3次元計測を用いた書籍電子化のための 高速かつ高精度なページ分割*

呂 彩 林** 渡 辺 義 浩*** 石 川 正 俊***

High-speed and High-accuracy Page Segmentation for Book Digitization based on 3D Sensing

Chaerim YEO, Yoshihiro WATANABE and Masatoshi ISHIKAWA

In this paper, we propose a new method to segment book shape represented as 3D point cloud into each page as an elemental technologies of book digitization based on 3D sensing. First, we focus on that each page of books can be treated as a two-dimensional surface embedded in three-dimensional Euclidean space. And then, we derive Lipschitz continuity of both normal and curvature to surfaces representing the shape of each page based on the nature of book shapes and the differential geometric properties of surfaces. The conditions for arbitrary two points to exist on the same surfaces are derived by Lipschitz continuity of both normal and curvature to surfaces. The proposal method achieves page segmentation by clustering point cloud using disjoint-set data structure based on this conditions. In the experiment, we evaluated calculation time and accuracy of the proposal method, and showed that our method achieved about 97% success rate with several milliseconds.

Key words: book digitization, 3D sensing, 3D point cloud segmentation, differential geometry, lipschitz continuity

1. 序 論

近年、スマートフォンやタブレットをはじめとする小型情報端末の普及により、電子書籍の需要が増えてきている。さらに、紙に印刷された書籍の電子化による保全も注目されている。このような背景の下、書籍電子化のニーズは、図書館の蔵書や企業の公的文書など大量の書籍を対象とする業務用途から、日常生活における手軽な電子化を対象とする個人用途まで、多岐に渡っている。

あらゆる局面に渡る書籍電子化のニーズに応えるために、書籍電子化システムは、高速、高精度、非破壊かつ手軽なものが要求されるようになり、様々なシステムが開発されている¹⁾。中でも、3次元計測を用いたシステムは、取得された形状情報を用いることで、書籍画像の歪み補正を高精度に行える点で有望である。これにより、ページが変形した状態でも書籍を電子化することが可能となり、書籍の裁断や紙面の圧着を行う必要もなく、高速に電子化が実施可能である。

このようなシステムの確立に向けて、書籍の3次元形状を各ページに分割することは、重要な課題の一つである。特に、ページ分割は書籍画像を取り込んだあとに実施されるだけでなく、取り込む瞬間をリアルタイムに認識するためにも重要な役割を担っている。この場合、ページがめくられている途中の状態でもページ分割を適用できる必要がある。さらに、高速な書籍電子化に向けて、ページ変形中にその動作を止めることなく、連続的に書籍を取り込むスタイルも新たに注目されている²⁾。このように、ページ分割は、多様なページの変形に対して、高精度かつ高速に適用できる性能が求められている。

しかし、このような要求に対して、従来の書籍電子化システム^{3)~5)}におけるページ分割手法は、その対象を、完全に開かれた状態の書籍や、単純なページ変形が生じた書籍に制限し、計測装置に対する書籍位置が既知であると仮定した上で、実現されている。このため、これらの手法は、書籍を自由に配置したい場合や、ページに複雑な変形が生じた場合には適用することが難しかった。これらの制限は、書籍電子化システムに求められる手軽さと高速性を低下させる要因となっている。

このような背景のもと、本論文では新たなページ分割手法を提案する。特に、提案手法では、計測装置に対する書籍の位置が任意であること、手でページをめくることによって複雑なページ変形が生じた場合でも高精度にページ分割ができること、高速な書籍電子化にも適用できるように短時間で処理が完了することを目標としている。

本論文ではこのために、ページの形状を表す曲面における法線と平均曲率のリプシッツ連続性を導出する。このリプシッツ連続性から、同じページ曲面上に存在する任意の2点に関する幾何学的な制約を導くことができる。この制約を用いることで、書籍のページ分割は、3次元点群に対する高速なクラスタリングの問題として記述することができる。具体的には、疎集合データ構造を用いた3次元点群の凝集的クラスタリングによって、高速かつ高精度なページ分割が可能である。提案する手法を評価した結果、ページ分割を97%程度の成功率で達成できることを示すと同時に、数msで処理を実行できる高速性も確かめることができた。

2. 関 連 研 究

2.1 書籍電子化におけるページ分割

既存の書籍電子化手法^{3,4)}は、書籍の位置を固定し、ページ変形による形状も単純なものに制限することで、ページ分割を簡易に実現している。また、変形したページを電子化した場合は、その書籍画像を歪み補正する必要がある。このような歪み

* 原稿受付 平成 29 年 5 月 10 日

掲載決定 平成 29 年 8 月 31 日

** 東京大学大学院情報理工学系研究科

(現、(株)サイバーエージェント；東京都渋谷区道玄坂 1-12-1)

*** 東京大学大学院情報理工学系研究科 (東京都文京区本郷 7-3-1)

補正では、単一のページを対象とすることが一般的である。このため、複数のページを同時に捉えたデータに対して、ページ分割を先に適用することが必要不可欠である。

一方、上記のシステムが書籍を静止させて電子化をするものであったのに対して、ページめくり動作中の書籍に対して電子化をするシステムも提案されている。例えば、BFS-Auto と呼ばれるシステム⁵⁾は、自動めくり機⁶⁾を用いて、ページめくり中の書籍を電子化している。しかし、同システムのページ分割手法は、自動めくり機によるページ変形が円柱形状になることや、書籍の位置が計測装置に対して固定されていることを前提としているため、自動めくりを必要としない書籍電子化のニーズに柔軟に応えることが困難である。

このほか、書籍画像のみを用いてページ分割を行う手法⁷⁾も報告されている。この手法は、開かれた書籍を撮像した画像から、書籍のエッジ検出に基づいて抽出された書籍の枠線を用いて、ページ分割を実現しているため、先述した一般的な環境での適用は困難である。

2.2 ポイントサンプルジオメトリ

コンピュータグラフィックスにおいて、3次元形状の表現形式は、主にポリゴンや曲面といった面を基本要素とするものが典型的であったが、ポイントサンプルジオメトリ (point-sampled geometry) と呼ばれる新たな形式が提案され^{8)~10)}、その重要性が高まっている。ポイントサンプルジオメトリにおける3次元形状は、物体の表面をサンプリングした点群で表され、点間の接続情報は保持しない。このため、データ保持の単純化や、メモリの削減が可能となる。

3次元計測において、ポイントサンプルジオメトリは重要な役割を果たす。一般的に、実物体の3次元計測によって得られる形状は、計測対象の表面をサンプリングした点群であるため、3次元点群の特徴抽出には、ポイントサンプルジオメトリが有効であると考えられる。

一般的に、形状の特徴として、法線や曲率が用いられる。3次元点群における法線推定は、近傍点群の共分散分析により推定する手法^{11)~13)}が典型的であるが、MLS 曲面¹⁴⁾や、Point-Set 曲面¹⁵⁾などの曲面モデルを用いた当てはめにより推定する手法も提案されている。曲面当てはめによる推定は精度が高いものの、処理に多くの時間を要する問題がある。

曲率推定においては、曲面当てはめによる推定方法が一般的であったが、曲面変化度 (surface variation)¹⁶⁾ という、曲面の平均曲率と強い相関を持つ変量が Pauly らにより提案された。曲面変化度は、0 以上 1/3 以下の実数値で、下限は近傍点群が平面を構成している場合、上限は近傍点群が等方的に分布している場合である。曲面変化度は、近傍点群の共分散分析により計算ができるため、高速な計算が可能である。

3. 書籍形状の幾何学的性質

3.1 書籍形状の定式化

書籍は複数のページで構成されており、それぞれのページの変形により、書籍全体の形状が定まる。各ページの3次元形状は、その厚さが無視できるほど薄く、折り目や切れ目が存在しないと仮定することで、 $\mathbb{R}^3$ に埋め込まれた2次元曲面により表現することができる。以降、ページの形状を表す曲面をページ曲面と呼ぶ。幾何学における書籍形状は、複数のページ曲面の集まりとして考えることができる。

しかし、幾何学の世界と違い、実世界における3次元計測で

Fig. 1 Concept of book shape formalization

は、一般的に対象物体の表面の形状情報しか得られないため、すべてのページの形状を得ることは不可能である。例えば、計測装置が上方にある場合は最も上のページの形状のみ計測される。本論文では、ページの表面や裏面が計測可能な、すなわち、形状情報を取得することが可能なページのみを考える。

以上のことを踏まえて、ページ曲面および書籍形状を定式化する。各ページ曲面が $S_i (i = 1, \dots, M)$ で与えられた時、書籍形状 $\mathcal{B}$ は、

$$\mathcal{B} = \{S_1, \dots, S_M\} \quad (1)$$

で表現できる。 M は計測可能なページの数である。図1に書籍形状の定式化の概念を示す。

3.2 微分幾何に基づく曲面の性質

ページ曲面の性質を理解するためには、その本質である $\mathbb{R}^3$ に埋め込まれた2次元曲面の性質を理解する必要がある。以降、特に断らない限り、本論文における曲面とは、 $\mathbb{R}^3$ に埋め込まれた2次元曲面を意味する。本節では、ページ曲面の性質を導くために、曲面の微分幾何学性質について整理する^{17, 18, 19, 20, 21)}。

まず、曲面を定義することから始める。 $\mathbb{R}^3$ 上の動点 $p = (x, y, z)$ の各座標を、 $\mathbb{R}^2$ 上の領域 $D \subset \mathbb{R}^2$ で定義される関数 $x(u, v), y(u, v), z(u, v)$ を用いて、

$$x = x(u, v), \quad y = y(u, v), \quad z = z(u, v) \quad (2)$$

と表したとき、点 p により定義される領域 S を曲面とする。すなわち、 D から $\mathbb{R}^3$ への連続写像 φ

$$\varphi: D \rightarrow \mathbb{R}^3, \quad (u, v) \rightarrow (x(u, v), y(u, v), z(u, v)) \quad (3)$$

による D の像 $\varphi(D)$ が曲面 S を表す。この時、定義域 D と写像 φ を曲面 S のパラメータ表示、 D をパラメータ領域と呼ぶ。

特に、 φ が D に関して C^3 級であり、そのヤコビ行列

$$J = \begin{bmatrix} \frac{\partial p}{\partial u} \\ \frac{\partial p}{\partial v} \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \end{bmatrix} \quad (4)$$

が D の至るところで非退化であるとき、そのパラメータ表示は正則であるといい、正則なパラメータ表示をもつ曲面を正則曲面とする。

ここで、正則曲面 S を考える。便宜上、ここからは S 上の点を、 $S(u, v) = (x(u, v), y(u, v), z(u, v))$ で表す。 v を $v = v_0$ に固定し、 $S(u, v_0)$ は u に関する関数と考え、これは曲面 S 上の曲線である。よって、点 $S(u_0, v_0)$ における u 方向の速度ベクトルは、

$$S_u(u_0, v_0), \quad \text{ただし、} \quad S_u = \frac{\partial S}{\partial u} = \left(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u} \right) \quad (5)$$

で与えられる。同様に、点 $S(u_0, v_0)$ における v 方向の速度ベクトルは、

$$S_v(u_0, v_0), \quad \text{ただし, } S_v = \frac{\partial S}{\partial v} = \left(\frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v} \right) \quad (6)$$

で与えられる。正則曲面の定義から、 S_u と S_v は、曲面上のすべての点において一次独立である。これらは曲面上の任意の点 $S(u_0, v_0)$ における曲面 $S(u, v)$ の接平面の基底であるため、その接平面に垂直な単位法ベクトルは、

$$\tilde{n} = \frac{S_u \times S_v}{|S_u \times S_v|} \quad (7)$$

で与えられる。

曲面には、伸縮具合などの内在的性質を表す第一基本形式、そして空間への入り方などの外在的性質を表す第二基本形式という概念がある。曲面 $S(u, v)$ の第一基本形式 I は、Eq. (8) で与えられ、 E, F, G を第一基本量という。

$$I = Edu + 2Fdudv + Gdv^2 \quad (8)$$

$$E = S_u \cdot S_u \quad (9)$$

$$F = S_u \cdot S_v \quad (10)$$

$$G = S_v \cdot S_v \quad (11)$$

また、曲面 $S(u, v)$ の第二基本形式 II は、Eq. (12) で与えられ、 L, M, N を第二基本量という。

$$II = Ldu^2 + 2Mdudv + Ndv^2 \quad (12)$$

$$L = -S_u \cdot n_u \quad (13)$$

$$M = -S_u \cdot n_v = -S_v \cdot n_u \quad (14)$$

$$N = -S_v \cdot n_v \quad (15)$$

この時、平均曲率 H は、Eq. (16) のように第一基本量および第二基本量により定まる。

$$H = \frac{EN + GL - 2FM}{2(EG - F^2)} \quad (16)$$

3.3 ページ曲面の性質

本節では、前節で踏まえた一般的な曲面の性質に基づき、ページ曲面における法線と平均曲率がリプシッツ連続であることを導出する。以降、平均曲率を記号 H ではなく、記号 κ を用いて表す。

実世界におけるページの形状に自己交差が生じることは考えられないため、ページ曲面は、自己交差しない正則曲面である。ここで、ページ曲面 S のパラメータ表示

$$\varphi: D \rightarrow \mathbb{R}^3 \quad (17)$$

を考える。前節で、 φ による D の像を曲面 S としたため、写像 $\varphi: D \rightarrow S$ は全射である。また、 S には自己交差が生じないため、 $\varphi: D \rightarrow S$ は単射である。したがって、 $\varphi: D \rightarrow S$ は全単射であり、逆写像 $\varphi^{-1}: S \rightarrow D$ が存在する。

逆関数定理によると、 D 上のある点におけるヤコビ行列が非退化であるとき、その点の近傍で φ は、微分同型写像となる。このため、各近傍における逆写像は、 φ と同じく C^3 級であり、その近傍において全域的な逆写像 φ^{-1} と一致する。したがって、 $\varphi: D \rightarrow S$ は微分同型写像である²⁰⁾。

次に、 S の各点 $S(u, v) = (x(u, v), y(u, v), z(u, v))$ における法線 $\tilde{n}(u, v)$ と平均曲率 $\tilde{\kappa}(u, v)$ を考える。 $\tilde{n}, \tilde{\kappa}$ の写像は、以下の

ように書ける。

$$\tilde{n}: D \rightarrow \mathbb{R}^3, \quad \tilde{\kappa}: D \rightarrow \mathbb{R} \quad (18)$$

$\tilde{n}$ と $\tilde{\kappa}$ がリプシッツ連続であることを示すためには、定義域 D を知る必要がある。しかし、3次元計測により得られるものは、写像 φ の値域 S 、すなわち $\varphi(D)$ である。写像 φ の定義域 D は、 S を特定の曲面モデルへ当てはめられない限り、知ることができない。このため、 $\tilde{n}$ と $\tilde{\kappa}$ を、既知である S を定義域とする写像 n と κ に変換し、それらの写像に対して、リプシッツ連続であることを示す。

$\tilde{n}$ と $\tilde{\kappa}$ を、それぞれ φ^{-1} と合成した写像

$$n = \tilde{n} \circ \varphi^{-1}, \quad n: S \rightarrow \mathbb{R}^3 \quad (19)$$

$$\kappa = \tilde{\kappa} \circ \varphi^{-1}, \quad \kappa: S \rightarrow \mathbb{R} \quad (20)$$

を考える。 $\tilde{n}$ と $\tilde{\kappa}$ は、それぞれ φ の1階偏導関数と2階偏導関数から計算される。 φ が C^3 級であるため、 $\tilde{n}$ と $\tilde{\kappa}$ は、それぞれ C^2 級と C^1 級である。したがって、 n と κ は偏微分可能であり、それぞれの偏導関数は連続である。 n と κ の偏導関数が、有界であることを示すことで、リプシッツ連続性が導出できる。

極値定理によると、コンパクト集合を定義域とする連続写像は有界である。これによって、 S がコンパクト集合であると、 n と κ の偏導関数に有界となる。 S はページの形状であるため、有界な閉集合と考えることができる。ハイネ・ボレルの被覆定理によると、 $\mathbb{R}^n$ の部分集合において、有界閉集合とコンパクト集合は同値である。よって、 S はコンパクト集合であり、 n と κ の偏導関数は、コンパクト集合 S を定義域とするため、有界である。

以上から、 n と κ はリプシッツ連続であることが示され、 S 上の2つの点 $p = S(u_p, v_p) = (x_p, y_p, z_p)$ 、 $q = S(u_q, v_q) = (x_q, y_q, z_q)$ に対して、

$$\|n(p) - n(q)\| \leq L_n \|p - q\| \quad (21)$$

$$|\kappa(p) - \kappa(q)| \leq L_\kappa \|p - q\| \quad (22)$$

を満たす実数 L_n と L_κ が存在する。Eq. (21) と Eq. (22) をリプシッツ条件、 L_n と L_κ をリプシッツ定数と呼ぶ。

3.4 ページ曲面上に存在する点の性質

本節では、ページ曲面上の点同士が満たす性質を整理した上で、任意の2点が同一ページ曲面上に存在するための条件を定義する。

ページ曲面 S 上の点 p の近傍領域 $B(p)$ を、 p を中心とした半径 ε_p の閉球体内に存在する S 上の点と定義する。すなわち、 $B(p)$ は、

$$B(p) = \{x \in S \mid \|p - x\| \leq \varepsilon_x\} \quad (23)$$

と書ける。ここで、 $B(p)$ を S における局所領域とする。

$B(p)$ の任意の点 q を考える。 q は、 p の近傍領域内の点であるため、Eq. (24) を満たす。 p と q は、 S 上の点であるため、それぞれの点における法線 $n(p), n(q)$ は、適切なリプシッツ定数 ε_n に対し、リプシッツ条件の Eq. (25) を満たす。同様に、 p と q における平均曲率 $\kappa(p), \kappa(q)$ は、適切なリプシッツ定数 ε_κ に対し、リプシッツ条件の Eq. (26) を満たす。Eq. (24)、Eq. (25)、Eq. (26) をそれぞれ距離制約、法線制約、曲率制約、 $\varepsilon_p, \varepsilon_n, \varepsilon_\kappa$ を、それぞれ距離制約パラメータ、法線制約パラ

Fig. 2 p and q lie on same page surface, but $C(p, q)$ is false. Note that dashed circle is a local area whose radius is ε_p

メータ, 曲率制約パラメータと呼ぶ.

$$\|p - q\| \leq \varepsilon_p, \quad (24)$$

$$\|n(p) - n(q)\| \leq \varepsilon_n \|p - q\| \quad (25)$$

$$|\kappa(p) - \kappa(q)| \leq \varepsilon_\kappa \|p - q\| \quad (26)$$

$\mathbb{R}^3$ 上の任意の点 p, q に対して, 先述した 3 つの制約を同時に満たすか否かを示す真理関数 C を Eq. (27) に定義する. 真理関数 C は, 与えられた任意の 2 点が同一曲面上の同一局所領域内に存在するか否かを示すものである.

$$C(p, q) = \begin{cases} T & (3 \text{ つの制約を同時に満たす場合}) \\ F & (\text{otherwise}) \end{cases} \quad (27)$$

適切な定数 $\varepsilon_p, \varepsilon_n, \varepsilon_\kappa$ に対して, $C(p, q)$ が真となる場合は, p, q は, 同一ページ曲面上の同一局所領域内に存在すると考えることができる. 一方, $C(p, q)$ が偽となる場合は, p, q が異なるページ曲面に存在する場合と, 同一ページ曲面上の異なる局所領域内に存在する場合に大別できる. 後者の場合においては, S が連続曲面であることから,

$$C(p, r_1) \wedge C(r_1, r_2) \wedge \cdots \wedge C(r_{n-1}, r_n) \wedge C(r_n, q) \quad (28)$$

を満たす点列 $\{r_i\}_{i=1}^n$ ($n \geq 1$) が存在する. その例を, 図 2 に示す.

以上から, 任意の点 p, q が同一ページ曲面上に存在する条件は, Eq. (29) で与えられる.

$$C(p, q) \vee$$

$$(\exists n \geq 1, \{r_i\}_{i=1}^n : C(p, r_1) \wedge \bigwedge_{i=1}^{n-1} C(r_i, r_{i+1}) \wedge C(r_n, q))$$

4. ページ曲面の性質に基づいたページ分割

4.1 提案手法の概要

提案手法は, 入力として構造化光法などの 3 次元計測手法により得られる 3 次元点群 $P = \{p_i \in \mathbb{R}^3\}_{i=1}^N$ を持ち, 各点における法線と曲率を推定し, それらを用いてクラスタリングを行うことで, P を各ページを構成する点群 P_i ($i = 1, \dots, M$) に分割するものである. ここで, M は, P に存在するページ数であり, 予め分かっている.

計測において書籍以外の物体が計測される可能性もあり得るため, 入力点群 P は, 必ずしも書籍形状のみが含まれていると限らない. 例えば, 手でページをめくるときの図 3 のように, 手の一部も共に計測されてしまう. このため, 計測の際に書籍以外の物体の存在を許すが, そのサイズは書籍のサイズより十分小さいと仮定する. P は, 幾何学における書籍形状 B を, 有限点群にサンプリングしたものに, ノイズ点群が足されたと考えることができる. なお, ノイズ点群は, 3.4 節で導出

Fig. 3 Example of P including the part of hand

Fig. 4 Process flow of proposal method

した制約を満たさない. このように, 同制約は, ページを分割するとともに, 無関係の形状を分割することにも寄与すると考えられる.

P には, 書籍形状以外の点群が含まれていることから, 必ずしも $\bigcup_{i=1}^M P_i = P$ となるとは限らない. ここで, 書籍形状以外の全ての点群を P_0 とすることで, $\bigcup_{i=0}^M P_i = P$ となり, P_i ($i = 0, \dots, M$) を P の分割として扱うことができる. すなわち, 任意の $i, j \in \{0, \dots, M\}$ に関して, $i \neq j \iff P_i \cap P_j = \emptyset$ となるため, 素集合データ構造での保持および統合が可能となる. 提案手法における処理は, 図 4 に示すように, P の各点における特徴量の抽出, すなわち法線と曲率の推定をした後, 抽出された特徴量を用いてクラスタリングを行う.

4.2 書籍形状の特徴抽出

4.2.1 法線推定

本節では, P の各点 $p_j = (x_j, y_j, z_j) \in \mathbb{R}^3$ における法線 $n_j = (a_j, b_j, c_j) \in \mathbb{R}^3$ および, 曲率 $\kappa_j \in \mathbb{R}$ の推定について説明する.

提案手法では, Hoppe らが提案した手法¹¹⁾を用いて, 法線を推定する. Hoppe らの手法は, 近傍点群における共分散分析を用いるものである. p_j の k -近傍点群 $\tilde{Q}_k(p_j)$ に p_j を加えた点群 $Q_k(p_j)$ は,

$$Q_k(p_j) = \{p_j\} \cup \tilde{Q}_k(p_j) \quad (29)$$

と書ける. ここで, $\tilde{Q}_k(p_j)$ は, k -d 木²²⁾を用いて求める.

Fig. 5 Normal estimation based on covariance analysis. Note that blue points and red arrows are measured 3D points and normal vectors, respectively

$Q_k(p_j)$ の重心と共分散行列は、それぞれ

$$\bar{q}_j = \frac{1}{k} \sum_{q \in Q_k(p_j)} q \quad (30)$$

$$\Sigma_j = \frac{1}{k} \sum_{q \in Q_k(p_j)} (q - \bar{q}_j)^T (q - \bar{q}_j) \quad (31)$$

と書ける。 Σ_j が 3×3 の半正定値対称行列であるため、 Σ_j の固有値と固有ベクトルは、特異値分解により求めることができる。 Σ_j の3つの固有値を、 $\lambda_j^1 \geq \lambda_j^2 \geq \lambda_j^3 \geq 0$ とし、それらに対応する固有ベクトルを v_j^1, v_j^2, v_j^3 とすると、 p_j における法線 n_j は、 Σ_j の最小固有値 λ_j^3 に対する固有ベクトル v_j^3 となる。しかし、図5(a)に示すように、法線の向きは接平面に対して2通り存在するため、 n_j は、 $+v_j^3, -v_j^3$ のどちらも取りうる。このため、共分散分析により求められた法線の向きは、図5(b)に示すように一貫性が保たれていない。

Hoppe らの手法では、点群におけるユークリッド最小全域木に基づき、法線の向きを定めているが、提案手法では、3次元計測で用いられるカメラの位置 o を用いて、簡単に法線の向きを定める。 p_j における接平面によって、空間は2つに分けられる。Eq. (32) のように、 o と同一空間に存在するものを選ぶことで、法線の向きに一貫性を保たせることができる。Eq. (32) による法線の向きの判定結果を、図5(c)に示す。

$$n_j = \begin{cases} v_j^3 & (v_j^3 \cdot (p_j - o) > 0) \\ -v_j^3 & (\text{otherwise}) \end{cases} \quad (32)$$

4.2.2 曲率推定

本節では、 P の各点 $p_j = (x_j, y_j, z_j) \in \mathbb{R}^3$ における曲率 $\kappa_j \in \mathbb{R}$ の推定について説明する。3.章で、平均曲率を用いて

ページ曲面の性質を述べた。3次元点群の各点における平均曲率は、その近傍点群を2次曲面や、MLS 曲面¹⁴⁾などの曲面モデルに当てはめて計算するのが典型的である。しかし、曲面当てはめは一般的に計算量が高いため、高速な推定が困難である。

提案手法では、平均曲率の代わりに、それに強い相関を持つ曲率変化度¹⁶⁾を曲率の変量として用いることで高速な処理を実現する。曲率変化度は、曲面当てはめを行う必要なく、得られた3次元点群のみで計算が可能であるため、3次元点群を入力とする提案手法においても、有効に働くと考えられる。曲率変化度の計算は、4.2.1節で述べた法線推定と同様に、近傍点群における共分散分析に基づくため、法線推定と同時に行えることが可能で、高速な処理が実現できる。

p_j とその k -近傍点群で構成される点群 $Q_k(p_j)$ の共分散行列の固有値を、 $\lambda_j^1 \geq \lambda_j^2 \geq \lambda_j^3 \geq 0$ とした時、 p_j における曲面変化度 κ_j は、Eq. (33) で定義される。

$$\kappa_j = \frac{\lambda_j^3}{\lambda_j^1 + \lambda_j^2 + \lambda_j^3} \quad (33)$$

4.3 点群のクラスタリング

入力点群 $P = \{p_i\}_{i=1}^N$ を、各ページを構成する点群にクラスタリングする問題は、それぞれの p_i が頂点で、 $C(p_i, p_j)$ が真の場合、 p_i, p_j 間に無向辺が存在すると考えることで、無向グラフにおける連結成分を求める問題と同値となる。提案手法では、グラフを構成することを省き、 P を素集合データ構造で保持し、Eq. (27) に基づき統合していくことにより、凝集的なクラスタリングを行う。素集合データ構造とは、互いに素な動的集合の族 $S = \{S_1, \dots, S_k\}$ を管理するデータ構造であり、以下の3つの操作が定義されている。

MakeSet(x) x を唯一の要素と持つ新しい集合を生成し、 S に追加する。

Find(x) 要素 x を含む集合 $S_x \in S$ を返す。

Union(x, y) x を含む集合 $S_x \in S$ と y を含む集合 $S_y \in S$ を S から削除し、 $S_x \cup S_y$ を S に追加する。

素集合データ構造を用いたクラスタリングの骨子は、 $C(p, q)$ が真であれば、 p を含む集合と、 q を含む集合を統合することである。すなわち、 p, q が同一ページ曲面内のある局所領域に存在するならば、それらが属している集合 S_p, S_q の和集合 $S_p \cup S_q$ 内の任意の2点は、Eq. (29) を満たすため、同一ページ曲面上に存在すると考えることができる。

$C(p, q)$ の真偽値を計算するためには、距離制約パラメータ ε_p 、法線制約パラメータ ε_n 、および曲率制約パラメータ ε_κ が必要である。各パラメータの値が大きすぎると異なるページ曲面上の2点が同一ページに、各パラメータの値が小さすぎると同一ページ曲面上の2点が異なるページに判定されてしまうため、書籍のサイズ及び変形度合い、入力点群の解像度を考慮して決める。

すべての統合が終わって、 S に残っているすべての集合 S_i に対して、 S_i 内の任意の2点は、必然的に Eq. (29) を満たす。ここで、 S_i は、各ページを表す点群か、書籍以外の物体の点群か、ノイズにより統合されなかった点群かに大別される。書籍以外の物体は、書籍のサイズより十分小さいと仮定したから、事前に決めたノイズ判定パラメータ ε_T を用いて、 S_i の要素数が ε_T より小さい場合、書籍以外の物体の点群か、ノイズとみ

Algorithm 1 Clustering algorithm with disjoint-set data structure**Require:** $k, \varepsilon_p, \varepsilon_n, \varepsilon_k, \varepsilon_T$

```

1: for all  $p_i \in P$  do
2:    $MakeSet(p_i)$ 
3: for all  $p_i \in P$  do
4:   for all  $q \in \tilde{Q}_k(p_i)$  do
5:     if  $C(p_i, q)$  is true then
6:        $Union(p_i, q)$ 
7:  $M \leftarrow 0, P_0 \leftarrow \emptyset$ 
8: for all  $S_i \in S$  do
9:   if  $|S_i| < \varepsilon_T$  then
10:     $P_0 \leftarrow P_0 \cup S_i$ 
11:   else
12:     $M \leftarrow M + 1, P_M \leftarrow S_i$ 

```

なし、破棄する。これにより、各ページの点群のみ求めることができる。ノイズ判定パラメータの値が大きいと、有効なページがノイズとみなされ、破棄されてしまうため、書籍のサイズ、および入力点群の解像度を考慮して決める。

以上のことを踏まえた全体のアルゴリズムを Algo. 1 に示す。

5. 実験

5.1 実験概要

提案した手法の有効性を確認するために、実際の書籍を3次元計測して得られた形状を用いて、処理時間の評価実験と分割結果の精度評価実験を行った。前者の実験では、分割結果の精度を考慮せず、処理における最悪な場合を想定し、その場合における提案手法の処理時間を評価する。後者の実験では、提案手法により得られたページ分割結果と人間によるページ分割結果を比較することで評価を行う。

5.2 実験環境

3次元計測は様々な方法により実現可能であるが、書籍電子化においては非接触な3次元計測手法が望ましいため、グレイコードパターンを用いた構造化光法を採用した。構造化光法は、プロジェクタなどでパターン光を対象物体に照射し、そのパターンをもとにカメラ画像上の対応点を探索し、三角測量の原理により3次元点を求めることで、計測を行うものである。構造化光法を用いた3次元計測システムを図6に示す。使用した高速カメラとプロジェクタはそれぞれ、Basler社のacA2040-180kcと、Texas Instruments社のDLP LightCrafter 4500である。プロジェクタとカメラの内部パラメータと外部パラメータは、予めZhangの手法²³⁾を用いて定めた。また、同図のプロジェクタの右側のカメラは、キャリブレーションのためのもので、実際の計測には用いられていない。

計測に用いた書籍は、B5判(182 mm × 257 mm)、A5判(148 mm × 210 mm)、B6判(128 mm × 182 mm)の3種類である。B5判とB6判の書籍は、実験のために作成されたものであり、A5判の書籍は小林昭七著の「曲線と曲面の微分幾何」¹⁸⁾である。3種類のサイズの書籍において、手でページをめくったときの6通りの形状を取得した。形状IとIIは、理想的な見開きの状態と、計測装置に対して傾いて設置されている場合を示す。形状IIIからVIは、手めくり動作中に生じる代表的な場合である。形状IIIは、ページをめくる前に、手で書籍を支えている場合である。形状IVは、ページをめくり始めた後で、めくら

Fig. 6 3D shape acquisition system

れているページが綴じ部に対して右側に存在する場合を示す。一方、形状Vは、ページをめくり終える前で、めくられているページが綴じ部に対して左側に存在する場合を示す。最後に、形状VIは、ページを連続的にめくる場合を示す。形状の概要および取得された点群の情報と形状のテキストの一部を表1と図7に示す。

すべての実験において使用した計算機は、Dell Precision T7600で、プロセッサはIntel(R) Xeon(R) Processor E5-2687W 3.10 GHz、メモリは64.0 GB、OSは64ビット版のWindows 10である。

5.3 処理時間の評価実験

提案手法の処理時間に最も影響を与えるのは、入力点群のサイズ N と、近傍点群のサイズ k 、素集合データ構造における $UNION$ 操作の実行回数である。

実験において、B5判書籍の形状IIIの点群を、10通りのサンプリングレートでランダムサンプリングし、得られたそれぞれの点群を入力として用いた。また、近傍点群のサイズ k は8, 16, 32, 64の4通りを用いた。

$UNION$ 操作が最も行われる場合は、入力点群の任意の2点に対して、真理関数 C の値が常に真となる場合である。これは、 $\varepsilon_p, \varepsilon_n, \varepsilon_k$ が無限大に設定することで発生させることができるため、実験における $\varepsilon_p, \varepsilon_n, \varepsilon_k$ の値を無限大に相当する値に設定した。

10通りの入力点群と4通りの k に対し、提案手法の処理時間を100回計測し、平均処理時間を算出した結果を図8に示す。 $k = 8, 16, 32, 64$ の場合において、平均処理時間の処理時間の分布はそれぞれ、0.66 ms から 6.18 ms, 0.74 ms から 7.01 ms, 0.97 ms から 9.26 ms, 1.43 ms から 13.50 msであった。

5.4 分割結果の精度評価実験

また、提案手法によるページ分割の精度を評価する実験を行った。グレイコードパターンを用いた構造化光法による3次元計測では、表1に示したように、非常に密な点群が得られる。

動的に変形している書籍のリアルタイム計測では、高速な3次元計測が必要である。例えば、500~1,000 fpsを達成する高速な3次元計測²⁴⁾で得られる点群数は、数百~数千点程度で

Table 1 Book shape information

Case	With hand	No. of page	Description	No. of acquired point cloud		
				B5-size	A5-size	A6-size
I	No	2	Opened book	905,936	721,788	484,584
II	No	2	Rotated and Opened book	918,825	753,594	499,067
III	Yes	2	Opened book with grasping	928,282	619,086	422,374
IV	Yes	3	Turning page exists on right side of bind	951,975	657,205	426,279
V	Yes	3	Turning page exists on left side of bind	979,884	590,624	382,023
VI	Yes	4	Turning pages exist left and right side both	912,926	629,070	387,783

Fig. 7 Detail information of acquired textures of B5-sized book

Fig. 8 Upper bound of process time

Table 2 The size of down-sampled point cloud of each cases

Case	Size of down-sampled point cloud		
	B5-size	A5-size	B6-size
I	2,370	1,781	1,265
II	2,399	1,853	1,174
III	2,288	1,720	1,345
IV	2,365	1,868	1,409
V	2,353	1,749	1,135
VI	2,418	1,778	1,199

Table 3 Parameter setting

Book size	k	ε_p	ε_n	ε_k	ε_T
B5-size	8	10	0.03	0.005	150
A5-size	8	10	0.03	0.005	150
B6-size	8	10	0.03	0.005	100

ある。提案手法が、リアルタイム計測を伴う書籍電子化にも取り組めることを目標としていることを踏まえて、今回の実験では、18通りの書籍形状の点群のサイズが約1,000~3,000点になるように、ボクセルグリッドフィルターを用いてダウンサンプリングを行った点群を入力点群として用いた。ダウンサンプリング後の点群のサイズ及び実験における提案手法のパラメータ設定をそれぞれ表2と表3に示す。

比較対象としては、人間による分割結果を用いた。入力点群 P に対して、提案手法により分割された点群を $P_i (i = 0, \dots, M)$ とし、人間により分割された点群を $\tilde{P}_i (i = 0, \dots, M)$ とする。

Table 4 Concordance rate of the segmentation result by proposal method and human

Case	Concordance rate %		
	B5-size	A5-size	B6-size
I	99.92	96.17	97.89
II	98.39	95.33	98.17
III	99.79	99.21	97.71
IV	99.46	95.68	97.02
V	94.62	98.95	95.39
VI	98.99	97.64	93.19

P_i と $\tilde{P}_i$ は、同じページを指しているとする。 P の各点 p_i に対して、 $p_i \in P_j \wedge p_i \in \tilde{P}_j$ を満たす $j \in [0, \dots, M]$ が存在すれば、 p_i は正しく分類された点とする。全体の点の数 N に対する正しく分類された点の数の割合を一致率と定義し、これを精度評価の指標とした。すべての形状における一致率を表4に示す。同表に示したように、最低でも、93.19%が一致している。

手が入っている形状で、一致率が最も低い4つの形状の結果の詳細を、図9から図12に示す。それぞれの図に対して、(c)における紫色の点は破棄された点を、(d)における赤色の点は、人間による分割結果と一致していない点を示している。

6. 考 察

6.1 提案手法の高速性

リアルタイム3次元計測を用いた従来の電子化システム⁵⁾におけるページ分割の処理時間は、350点程度の入力点群に対して、約2msである。提案手法により処理可能な点群の数は、同処理時間で約3,000点程度であり、従来手法より約9倍の処理性能を達成している。これにより、提案手法はリアルタイム計測を用いたシステムに十分取り込める高速な処理を実現していると考えられる。

6.2 提案手法の高精度性

表4に示されるように、一致率はすべての形状において90%を超えている。特に、従来手法⁵⁾では形状I以外のページ分割は困難であったことを考えると、適用できるページ形状の自由度が向上したと考えられる。また、ページ分割の結果を、電子化タイミングの自動認識や、書籍画像の歪み補正に用

Fig. 9 Summary of case V with B5-size book

Fig. 10 Summary of case II with A5-size book

Fig. 11 Summary of case V with B6-size book

Fig. 12 Summary of case VI with B6-size book

いる場合、今回達成された一致率は十分な精度であると考えられる。

提案手法による分割と人間による分割が一致していない点は、図10、図11、図12に見られるように、主に書籍の綴じ部周辺とページの境界に分布されていた。この理由としては、綴じ部周辺やページの境界では計測ノイズが発生しやすいことが考えられる。これにより、法線と曲率が正しく推定されなく、法線制約と曲率制約が満たされなくなる。しかし、人間による分割結果と一致していない点は、上記でも述べたページ分割結果の適用先である書籍ページを電子化するタイミングの自動認識や画像の歪み補正においては、意味を持たない点であり、書籍電子化全体における精度には影響がないと考えられる。

また、その他の事例として、図9に示したB5判書籍の形状Vのように、ページの遮蔽が生じて、同一ページの点群が既に分割された状態で計測された場合は、提案手法による分割ができなかった。この理由としては、遮蔽により分割された同一ページ上の2つの点群間に距離制約が満たされなかったため、統合されずノイズとしてみなされ破棄されたことが考えられる。しかし、遮蔽が生じたページは、電子化することが不可能であるため、書籍電子化全体における精度には影響がないと考えられる。

以上のことから、提案手法によるページ分割は、書籍電子化を行うにあたって、十分な精度を持つと考えられる。

7. 結 論

本論文では、3次元計測を用いた書籍電子化の要素技術として、3次元点群で表された書籍形状を各ページに分割する手法を提案した。評価実験では、提案手法の処理時間とページ分割の精度を評価し、高速かつ高精度なページ分割が可能であることを確認した。これにより、書籍形状の幾何学的性質から導出された制約条件が有効であることとともに、このようにコンパクトに表現された制約条件のおかげでページ分割の効率的なアルゴリズムを設計可能であることが示された。

今後の課題として、提案手法におけるパラメータの自動推定が考えられる。現状では、パラメータを経験的に設定した。それぞれのパラメータは、3次元計測の手法、ページの変形具合、書籍のサイズに依存するため、あらゆる場合を網羅するパラメータを推定するためには、事前にデータベースを作成して、その情報を用いることで解決できると期待される。また、遮蔽にも頑固なページ分割手法への拡張のため、リアルタイム計測における形状の時系列情報を用いることも考えられる。

参 考 文 献

- 1) 渡辺義浩, 石川正俊: 書籍電子化技術の新展開 - ブックスキャナの技術動向と超高速化への挑戦 -, 現代の図書館, **51**, 4, (2013) 210.
- 2) T. Nakashima, Y. Watanabe, T. Komuro, and M. Ishikawa: Book Flipping Scanning, Adjunct Proceedings of Symposium on User Interface Software and Technology, (2009) 79.
- 3) A. Yamashita, A. Kawarago, T. Kaneko, and K. T. Miura: Shape reconstruction and image restoration for non-flat surfaces of documents with a stereo vision system, Proceedings of the 17th International Conference on Pattern Recognition, **1**, (2004) 482.
- 4) Z. Zhang, C. L. Tan, and L. Fan: Estimation of 3D shape of warped document surface for image restoration, Proceedings of the 17th International Conference on Pattern Recognition, **1**, (2004) 486.
- 5) 野口翔平, 溜井美帆, 山田雅宏, 渡辺義浩, 石川正俊: 自動めくり機を搭載する適応的撮像型の高速書籍電子化システム, 電子情報通信学会論文誌 D, **96**, 10, (2013) 2590.
- 6) 溜井美帆, 山田雅宏, 渡辺義浩, 石川正俊: 高速書籍電子化のための高速書籍自動めくり機, 日本ロボット学会誌, **31**, 7, (2013) 712.
- 7) V. Rodríguez-Orsoria, M. A. Nuño-Maganda, Y. Hernández-Mier, and C. Torres-Huitzil: Embedded image processing system for automatic page segmentation of open book images, Proceedings of the 10th International Symposium on Visual Computing, **2**, (2014) 531.
- 8) M. Levoy and T. Whitted: The use of points as a display primitive, Technical Report TR 85-022, University of North Carolina at Chapel Hill, Department of Computer Science, (1985).
- 9) H. Pfister, M. Zwicker, J. Van Baar, and M. Gross: Surfels: Surface elements as rendering primitives, Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, (2000) 335.
- 10) S. Rusinkiewicz and M. Levoy: Qsplat: A multiresolution point rendering system for large meshes, Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, (2000) 343.
- 11) H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, and W. Stuetzle: Surface reconstruction from unorganized points, Proceedings of the 19th Annual Conference on Computer Graphics and Interactive Techniques, (1992) 71.
- 12) N. Amenta, M. Bern, and M. Kamvyselis: A new voronoi-based surface reconstruction algorithm, Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques, (1998) 415.
- 13) N. J. Mitra and A. Nguyen: Estimating surface normals in noisy point cloud data, Proceedings of the 19th Annual Symposium on Computational Geometry, (2003) 322.
- 14) D. Levin: Mesh-independent surface interpolation, Geometric modeling for scientific visualization, Springer, (2004).
- 15) N. Amenta and Y. J. Kil: Defining point-set surfaces, ACM Transactions on Graphics, **23**, (2004) 264.
- 16) M. Pauly, M. Gross, and L. P. Kobbelt: Efficient simplification of point-sampled surfaces, Proceedings of the conference on Visualization, (2002) 163.
- 17) M. P. Do Carmo: Differential geometry of curves and surfaces, Prentice-Hall, (1976).
- 18) 小林昭七: 曲線と曲面の微分幾何, 裳華房, (1977).
- 19) 砂田利一: 曲面の幾何, 岩波書店, (2004).
- 20) 杉田勝実, 岡本良夫, 関根松夫: 理論物理のための微分幾何学 - 可換幾何学から非可換幾何学へ, 森北出版, (2007).
- 21) 田崎博之: 曲線・曲面の微分幾何, 共立出版, (2015).
- 22) J. L. Bentley: Multidimensional binary search trees used for associative searching, Communications of the ACM, **18**, 9, (1975) 509.
- 23) Z. Zhang: A flexible new technique for camera calibration, IEEE Transactions on pattern analysis and machine intelligence, **22**, 11, (2000) 1330.
- 24) 田畑智志, 野口翔平, 渡辺義浩, 石川正俊: 3視点拘束に基づくセグメントパターン投影型高速3次元計測, 計測自動制御学会論文集, **52**, 3, (2015) 141.