

令和元年6月25日現在

機関番号：14301

研究種目：挑戦的研究（萌芽）

研究期間：2017～2018

課題番号：17K19974

研究課題名（和文）三次元データ可視化技術を使った冊子体分析手法の開発

研究課題名（英文）Development of analysing a folded book using 3-D data visualization techniques

研究代表者

小山田 耕二（KOYAMADA, Koji）

京都大学・学術情報メディアセンター・教授

研究者番号：00305294

交付決定額（研究期間全体）：（直接経費） 5,000,000円

研究成果の概要（和文）：本研究では、冊子体ボリュームデータから、冊子体を構成するページが必ずしも平面でなくても、1枚のページであると認識する手法を開発した。具体的には、研究代表者が開発したActive Grid法を使い、与えられた3次元CTボリュームデータから冊子に対応する空間領域を直方体格子に写像することができた。Active Gridは、三次元格子モデルを前提とし、ボリュームデータ内部の三次元領域に適合するように各格子を変形させる。格子を変形させるために格子の内部エネルギーを対象にしたエネルギー最小化問題を解くものである。

研究成果の学術的意義や社会的意義

本研究において開発する技術により、古文書分析が促進されるとともに、書籍・アルバムなどから、物理的な解体を行わずに、ページ単位に情報を切り出すことが可能になるものと期待される。これにより、より多くの古文書の情報が抽出され、文献学をはじめとして、人文科学・社会科学の研究を進化させるきっかけになる。また、多くの冊子体のデジタル化において、冊子の背表紙を取り除き、物理的に解体して、ページスキャナにかけることになっている。本システムの開発により、冊子体に傷をつけることなくデジタル化が実現され、文化財保存に関する災害への耐性強化に貢献する。

研究成果の概要（英文）：In this project, we developed a technique to recognize one page from the volume data of the booklet, even if the pages constituting the booklet are not necessarily flat. Specifically, using the Active Grid method developed by us, it was possible to map the space area corresponding to the booklet from the given 3D CT volume data to a rectangular grid. Active Grid assumes a three-dimensional grid model and deforms each grid so as to fit in a three-dimensional region inside the volume data.

研究分野：視覚的分析

キーワード：データ可視化

1. 研究開始当初の背景

現在、大きな自然災害や戦禍などで被災した重要文献は多く、先進技術を活用した復元に対して、社会から大きな期待がかかっている。歴史文書館では、このような文献の復元の依頼を受け、物理的に文献の各ページを剥がす作業を行って、古文書を取り出してきたが、このような方法では、もとの文献に傷をつける可能性が指摘されてきた。

2011年3月11日の大津波によって海水損した陸前高田市立図書館所管古文書が4月1日～4月3日に回収されて、物理的な復元作業が実施されてきた。また、日本の古文書は、襖の下張りとして使われていたことが多く、歴史文書館では、古い家屋の所有者の方から、古い襖などを預かり、下張りを剥がす作業を行って、古文書を取り出している。貴重な古文書を発見するために、この下張りを剥がすことになり、襖自身については、解体せざるを得ず、非侵襲での復元を期待されている。

西アフリカ・マリで、世界遺産都市トンブクトゥでは、2012年4月にイスラム武装勢力が街を掌握し、世界遺産の古文書や寺院を偶像崇拜だとみなして破壊した。数十万点の古文書が保管されていたアフマド・ババ研究所も2013年1月に放火された。イスラム武装勢力による破壊の被害は逃れたものの、マリ北部の砂漠地帯で乾燥に長年さらされていた羊皮紙の書物は劣化が激しく、このままでは消滅してしまう恐れがあるため、これらのデジタル化が期待されている。

文献に傷をつけることなく、ページごとの文章情報を読み取る技術はこれまでに提案されておらず、本研究が世界に先駆けて実施する初の取り組みとなる。

近年では、文献より断片的に文字情報を抽出する技術報告があった。2015年2月、パピルス古文書について、これに対して、位相コントラスト X 線 CT 装置を用いて、非侵襲にデジタル化し、そのデジタル情報をもとに文書データを抽出する研究成果がイタリアのマイクロエレクトロニクス・マイクロシステム研究所モセラ氏率いる研究チームより発表された。この方法では、古文書に書かれていたギリシャ文字 24 文字を抽出できたと言われるが、単語や文については、抽出できておらず、ページ領域を正確に抽出できていなかったことが原因と推測する。

2016年9月、米マサチューセッツ工科大学 (MIT) メディアラボとジョージア工科大学 (ジョージアテック) の研究者がテラヘルツ波を利用した手法を提案した。現在、検出できるのは、表面に文字が1つずつ書かれた紙を9枚重ねた状態までであり、どのページに何の文字が載っているかを認識するレベルに留まっている。

本研究によって開発が行われる、世界的に類を見ない、吸収型 X 線 CT 撮影装置を用いた冊子体デジタル化技術については、社会実装に大きな期待がかかっている。位相コントラスト型やテラヘルツ波による撮像に比べて、ハード開発のためのコストが小さく、量産への道のりが短い。このため、産業振興の観点で重要な技術開発である。また、位相コントラスト型やテラヘルツ波による撮像であっても、いったんポリリュームデータ化したものについては、本研究開発におけるセグメンテーション技術の適用が可能であり、技術移転の観点でも重要な技術開発である。

2. 研究の目的

本研究では、ページを剥がせなくなった状態の冊子体に対して、三次元データ可視化技術を利用して、これに物理的な変形・破壊を与えない非侵襲なページ情報抽出技術を開発する。この技術では、まず、対象となる冊子体に対して、X 線 CT 装置によりデジタル化し、三次元画像 (ポリリュームデータ) を生成する。次に、このポリリュームデータから等値面可視化技術をつかって、ページ面を等置面として生成する。この目的達成のために、以下の研究的疑問を設定する。

研究的疑問 1 X 線 CT 装置により、冊子体ポリリュームデータが生成できるのか？

冊子体ポリリュームデータは、解像度においても適切さが要求される。すなわち、高解像度すぎるとページ情報ではなく、ページを構成する繊維構造を抽出することになり、ページ上の印字情報を取得する目的からは適切とは言えない。このために、低解像度の撮像に対して、超解像技術を適用する。ただ、線吸収率については、紙とインクとの差をできるだけ高解像度で再構成する必要がある。

研究的疑問 2 冊子体ポリリュームデータから印字情報付きのページ面を等値面として抽出できるのか？

冊子体ポリリュームデータからページ領域に対応したスカラ場を定義できれば、このスカラ場におけるページ情報を等値面として抽出できるが、隣接するページが接触している場合、等値面抽出では、隣のページを対象にしてしまい、迷走する可能性がある。冊子体を構成するページであるという潜在的な意味情報を利用した抽出法を検討する必要がある。

これら提案手法については、ファントム冊子体を使って、ページ情報がどの程度復元できたかを計測して、その有用性を明らかにする。

3. 研究の方法

研究目的で述べた研究的疑問に対して、以下の仮説を設定する。

仮説 1 空間や線吸収率において超解像技術が開発できるなら、X 線 CT 装置により、冊子体のポリウムデータが生成できる。

仮説 2 潜在的な意味情報を使ったセグメンテーション技術を開発できるなら、冊子体ポリウムデータから印字情報付きのページ面を等値面として抽出できる。

仮説 1 の前提「空間や線吸収率において超解像技術が開発できる」に対応して、以下の研究計画を設定する。

研究計画 1-1 超解像技術を使った高解像度化技術の開発

低コストの X 線管や X 線検出器を使ったイメージングでは、十分な解像度のポリウムデータを取得することができない場合がある。この場合、ソフトウェアの活用により、ハードウェアの性能を補償する必要がある。具体的には、冊子体に対して、二種類（高解像度局所撮影、低解像度全体撮影）の投影データを用いて、関心領域における画像再構成を行う。空間分解能とは、二つの物体を識別可能な最小の距離、と定義される。したがって、冊子体を構成するページ情報が再構成画像中で識別可能かどうかによって、空間分解能の評価を行う。空間分解能とは、二つの文字またはページ領域を識別可能な最小の距離と定義する。

研究計画 1-2 吸収率データにおけるコントラスト増幅

吸収型の X 線 CT では、現実問題として、紙とインクの線吸収率に十分な差が出ない場合を想定しないといけない。このような場合、インク領域と紙領域との違いは、ほとんど認識されない状況が発生する。このような場合、線吸収率におけるコントラストを高めるために、X 線の波長を可能な範囲で変化させ、与えられた紙とインクのペアにおいて、コントラストを最大化する波長を特定する。市場で流通する印刷紙・インクにおける線吸収率の波長依存特性を事前に作成しておけば、印刷紙・インクの組み合わせ毎にコントラスト最大化させる波長を特定することができる。この組み合わせは有限なので、コントラスト最大化波長の候補リストを作成し、これを保存しておくようにする。冊子体からポリウムデータを作成するとき、このリストから順に X 線波長を取り出し、コントラストを最大化する波長を選び出すことができるかどうかを確認する。

研究計画 1-3 ページ情報におけるインク発色再構成

ページ上にカラー印刷されていても、X 線 CT では、原則グレースケールで再構成される。一般に、印刷物における色の認識は、使用されているインクの分光カーブと強い関係を持つ。分光カーブとは、反射率の波長特性を表し、通常波長は、可視光のそれを利用する。本研究計画では、可視光の代わりに X 線を用いた吸収率に関する分光カーブを作成する。この作成のために、X 線の波長変化に対する吸収率の変化を調べ、各インクの分光カーブを作成する。対話処理として、カラーページに対する分光カーブ選択を取り入れることにより、適切な色表現を実現することができるかを確認する。

次に、可視光における分光カーブと X 線におけるそれとを関係づけるモデルを作成する。市販されているインクの分光カーブについては、波長に対する反射率データとして保持し、新たに作成する X 線についての分光カーブとの対応関係を作成する。事前にインクの情報が与えられる場合には、この対応関係を使って、色の復元が可能となる。そうでない場合も、対話処理により、対応関係空間の探索を行い、適切な分光カーブを選択し、色の復元が可能になるかどうかを確認する。

仮説 2 前提「潜在的な意味情報を使ったセグメンテーション技術を開発する」に対応して、以下の研究計画を設定する。

研究計画 2-1 ページ領域の空間セグメンテーション技術の開発

本計画では、ページのよれや折れがあっても、一枚の印刷物であると確実に認識できる手法を確立する。紙 + インクと空気領域とが明確に分離するためには、特に、ページの重なり方向における十分な解像度が必要である。ただ、解像度が高すぎると紙を構成するパルプ繊維を抽出することになってしまい、最適な解像度が存在する。

空気領域が徐々に狭くなり、厚さゼロになってしまう接触点を特異点として抽出し、連続するページ面が交差しないという制約、一種の潜在的な意味情報、を課すことにより、迷走しないような等値面抽出手法を開発する。特異点抽出のために、与えられた 3 次元 CT ポリウムデータからヘッセ行列計算を行ってえられるテンソルポリウムデータにおける縮退点を評価する。

何も印刷されていない冊子体から X 線 CT データを取得し、正しくページ領域が抽出されているかどうかを確認する。冊子体については、ページの燃れや折れを人為的に生成しておき、どの程度の頻度の燃れや折れならば、ページ面抽出がうまくいくのかを明らかにする。

研究計画 2-2 文脈情報を使った文字認識技術の開発

本計画では、あらかじめ指定した文書情報に基づき、意味的に整合性のある文字認識を行う技術を開発する。1 ページに印刷された文字情報を抽出するためには、既存の文字認識技術をそのまま利用することができる。

本研究で特徴的なのは、連続する見開きになるページのペアにおいて、正確にページ上で文字認識を行うことである。このページペアにおいて、接触点が存在する場合、どちらか一方のページ上でインク領域を抽出したとしても、もう片方のページにおけるインク領域のページ情報が混入している可能性がある。インク領域において文字認識する場合、文字単位での正確な認識に加えて、生成文における意味的な自然さを実現するための技術を開発し、ペアのページにおいて、正しい文が抽出されているかどうかを確認する。

4. 研究成果

2017年度では、ページを剥がせなくなった状態の冊子体に対して、三次元データ可視化技術を利用したページ情報抽出技術を開発する。H29年度に実施した研究の成果について、冊子体から空間や線吸収率において十分な解像度をもつ三次元データを生成するための準備を行った。

低コストのX線管やX線検出器を使ったイメージングでは、十分な空間解像度のボリュームデータを取得することができない場合がある。この場合、ソフトウェアの利活用により、ハードウェアの性能を補償する必要がある。具体的には、冊子体に対して、二種類(高解像度局所撮影、低解像度全体撮影)の投影データを用いて、関心領域における画像再構成を行った。空間分解能とは、二つの物体を識別可能な最小の距離、と定義される。したがって、冊子体を構成するページ情報が再構成画像中で識別可能かどうかによって、空間分解能の評価を行う。空間分解能とは、二つの文字またはページ領域を識別可能な最小の距離と定義する。

吸収型のX線CTでは、紙とインクの線吸収率に十分な差が出ない場合を想定する。このような場合、インク領域と紙領域との違いは、ほとんど認識されない状況が発生する。このような場合、線吸収率におけるコントラストを高めるために、X線の波長を可能な範囲で変化させ、与えられた紙とインクのペアにおいて、コントラストを最大化する波長を特定する。市場で流通する印刷紙・インクにおける線吸収率の波長依存特性を事前に作成し、印刷紙・インクの組み合わせ毎にコントラスト最大化させる波長を特定するための準備を行った。この組み合わせは有限なので、コントラスト最大化波長の候補リストを作成し、これを保存しておくようにする。冊子体からボリュームデータを作成するとき、このリストから順にX線波長を取り出し、コントラストを最大化する波長を選び出すことができるかどうかを確認するための準備を行った。

2018年度では、冊子体ボリュームデータから、冊子体を構成するページが必ずしも平面でなくても、1枚のページであると認識する手法を開発した。具体的には、研究代表者が開発したActive Grid法を使い、与えられた3次元CTボリュームデータから冊子に対応する空間領域を直方体格子に写像することができた。Active Gridは、三次元格子モデルを前提とし、ボリュームデータ内部の三次元領域に適合するように各格子を変形させる。格子を変形させるために格子の内部エネルギーを対象にしたエネルギー最小化問題を解くものである。

Active Grid法を使った、ボリュームデータからのページ領域の空間セグメンテーション技術の有効性を評価するため、金属元素を含むインクを用意して、印字したページから構成される冊子体を用意し、これを解像度1024×1024×557のボリュームデータを作成した。本研究により、少し撓んだ冊子体を直交体格子に写像し、冊子体を構成する1ページの情報を抽出できることを確認した。

さらに、深層学習技術を使った冊子体を構成するページ文字認識技術を開発した。文字情報を含むページを画像データとして作成し、これを複数積み上げることにより、ボリュームデータを構成した。このボリュームデータを教師データとし、他に元のボリュームデータにアフィン変換(回転、並進、拡大縮小)を適用して得られた、疑似冊子体ボリュームデータを複数作成し、FCNベースの深層学習を実施し、疑似冊子体ボリュームデータから元のボリュームデータが復元できた。

5. 主な発表論文等

〔雑誌論文〕(計 1件)

1. Yuma Hayashida, Tomoya Uetsuji, Yasuo Ebara, Koji Koyamada, "Category Classification of Text Data with Machine Learning Technique for Visualizing Flow of Conversation in Counseling," *Nicograph International (NicoInt)*, 2017, 10.1109/NICOInt.2017.35

〔学会発表〕(計 2件)

1. 中西徹、小山田耕二、三次元CT装置を使った冊子体データ化技術の開発、第46回可視化情報シンポジウム講演論文集
2. 韓忠江、雷溥文、小山田耕二、三次元CT装置を使った冊子体データ化技術の開発、日本シミュレーション学会 AI研究委員会研究会

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

名称：
発明者：
権利者：
種類：
番号：
出願年：
国内外の別：

取得状況(計 1 件)

名称：書物電子化方法および書物電子化装置
発明者：中西徹、小山田耕二、坂本尚久
権利者：シャープ株式会社
種類：特許公報(B2)
番号：6432871
取得年：2018 年
国内外の別：国内

〔その他〕

ホームページ等

6 . 研究組織

(1)研究分担者

研究分担者氏名：

ローマ字氏名：

所属研究機関名：

部局名：

職名：

研究者番号(8桁)：

(2)研究協力者

研究協力者氏名：中西徹

ローマ字氏名：Tohru NAKANISHI

研究協力者氏名：坂本尚久

ローマ字氏名：Naohisa SAKAMOTO

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。