

令和 3 年 6 月 14 日現在

機関番号：13801  
研究種目：基盤研究(C)（一般）  
研究期間：2018～2020  
課題番号：18K11352  
研究課題名（和文）疎な2次元入力に基づく高画質な3次元形状モデリング  
  
研究課題名（英文）3D Shape Modeling from Sparse Multi-view Images  
  
研究代表者  
岡部 誠（Okabe, Makoto）  
  
静岡大学・工学部・准教授  
  
研究者番号：40557211  
交付決定額（研究期間全体）：（直接経費） 3,500,000円

研究成果の概要（和文）：疎な多視点画像（例えば、単一視点や直交2視点の画像）から3次元ボリュームをモデリングするための手法を提案する。入力画像が疎だと、既存手法でモデリングしたボリュームは不自然にぼやけて見える。一方、提案手法は入力画像の見た目情報を転送しつつモデリングすることで、あらゆる方向から見た時のボリュームの見た目を保存する。見た目情報としては入力画像から学習したテクスチャ特徴量を用いる。提案手法を用いることで、疎な多視点画像からでも自然な見た目の3次元ボリュームが得られることを示す。

#### 研究成果の学術的意義や社会的意義

CTによるボリュームのモデリングには、相当数（通常、数百枚程度）のソノグラム画像（X線写真など）が必要となる。ソノグラム画像の枚数を大幅に減らしても高画質なボリュームをモデリングできる技術を開発することで、1)医療でCTを用いる際、撮影時間を短縮し被写体に与える負担を軽減できる、2)自然現象（煙/炎/液体などの流体、樹木などの植物）のモデリングにCTを用いる際、必要なカメラの台数が減らせる、3)ソノグラム画像が少数なら各画像をユーザが直接編集でき、形状デザインのためのインタラクティブなソフトウェアが開発できる、という3つのメリットがある。

研究成果の概要（英文）：We propose a method of three-dimensional (3D) volume modeling from sparse multi-view images (e.g., only a single-view input or a pair of front- and side-view inputs). The volume determined from such sparse inputs using previous methods appears blurry and unnatural with novel views; however, our method preserves the appearance of novel viewing angles by transferring the appearance information from input images to novel viewing angles. For appearance information, we use texture features learned from the input images. We demonstrate our method successfully provides natural-looking volumetric shapes from sparse multi-view input images.

研究分野：コンピュータグラフィックス

キーワード：Computed tomography 機械学習 モデリング レンダリング

### 1. 研究開始当初の背景

Computed tomography (CT)によるポリウムモデリングは、物体の表面形状だけでなく、物体の内部構造を可視化できるため、医療や教育、また、エンターテインメントのためのコンテンツ生成技術として広く利用されている。CTでポリウムをモデリングするには物体を多視点で撮影した画像が必要だが、それらの撮影には特別な装置が必要であるなど、しばしば多大なコストを要する。また、医療分野では多数のX線画像を撮影する場合に、X線が人体に与える悪影響についても議論されている。そこで研究代表者は、既存のCT技術を拡張し、なるべく少ない枚数の画像集合からでもポリウムをモデリングできるような技術を研究している。

画像に基づくポリウムモデリングに関する既存研究について紹介する。Gregsonらは複数台のビデオカメラで撮影した映像を用いて流体现象をポリウムアニメーションとしてモデリングできる、stochastic tomographyを提案した[1]。この手法の目的の1つは必要なカメラの台数を減らすことだが、依然として複数台のカメラが必要であった。研究代表者らは過去の研究で1台のビデオカメラの映像から流体现象をモデリングした[2]。この手法はモデリングできる形状に限界があり、樹木などの構造はモデリングが不可能とされている。近年、単一画像から3次元形状をモデリングする手法が盛んに研究されている。これらの中には形状表現にポリウムを用いるものもあるが、基本的には物体の表面形状を扱っており[3, 4, 5]、樹木のような内部構造のモデリングは扱われていない。

### 2. 研究の目的

少ない枚数の画像集合からポリウムをモデリングすることについて議論する。ここでは、樹木を正面(方位角が $0^\circ$ )と真横(方位角が $90^\circ$ )から撮影した2枚の画像のみが与えられているとする。CT技術の1つである最小二乗法に基づく手法[6]を用いてポリウムをモデリングし、それを様々な視点からレンダリングした画像を図1に示す。入力2枚の画像と同じ方向からレンダリングすると、入力画像と同じ見た目の画像が得られる(図1の上段の $0^\circ$ と $90^\circ$ )。ところが、視点を変えるとレンダリングされる画像はぼやけた画像となり、そこに樹木の構造は見られない(図1の上段の $20^\circ$ と $45^\circ$ と $70^\circ$ )。また、図1の下段に示すのは、上段と同じ方位角で、少し上空からポリウムを見下ろすようにレンダリングした画像だが、ぼやけて見えると同時に、入力2枚の画像が最小二乗法によって投影された痕跡がグリッド線として見える。もし、樹木のポリウムが正しくモデリングされていたなら、視点を変えたとしても、このようにぼやけて見えたり、グリッド線が見えたりするようなことはないはずである。

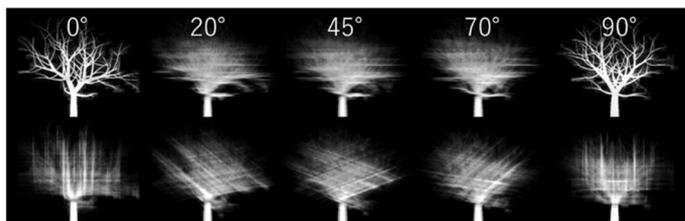


図1) 2枚の樹木の画像からモデリングしたポリウム。

そこで研究代表者のアイデアは、視点を変えてレンダリングしても、入力画像と同様の質感となるようにポリウムを修正しようというものである。具体的には、レンダリングされた画像の質感と入力画像の質感との差をコスト関数とし、それを最小化するようにモデリングを行うことで、どこから見ても入力画像と同様の質感に見えるようなポリウムをモデリングする。結果的に少ない枚数の画像集合からでも良いポリウムがモデリングできる。今回は対象を樹木に限定して実験を行ったので、その結果を報告する。提案手法を用いることで、入力画像が1枚しか与えられない場合でも図2のような樹木のポリウムをモデリングすることができる。

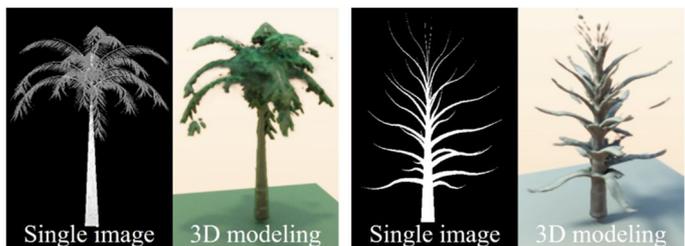


図2) 1枚の画像からモデリングした樹木のポリウム。

### 3. 研究の方法

本研究では画像の質感の記述にテクスチャ画像合成技術を用いる。この技術はユーザの指定した質感を持つ画像を作る技術である。図3に研究代表者が開発した技術による例を示す。上段はユーザが与えた入力画像、下段はコンピュータが合成したテクスチャ画像である。上段と下段の画像は同一ではないが両者の質感は類似している。コンピュータは、入力画像(上段)を分析し、テクスチャ特徴量を抽出し、抽出された特徴量を再現するような画像(下段)を合成する。上記で述べたように、角度を変えて描画しても、入力画像( $0^\circ$ と $90^\circ$ )と同様の質感(テクスチャ



図3)上段が入力画像、下段が合成されたテクスチャ画像。

ャ)が得られるようにポリュームを修正する方法について述べる。テクスチャ画像合成手法の適用結果を図4に示す。図4-aが入力画像、図4-bはポリュームの角度を変えて描画したぼやけた画像である。ここでテクスチャ合成を適用すると、ぼやけた画像(図4-b)を入力画像(図4-a)と同じ質感(同じテクスチャ特徴量を持つよう)に変化させることができる。既存のテクスチャ合成手法[7,8]の結果(図4-cと図4-d)に比べ、提案

手法の結果(図4-e)は、幹から先端に向かうに従い枝が細くなるような特徴など、入力画像(図4-a)の樹木の特徴をより良く捉えている。

この結果を踏まえ、ポリュームの角度を変えて描画した画像の質感と入力画像の質感の差をコストと考え、それを最小化しながらポリュームをモデリングすることで、どの角度から見ても入力画像と同じ質感を持つポリュームが得られる。

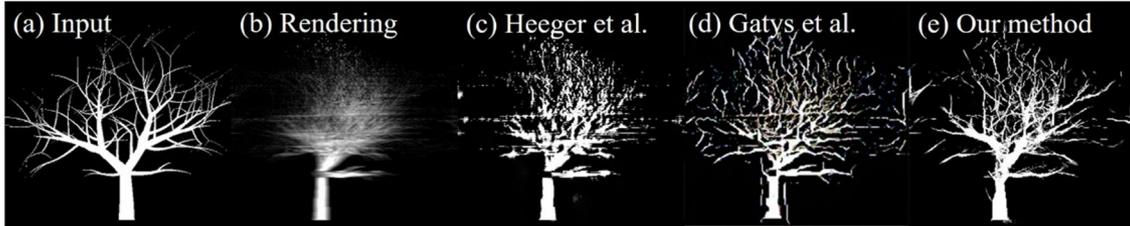


図4)テクスチャ合成手法を樹木のモデリングに適用した際の結果。

#### 4. 研究成果

ここまでの研究で2枚の入力画像のみからでもある程度の樹木をモデリングできることが分かった(より多くの入力画像を準備できるなら、より高精度な形状をモデリングできる)。

実験に用いた入力画像の解像度は全て  $256 \times 256$  ピクセルで、モデリングしたポリュームの解像度は全て  $256 \times 256 \times 256$  ボクセルである。今回用意した入力画像は全て Weber らの手法[10]で作成した樹木のCGモデルをレンダリングすることによって入手した。実際の樹木の写真からのポリュームモデリングは近い将来実施したい。1枚の画像からモデリングした結果を図2に示す。入力画像が1枚しかない場合は、樹木を正面(方位角が  $0^\circ$ )と真横(方位角が  $90^\circ$ )から撮影した2枚の画像が同一であるものとしてモデリングを行う。そのため完成するポリュームは対称性を帯びることになる。図2-右の結果については、モデリングにはトータルで396秒を要した。

樹木を正面(方位角が  $0^\circ$ )と真横(方位角が  $90^\circ$ )から撮影した2枚の入力画像からモデリングした樹木のポリュームを図5に示す。ポリュームのレンダリングには Kroes らの Exposure Render を用いた[9]。この手法はポリュームレンダリング方程式の積分を確率的に効率よく行うことで、インタラクティブに伝達関数や環境照明を変更することが可能になっている。図5-下段では樹木の枝が茶色に、葉が緑色になるようにレンダリングしているが、これは枝のボクセル値と葉のボクセル値の違いを利用し、色を割り当てるような伝達関数を用いてレンダリングしている。

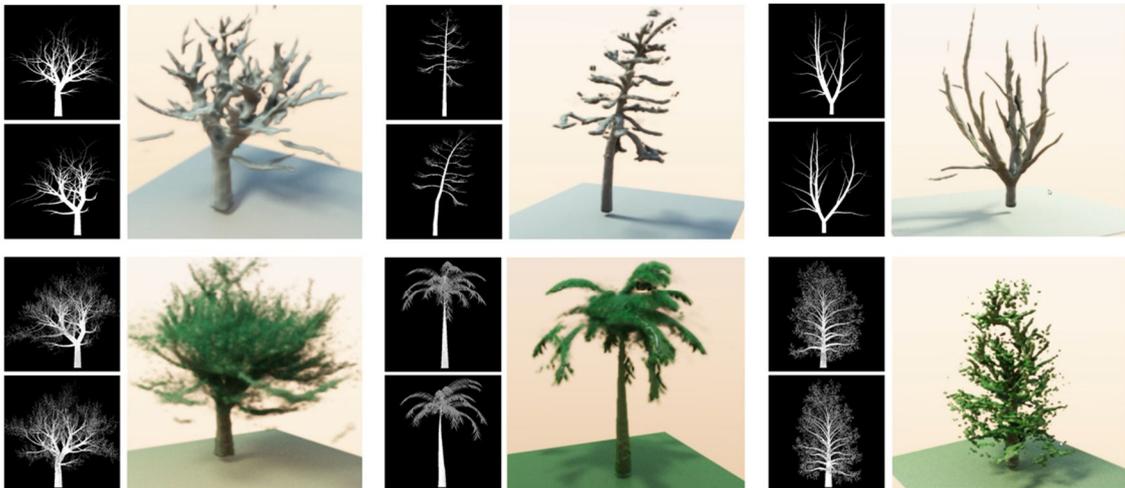


図5)樹木を正面( $0^\circ$ )と真横( $90^\circ$ )から撮影した2枚の入力画像からモデリングした樹木のポリューム。テクスチャ合成手法を樹木のモデリングに適用した際の結果。

提案手法にはいくつか改良の余地がある。そのうちの1つは、現在、Portilla らと同様に相互

相関を制約にしたテクスチャ合成手法を用いているが、低レベルな特徴量しか用いていないために良い画像が合成できない場合がある。例えば、図 5-下段-中央のヤシの木のような例では、木の幹の部分に葉のギザギザな特徴が合成されてしまうことがある。これはテクスチャ合成手法が木の幹と葉の区別をしていないために起こることである。画像中に現れるパーツを分類して理解できるような高レベルな特徴量が扱えるような画像合成手法を開発することが、より高精度なボリュームモデリングするために必須であることが伺える。モデリングに掛かる計算時間が長いことも課題である。現在のボトルネックはテクスチャ合成の処理に時間が掛かっていることと同時に、最適化が収束するまでの繰り返しの多さも問題である。手法を見直して高速化を図り、スケッチベースモデリングなど、インタラクティブなモデリングのためのユーザインタフェースを開発したい。今回はボリュームモデリングについて技術開発を行ったが、同様の考え方はサーフェスマデリングにも適用できると考えている。Structure from motion などの手法と、提案手法の考え方を組み合わせるような実験も行っていきたい。

#### <参考文献>

- (1) J. Gregson, I. Ihrke, N. Thuerey, and W. Heidrich, "From capture to simulation: Connecting forward and inverse problems in fluids," *ACM Trans. Graph.*, vol. 33, no. 4, pp. 139:1-139:11, 2014.
- (2) M. Okabe, Y. Dobashi, K. Anjyo, and R. Onai, "Fluid volume modeling from sparse multi-view images by appearance transfer," *ACM Transactions on Graphics (Proc. SIGGRAPH 2015)*, vol. 34, no. 4, pp. 93:1-93:10, 2015.
- (3) X. Yan, J. Yang, E. Yumer, Y. Guo, and H. Lee, "Perspective transformer nets: Learning single-view 3d object reconstruction without 3d supervision," in *Advances in Neural Information Processing Systems* 29.
- (4) H. Kato, Y. Ushiku, and T. Harada, "Neural 3d mesh renderer," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.
- (5) S. Lutz, Y. Li, A. Fitzgibbon, and N. Kushman, "Inverse graphics gan: Learning to generate 3d shapes from unstructured 2d data," 2020.
- (6) I. Ihrke and M. Magnor, "Image-based tomographic reconstruction of flames," in *Proc. of SCA '04*, 2004, pp. 365-373.
- (7) Heeger et al., "Pyramid-based texture analysis/synthesis," in *Proc. of SIGGRAPH '95*, 1995.
- (8) Gatys et al., "Texture synthesis using convolutional neural networks," in *NIPS 2015*.
- (9) T. Kroes, M. Eisemann, and E. Eisemann, "Visibility sweeps for joint-hierarchical importance sampling of direct lighting for stochastic volume rendering," in *Proceedings of Graphics Interface 2015*.
- (10) J. Weber and J. Penn, "Creation and rendering of realistic trees," in *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques*, 1995, pp. 119-128.

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件/うち国際共著 1件/うちオープンアクセス 0件）

1. 著者名 Makoto Okabe, Keita Noda, Yoshinori Dobashi, Ken Anjyo	4. 巻 40
2. 論文標題 Interactive Video Completion	5. 発行年 2020年
3. 雑誌名 IEEE Computer Graphics and Applications	6. 最初と最後の頁 127-139
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/MCG.2019.2950176	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

〔学会発表〕 計21件（うち招待講演 2件/うち国際学会 2件）

1. 発表者名 岡部 誠
2. 発表標題 疎な多視点画像からの樹木のポリウムモデリング
3. 学会等名 Visual Computing 2020
4. 発表年 2020年

1. 発表者名 坪田 颯生, 岡部 誠, 工藤 隆朗, 由良 俊樹, 本間 祐作
2. 発表標題 SiamMaskを用いた動画修復とマスク生成への応用
3. 学会等名 Visual Computing 2020
4. 発表年 2020年

1. 発表者名 Satsuki Tsubota, Makoto Okabe
2. 発表標題 Interactive Video Completion with SiamMask
3. 学会等名 Pacific Graphics 2020 (国際学会)
4. 発表年 2020年

1. 発表者名 笹俣 匡平, 岡部 誠
2. 発表標題 物体構造情報を用いた画像の超解像度化
3. 学会等名 第178回 CGVI研究会
4. 発表年 2020年

1. 発表者名 岡部 誠, 野田 啓太, 土橋 宜典, 安生 健一
2. 発表標題 Interactive Video Completion
3. 学会等名 Visual Computing 2020 (招待講演)
4. 発表年 2020年

1. 発表者名 渡邊 祐大, 岡部 誠
2. 発表標題 監視カメラ映像の異常検出手法の改良
3. 学会等名 Visual Computing 2020
4. 発表年 2020年

1. 発表者名 三ツ井 慧太郎, 岡部 誠
2. 発表標題 光源位置を考慮した画像の合成ソフトウェアの開発
3. 学会等名 Visual Computing 2020
4. 発表年 2020年

1. 発表者名 池上 敏史, 岡部 誠
2. 発表標題 対象物をイメージ通りの向きや画角でとるためのドローン制御の研究
3. 学会等名 Visual Computing 2020
4. 発表年 2020年

1. 発表者名 山梨 傑, 岡部 誠
2. 発表標題 一枚の画像による3Dモデル再構成
3. 学会等名 Visual Computing 2020
4. 発表年 2020年

1. 発表者名 松本 健吾, 岡部 誠
2. 発表標題 対象を絞った動画の説明文生成
3. 学会等名 Visual Computing 2020
4. 発表年 2020年

1. 発表者名 Makoto Okabe
2. 発表標題 Video analysis and animation synthesis
3. 学会等名 3D Vision & Computer Graphics Online Workshop (招待講演) (国際学会)
4. 発表年 2020年

1. 発表者名 岡部 誠
2. 発表標題 動画中の異常検出技術の実装と分析
3. 学会等名 第175回 CGVI研究会
4. 発表年 2019年

1. 発表者名 毛島 由裕, 岡部 誠
2. 発表標題 ディープラーニングを用いた脳波からの運動意図の識別
3. 学会等名 Visual Computing 2019 ポスター発表
4. 発表年 2019年

1. 発表者名 杉本 達哉, 岡部 誠
2. 発表標題 学習的画像合成による立体誇張の高度化
3. 学会等名 Visual Computing 2019 ポスター発表
4. 発表年 2019年

1. 発表者名 原田 望, 岡部 誠
2. 発表標題 ラフスケッチの線画化技術を用いたリアルタイムの線画修正
3. 学会等名 Visual Computing 2019 ポスター発表
4. 発表年 2019年

1. 発表者名 岡部 誠, 野田 啓太, 土橋 宜典, 安生 健一
2. 発表標題 インタラクティブな動画修復
3. 学会等名 Visual Computing 2019 ポスター発表
4. 発表年 2019年

1. 発表者名 朝倉 康弘, 岡部 誠
2. 発表標題 単眼カメラからリアルタイムで人の姿勢を推定する手法の姿勢矯正への応用
3. 学会等名 Visual Computing 2019 ポスター発表
4. 発表年 2019年

1. 発表者名 齋藤 颯太, 岡部 誠
2. 発表標題 ShadowDrawにおける影の表示精度の向上
3. 学会等名 Visual Computing 2019 ポスター発表
4. 発表年 2019年

1. 発表者名 笹俣 匡平, 岡部 誠
2. 発表標題 画像の高解像度化技術の改善
3. 学会等名 Visual Computing 2019 ポスター発表
4. 発表年 2019年

1. 発表者名 野田 啓太, 岡部 誠
2. 発表標題 時間的一貫性を考慮した動画修復手法の高速化
3. 学会等名 Visual Computing 2018
4. 発表年 2018年

1. 発表者名 杉山 拓海, 朝倉 康弘, 岡部 誠
2. 発表標題 スタイル変換技術を用いて写真をイラスト化するためのワークフロー
3. 学会等名 Visual Computing 2018
4. 発表年 2018年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------