. Pratt, *Digital Image Processing*. Wiley, 1991.