Three-dimensional book data page segmentation and extraction method using Laplace equation

Jiarui Ou\textsuperscript{1,*}, Zhongjiang Han\textsuperscript{2}, Koji Koyamada\textsuperscript{3}

\textsuperscript{1}Academic Center for Computing and Media Studies, Kyoto University, Japan
\textsuperscript{2}Graduate School of Engineering, Kyoto University, Japan
\textsuperscript{3}Academic Center for Computing and Media Studies, Kyoto University, Japan

\textsuperscript{*}oujiarui@outlook.com

Received: January 9, 2020; Accepted: October 29, 2021; Published: November 25, 2021

Abstract. Our work aims to interpret historical documents in closed booklet form without physical unfolding through CT scanning techniques. X-ray based CT scans can transform booklets into physical data for analysis, and current research in this direction performs page segmentation at the two-dimensional level and does not analyze data from historical documents at the three-dimensional level. Our proposed method uses a self-developed web tool for 3D dimensional annotation and solves Laplace’s equation to construct the 3D spatial structure of historical documents and perform page segmentation. We experiment with data in booklet form and show how to segment and extract pages in three dimensions by using Laplace’s equation.

Keywords: CT, Book Data, Page Segmentation, Laplace Equation

1. Introduction

Recently, research cases using CT scanning devices have been published regarding the decoding of old literature that has become fragile and cannot be opened. Recently, methods for digitizing ancient documents buried in the ashes of Mount Vesuvius using phase CT scan data and particle synchrotrons, have been proposed. The analytical approach to digital literature is constant, and we are developing an analysis method that assumes the literature is a booklet.

X-rays are widely used in non-invasive research related to the analysis of historical documents. For example, Oksana Samko [1] et al. used x-rays to digitize historical parchments and perform virtual unfolding. Inna Bukreeva [2] et al. used x-rays to successfully parse a series of burned historical documents. However, the methods proposed in these studies are still based on analysis on two-dimensional images and do not perform page segmentation on a three-dimensional level. Therefore we propose a method to perform page segmentation on three dimensions using Laplace’s equation to fit the spatial structure of the target data.
This study developed a method to efficiently extract page data from 3D booklet images acquired using 3D CT equipment using the Laplace equation and to evaluate its effectiveness. Our purpose is to propose one method with strong operability and versatility, which can improve the accuracy of page extraction while dealing with more complex data. How to correctly extract page information from the 3D images generated by the booklet is our main academic issue. We propose the following hypothesis that if each page surface of a booklet is treated as an isosurface and a scale field is generated appropriately, the page information can be correctly extracted from the three-dimensional image. The booklet is a multi-page printed in ink. The main purpose of binding is to display the pages that make up the booklet without flipping the page.

2. Methods

The advantage that nondestructive testing CT technology provides leads one way to the solution of the 3D booklet reconstruction problem, which is helpful to digitize ancient books. In this research, we explored extracting each page from CT scanned 3D data of the booklet type by establishing and solving one Laplace equation with the principle of electrostatic field analysis among capacitors. Having generated the order of pages, we could separate page models. We make an overview of the research in Fig. 1.

In the case of the absorption X-ray CT used in this research if there is not enough difference in the linear absorptance between paper and ink, the difference between the ink areas and the paper areas is hardly recognized [3, 4]. In such a case, to increase the contrast in the linear absorptance, the wavelength of the X-ray is changed within a possible range, and the wavelength that maximizes the contrast is specified for a given paper and ink pair [5]. In the previous research, we investigated the wavelength-dependent characteristics of the linear absorptance of printing paper and ink distributed in the market and investigated the combination of printing paper and ink. We have developed a method that can identify the wavelength that maximizes the contrast. In this research, ink containing metal elements is prepared, a booklet composed of printed pages is prepared, and a three-dimensional image with a resolution of 512 x 512 x 276 is created from this. Specifically, the CT to be used this time is inspeXio SMX-90CT Plus manufactured by Shimadzu Corporation, and the booklet
sample is A4 paper and ink containing 40-50% iron oxide (Canon Cartridge 039 H). In this section, we perform a page extraction experiment based on solving the Laplace equation for this data to verify its feasibility in the research of page segmentation extraction.

2.1. Page extraction

Considering that the structure of the booklet pages we analyzed has a high degree of similarity to the plate capacitor structure in the electrostatic field problem, we made the hypothesis “page surfaces of the booklet could be extracted by one scalar field describing them as isosurfaces”. Therefore, this scalar field could be generated by the Laplace equation, by which we could also solve the electric potential problems inside capacitors. Since the undetermined coefficient of the exact solution of the Laplace equation is determined from the boundary condition, the 3D coordinates with the page number should be specified in advance. That’s also the sampling process for which we have designed the GUI (Graphical User Interface) above. Above all, the exact solution of the Laplace equation determined in this way is calculated to generate a scalar field, the page surface is extracted as an isosurface and a 3D image will be mapped to the isosurface.

For the Laplace equation, it could be shown as follows in 3D space of the Cartesian coordinate system:

\[
\nabla^2 \psi = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} = 0
\]

(1)

For the Laplace equation above as a linear partial differential equation, the principle of linear superposition is satisfied. In our experiments, considering the surface should be a developable surface, we determined the solution form as

\[
\psi = (A_1 e^{\sqrt{-\lambda_1} x} + B_1 e^{-\sqrt{-\lambda_1} x} + C_1 x + D_1) \\
\times (A_2 e^{\sqrt{-\lambda_2} y} + B_2 e^{-\sqrt{-\lambda_2} y} + C_2 y + D_2) \\
\times (A_3 \cos \sqrt{-\lambda_3} z + C_3 z + D_3)
\]

(2)

In the experiment, we used the previous UI (User Interface) to collect coordinate points and used the above formula to estimate the parameters of the scalar field. In Equation 2, \(x\), \(y\), and \(z\) represent the coordinates in three-dimensional space. \(\psi\) represents the number of pages where the corresponding coordinate points are located. By substituting the annotation data set \((x, y, z, \psi)\) into the formula [6], the values of the other parameters can be approximated using the least-squares method. At this time, the three-dimensional spatial structure is represented by Laplace’s equation and can be approximated as the spatial distribution of the book pages.

2.2. Annotation tool

In this section, we build an interactive environment for page numbers and 3D coordinates for specifying points within the booklet using WebGL 3D graphics technology. The number of pages and the sufficient number of three-dimensional coordinates of the interior points
of the booklet is specified in the interactive environment. The exact solution of the Laplace equation determined in this way would be used to generate a scalar field, extract the page surface as an isosurface, and map the three-dimensional image to the isosurface.

The WebGL (Web Graphics Library) has become one new standard of Web 3D graphics protocol, which could create customized working environments of various projects calling for complex 3D graphic rendering and interaction with users. This tool combines the advantages of JavaScript and OpenGL ES 2.0 (OpenGL for Embedded Systems). We develop scripts to render the scene in HTML5 (HyperText Markup Language) Canvas with the hardware rendering acceleration by OpenGL. It is beneficial for promoting the program, we built on the internet because HTML script does not require any extra-plugin beyond the browser itself while the OpenGL is also a cross-platform standard.

With this technique, we built the working platform to visualize the whole 3D volume data containing the booklet information scanned by the CT machine. It means that we provide one 3D space following the principle of perspective and set the main object (volume data) in the center of the space and the camera above it at the very beginning. The user could do some basic interactions through simple actions, like rotating the object by dragging with the left button, adjusting the distance from the camera to the object by scrolling the mouse wheel, and moving the object by dragging it with the right button. With those actions, the user could check detail structures on of the object. To observe the internal structure of the object, we made three slices perpendicular to each other with changeable coordinates, which is shown in Fig. 2. As shown in Fig. 3, to specify the boundaries of pages and make it easier to obtain the samples of the feature points on the page, the function to specify the range of gray values that make the points whose gray value is out of that range invisible. By testing, the users could catch a very clear vision on the page boundary in the local image by limiting the range, even in some complex structures.

In the very first version of our program, the user should select the points on the slice as the point on the actual pages by double click on the points, then the 3D coordinates will be recorded with their page numbers in the CSV file, which is also accessible in the console.
of the browser for users’ checking. However, we found this way of sampling is of very low efficiency. After the redesign, a sphere following the pointer is created with an adjustable radius. The user should set an appropriate radius, which is shorter than the distance from the selected point on the page to any other points on the next pages or other visible points. After that, all the points in this sphere should be on the exact page we are searching for, then the information of coordinates about the group of points will be recorded in the CSV file when users double-clicking on it. The option of turn on and off the rotating function was also added in the control bars, considering the conflict between rotating function by single click and sampling by double click.

Figure 3: Only the page borders are displayed by adjusting the gray scale range (a); Adjust the radius of the sphere for sampling (b); The labeled coordinates can be viewed on the console screen (c)

According to the obtained three-dimensional coordinates, we can obtain a series of parameters by solving the Laplace equation [7, 8, 9]. This allows us to calculate the scalar of each coordinate point in a three-dimensional space of the same size as the volume data and generates a scalar field to simulate the distribution of pages in the volume data.
As in Fig. 4, when collecting the coordinates of the internal points, each page is given a different mark, so that in the generated scalar field, the scalar values of these coordinate points are also equal to the mark value (page number) of the page.

By generating such a scalar field, we can extract the pages of corresponding values (number of pages) by extracting the isosurface of different values. Finally, we corresponded the isosurface coordinates with the data in the volume data and extracted each page in the volume data according to the value of the isosurface.

3. Results

In this section we will describe the experiment with its results, the process, and considerations for labeling the data, generating a scalar field using the derived Laplace equation, and extracting each page using the scalar field.

3.1. Annotation

We use the annotation tool described in the previous section to perform page annotation on volume data. Thanks to the range annotation function, we can annotate approximately 500 points per minute. We have summarized the following notes to reduce the errors for the subsequent experiments. First, the selected annotation points should be distributed as evenly as possible in the volume data. This is to make the spatial structure fitted by the derived Laplace equation coincides with the actual spatial structure as much as possible during the calculation process. Second, the number of labeling points should be appropriate. Too many data points can make the process of deriving Laplace’s equations difficult, i.e., the errors in the derived equations are too large. Too few data points will make it difficult to fit the exported equations to the page structure in detail, which will make the extracted pages have
errors in detail. The above error analysis will be discussed in a subsequent section of the results.

3.2. Scalar Field Generation

As shown in Fig. 5, we sample the coordinates of 3515 points to calculate the Laplace equation and reduce the total error (the square of the difference between the value of each point and the number of pages) to 1400. We visualize the resulting scalar field and fit each page by generating isosurface.

![Figure 5: Visualization results of scalar field isosurface](image)

By visualizing the results, we can conclude that if the error of the scalar field obtained by the parameters of the Laplace equation is small enough, it can fit each page in the volume data [10, 11]. We calculated from the coordinate values of different numbers of points and obtained various scalar field parameters. We searched for the optimal solution among these parameters to extract the corresponding pages in subsequent experiments. Based on the results, it is easy to extract each page from the volume data through the program and evaluate it in subsequent experiments to determine the feasibility of the research method.

Regarding the appropriateness of the scalar field constructed using the interactive environment, a three-dimensional image is generated using the exact solution of the Laplace equation for which the undetermined coefficient is predetermined, and the difference from the coefficient is determined through this interactive environment is calculated. The closer the value (error) is to zero, the more appropriate the scalar field is generated. Using this result, it is possible to determine a criterion for the number of internal points to be specified that are necessary for generating the scalar field.

We experimented with different numbers of internal points extracted from the same individual data. By determining the parameters of the exact solution of the Laplace equation, the error in the scalar field is minimized as much as possible. The data in the above exper-
The experimental process shows that the larger the number of internal points, the easier it is to fit the curved shape of the page, but simultaneously, it will also cause the parameters of the Laplace equation that cannot perfectly match the volume data. As shown in Fig. 6, when using the coordinate data of 881 internal points to calculate and reduce the error to 716, although the overall brochure shape is consistent with the real situation, the details are warped, which makes it difficult to extract the pages. It causes the loss of data, which will greatly reduce the experimental results.

Figure 6: Isosurface visualization results when the number of points is insufficient

The conclusion is that when sampling points, it is necessary to collect as uniformly as possible. Simultaneously, to ensure the fitting effect, it is necessary to collect the feature points of the page (sudden distortion, flicking). Based on this, we must collect more points while ensuring the validity of the Laplace equation parameters.

4. Discussion

According to the results, we successfully obtained the coordinate range of each page in the book data through the calculated scalar field. Based on these results, we can easily extract each page of the book data from the volume data and visualize it.

4.1. The calculated field

With the derived Laplace equation, we can fit a scalar field in a computer. As shown in Fig. 7, we selected one of the facets and highlighted the field range of the corresponding page separately. The results show that in the case where the scalar field represented by Laplace’s equation has little error with the real page structure, we can extract the page fields from real data by selecting page numbers in the scalar field.
Simultaneously, in the boundary region of each page, the page fit in the boundary region may be biased because the points are not completely taken to the outermost points, but this does not affect the final page extraction results, which will be shown in the next section.

In the actual situation, we found that because the parameters that generate the scalar field are not unique, this will cause the scalar field to have some errors with the real volume data. This result will cause us to specify the number of pages we want to extract, not a certain number, but a range. For example, we extracted the fourth page in the experiment. However, in the scalar field, the value of the coordinates representing the data of the fourth page will fluctuate between 3.4 and 4.1 due to the error. This direct result of this error is that the extracted page will be thicker than the original page, and when the page gap is small, the data of other pages will also be included. However, we believe that this problem can be solved by reducing the error of the Laplace equation parameters.

4.2. The tolerance of page extracting

For dealing with the problem above, we must set one tolerance of errors to make sure to keep the valid data from discarding. This means that the extracted volume data of one page should not be a single layer with the exact $\psi$, but one region with sufficient thickness contains those valid data, which call for one range $[\psi - \delta_1, \psi + \delta_2]$ when extracting page data instead of $\psi$. To show this model with tolerance, we do one experiment on page 4, as shown in Fig. 8.

As shown in the result, the right range would lose much valid data, but when the range was widened into the left one, it owned a good ability to contain the entire page data.
4.3. The visualization of pages

The output of the extracted pages above is series of coordinates of points. Then, the problem comes that can we judge the result. For reflecting the main information of those data, we projected the volume data onto one plane. More detailed speaking, we do the compression through the thickness direction, using the average of gray values, and increase the contrast of the image. The images generated from the 7 pages of our sample are shown in Fig. 9.

We found that the results are identifiable as simple text, and they could be recognized by the OCR (Optical Character Recognition) [12] system from Google with correct contents. The output so far is acceptable in the OCR system of Google. The recognition results with the original page image are shown in Table 1.

As the form shows, all of the internal texts could be recognized correctly, except the incomplete letter “u” misrecognized into “I”. The test samples we used in this experiment are single-sided printing. In a further study, double-sided printing will be tested. After all, considering the conditions of old books, it is proved that the data obtained from this method preform will be in OCR, and are theoretically universal. However, due to the scarcity of CT data in books and experimental funding, there are currently no more experiments for even more complex conditions testing.

5. Conclusion

We provide a novel approach for page segmentation and extraction of data in booklet form without physical unfolding. This is the first attempt to simulate the spatial structure of booklets and page extraction at the 3D level in a similar study. In contrast to the analysis on the 2-dimensional section, we perform 3D annotation from the volume data and fit the 3D spatial page distribution of the booklet by solving Laplace’s equation to complete the page extraction. This method gets rid of the limitation of 2-dimensional data and fits the volume data in terms of spatial structure.

In the study, we used the booklet data for experiments and completed the extraction, and we will use more data for experiments and optimize the proposed method in future studies.
Figure 9: The projection of the page (a) to page (g) on plane

Furthermore, the deep learning network is also being taken into consideration for possible promotion in our research.
Table 1: Result comparison

| Output of our system | Original page image | Output of OCR |
|----------------------|---------------------|---------------|
| &Image;              | &Image;             | alization     |
| &Image;              | &Image;             | alization     |
| &Image;              | &Image;             | alization     |
| &Image;              | &Image;             | alization     |
| &Image;              | &Image;             | alization     |
| &Image;              | &Image;             | alization     |
| &Image;              | &Image;             | alization     |
| &Image;              | &Image;             | alization     |
| &Image;              | &Image;             | alization     |
References

[1] Oksana Samko, Yu-Kun Lai, David Marshall, and Paul L. Rosin. Virtual unrolling and information recovery from scanned scrolled historical documents. *Pattern Recognition*, 47(1):248-259, 2014.

[2] I. Bukreeva, A. Mittone, A. Bravin, G. Festa, M. Alessandrelli, P. Coan, V. Formoso, R. G. Agostino, M. Giocondo, F. Ciuchi, M. Fratini, L. Massimi, A. Lamarra, C. Andreani, R. Bartolini, G. Gigli, G. Ranocchia, and A. Cedola. Virtual unrolling and deciphering of Herculaneum papyri by X-ray phase-contrast tomography. *Scientific Reports*, 6:1–7, 2016.

[3] P. Campadelli, E. Casiraghi, and G. Lombardi. Automatic liver segmentation from abdominal ct scans. In *14th International Conference on Image Analysis and Processing (ICIAP 2007)*, 731–736, 2007.

[4] Akinobu Shimizu, Rena Ohno, Takaya Ikegami, Hidefumi Kobatake, Shigeru Nawano, and Daniel Smutek. Segmentation of multiple organs in non-contrast 3d abdominal ct images. *Int. J. Computer Assisted Radiology and Surgery*, 2:135–142, 2007.

[5] T. Shinohara, J. Takayama, S. Ohyama, and A. Kobayashi. Analysis of textile fabric structure with the ct images. In *SICE 2003 Annual Conference (IEEE Cat. No.03TH8734)*, 1:428–432, 2003.

[6] Ken ichi Yoshida, Naoshi Nishimura, and Shoichi Kobayashi. Application of new fast multipole boundary integral equation method to crack problems in 3d. *Engineering Analysis with Boundary Elements*, 25(4):239–247, 2001.

[7] B. Boroomand, S. Bazazzadeh, and S.M. Zandi. On the use of laplace’s equation for pressure and a mesh-free method for 3d simulation of nonlinear sloshing in tanks. *Ocean Engineering*, 122:54–67, 2016.

[8] Hiroshi Serizawa. River channel networks created by poisson equation and inhomogeneous permeability models. *Journal of Advanced Simulation in Science and Engineering*, 4(2):176–208, 2017.

[9] Michael Kazhdan and Hugues Hoppe. Screened poisson surface reconstruction. *ACM Trans. Graph.*, 32(3):13, 2013.

[10] Ragnar Bade, Jens Haase, and Bernhard Preim. Comparison of fundamental mesh smoothing algorithms for medical surface models. In *SimVis*, 1:289–304, 2006.

[11] Y. Y. Boykov and M. Jolly. Interactive graph cuts for optimal boundary region segmentation of objects in n-d images. In *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, 1:105–112, 2001.
[12] Tomas Mikolov, Anoop Deoras, Daniel Povey, Lukas Burget, and Jan Cernocky. Strategies for training large scale neural network language models. *2011 IEEE Workshop on Automatic Speech Recognition and Understanding, ASRU 2011, Proceedings*, 196–201, 2011.