拡張現実のための推定深度に基づくリアルタイム背景除去手法の検討

坪木 良宏[1,a)]　川上 朋也[1]　松本 哲[2]　義久 智樹[2]　寺西 裕一[3,2]　下條 真司[2]

概要：仮想空間技術の一つに拡張現実（Augmented Reality, AR）があり, カメラと携帯端末を用いた視覚的なアプリケーションとして, 撮影された現実の映像にさまざまな情報を重畳表示する仕組みが提案されている. 一方, 視覚的な AR アプリケーションとしては, 撮影端末から距離の近い物体以外を背景として映像から除去し, 任意の背景に置き換えた映像をリアルタイムに表示することなども考えられる. 付近に実在する物体に対して任意の背景を適用することで, エンターテインメント性や臨場感の高い映像体験を利用者に提供できる. 本研究ではスマートフォンやタブレット端末を想定し, カメラで投影されるリアルタイムの映像から背景を除去する手法を検討する. 映像内の除去対象の決定には深度推定と画像処理を用い, あらかじめ設定した閾値よりも高い推定深度の領域に対し画像処理を用い, 検出された領域を背景として除去する.

1. はじめに

仮想現実 (VR)/拡張現実 (AR) 技術の発達により，ユーザが仮想的に異なる位置や時刻に居るかのような体験（時空間移動体験と呼ぶ）を実現することが可能となりつつある．より臨場感の高い時空間移動体験を得るには，ユーザの物理的な移動によって視点が自然に変化する必要がある．ヘッドマウントディスプレイ等を用いて VR 映像をユーザに提供する形態では，ユーザの顔の動きや向きをセンシングして視点を変化させる．しかし，ユーザに見えている映像はユーザが居る現実世界と全く異なるため，ユーザ自身の現実世界での行動は限定的となる．行動範囲を大きくすると，現実世界の周囲に存在する人やモノに衝突するなどの問題が生じてしまう．

ユーザの物理的行動を制限せずに時空間移動体験を得る方法として，室内に VR 映像をプロジェクション投影する設備を用いる方法がある．この方法により部屋全体が異なる時空間に移動したかのような体験が可能となる．しかし，VR 映像投影可能な特別な設備を備えた部屋が必要となる．また，一般にユーザの室内での移動は制限される．特別な設備を備えた部屋が不要な例としては，Atmoph Window [1] がある．Atmoph Window は，室内に設置する窓型のディスプレイ装置であり，ディスプレイに異なる空間を映し出すことで，異なる空間に存在する部屋に居る感覚を得られるようにする．ディスプレイ上の映像は，ユーザの顔位置を検出した上で角度に応じて変化し，視点変化に追従する．しかし，ユーザの体験はこの窓型の装置が置かれた付近に限定される．

本研究では，時空間移動体験の臨場感を維持しつつ，実世界でのユーザの行動を制限せず，自由な視点移動を可能とする方法として，ユーザから一定距離以上の室内空間に AR を用いて時間・場所が異なる空間を投影する「AR 室内旅行」を実現するシステムを提案する．

提案システムでは，屋内の一部を背景とみなし，仮想的に異なる時空間に置換する AR アプリケーションを実現する．提案システムによる AR の実現イメージを図 1 に示す．ユーザがモバイル端末，AR グラスなどのディスプレイを通して室内を見渡すと，実世界の物体の一部が任意の VR 空間に置き換わる．

本稿では，提案する AR アプリケーション実現のために必要となる基礎技術として，屋内映像の一部や一定以上の距離にある物体を背景として認識し，リアルタイムに背景除去を行う方法について述べる．

2. 関連研究

仲野らはまち歩きイベントに AR アプリケーションを用いることで，イベント参加者の知識獲得，満足度の向上を支援した [2]．この AR アプリケーションは，異なる空間

[1] 福井大学
Uniersity of Fukui, Fukui 910–8507, Japan
[2] 大阪大学
Osaka University, Ibaraki, Osaka 567–0047, Japan
[3] 国立研究開発法人情報通信研究機構
National Institute of Information and Communications Technology, Koganei, Tokyo 184–8795, Japan
[a)] hb180738@u-fukui.ac.jp

©2021 Information Processing Society of Japan

図 1 元映像（左）と「AR 室内旅行」の使用イメージ（右）

への移動は考慮しておらず，時間を遡った街の様子を重畳表示するものである．森らは実在する物体を視覚的に消し去ったうえで体験者に除去したことを分からせないようにする隠消現実感（Diminished Reality, DR）システムを提案している [3,4]. 本論文で目指す AR アプリケーションは DR とは大きく異なり，実物体を除去し違和感なく背景を重畳表示するのではなく，実物体に任意の背景を重畳表示し，まったく別の空間のように見せることを目指している．

3. 提案手法

提案する AR アプリケーションを実現するためには，カメラの入力画像から背景とする領域を検出することが必要になる．本研究では，幅広い機器で適用可能とするため，その領域検出を機械学習による深度推定と画像処理を併用する．具体的には，入力画像の深度推定の結果から明らかに背景ではないと判断できる領域を先に抽出し，その領域以外の背景の候補領域に対し画像処理を用いて背景を検出する．提案手法の流れを図 2 のフローチャートに示す．提案手法を用いることで，ユーザが実世界の物体に衝突し，転倒してしまう危険性を排除したうえで，距離的な制限がなく，エンターテイメント性や臨場感の高い AR 空間を構築する．

4. 実装

筆者らは 3 章で述べた提案手法に従い，試験的なリアルタイム背景除去システムを 1 台の Windows 10 PC 上に Python で実装した．実装に用いた PC は CPU が Intel Core i5-7300U（最大 3.50GHz, 2 コア/4 スレッド），メモリが 8.00GB で，GPU はオンボードである．機械学習による深度推定も Python で実行し，機械学習のモデルには Laina らの手法を用いた [5]．サイズが 480×640 の入力画像での深度推定結果を図 3 に示す．画像処理機能も同じく Python により実装しており，入力画像のエッジ抽出やモルフォロジー変換を行っている．また，ユーザの視線角度によって，大まかな背景領域と除去が不要な実世界領域を決定する処理も行っている．画像処理の過程を図 4 に示す．図 4b のエッジ抽出（Canny 法）と図 4c のモルフォロジー変換には，それぞれ OpenCV [6] の cv2.Canny() と cv2.dilate(), cv2.erode() を用いた．

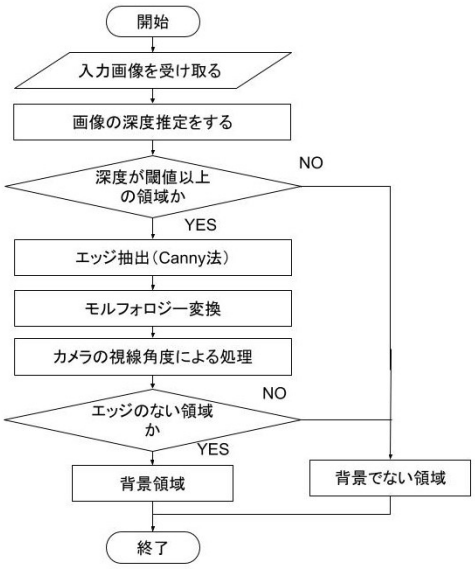

図 2 提案手法の流れ

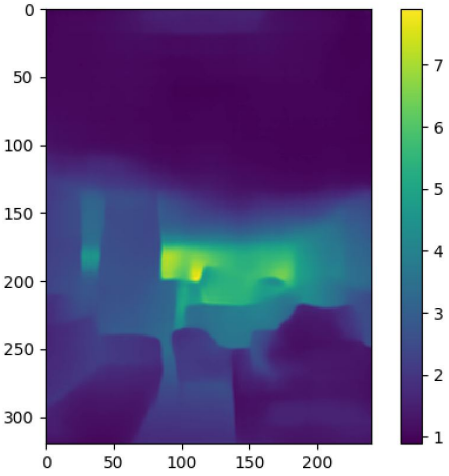

図 3 機械学習による深度推定結果

5. 評価

4 章で述べた試験的な実装システムを用いて，リアルタイムな AR アプリケーションへの適用可能性を評価した．機械学習による深度推定に関しては，画像サイズやプログラムを変更した際の実行時間を測定した．画像処理に関しては，エッジ抽出とモルフォロジー変換，視線角度によって決定される背景除去領域の正確さを測定した．

深度推定の結果については，入力画像サイズが 480×640 の場合，実行時間は約 1.82 s であった．本研究で目指すリアルタイムな AR アプリケーションでは処理遅延はユーザの臨場感の低下につながり，現在の試験的な実装システムによる深度推定の結果では不十分と考えられる．そのた

©2021 Information Processing Society of Japan

図 4 画像処理の過程

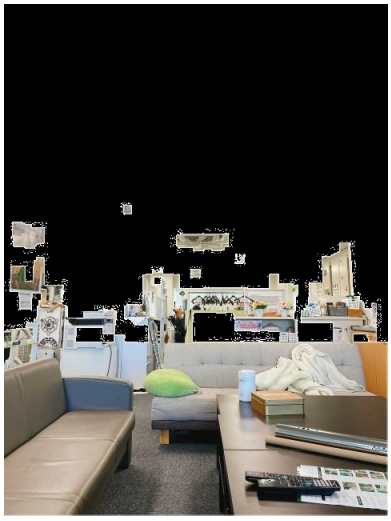

図 5 画像処理により決定された領域

め，高性能な GPU を搭載した PC 上での実装やシステムの分散化を今後検討している．また，画像処理の結果からは，入力画像に Canny 法を用いたエッジ抽出とモルフォロジー変換を行った後，ユーザの持つカメラの視線角度から大まかに背景除去したい領域とそうでない領域を決定できることが視覚的に確認できた．画像処理により決定された領域を図 5 に示す．現時点の実装システムでは，機械学習による深度推定結果と画像処理結果を組み合わせての背景領域の決定は行っていない．2 つの仕組みを組み合わせ，ユーザ周辺の実世界を残しつつ，臨場感の高い AR 空間が実現できる背景抽出除去を今後検討する．

6. まとめ

本論文では，拡張現実のための推定深度に基づくリアルタイム背景除去手法について検討した．検討手法の試験的な実装と評価の結果，機械学習を用いてある程度の精度で深度を推定できることを確認した．また，画像処理の結果からは，エッジ抽出やモルフォロジー変換などを組み合わせることにより，大まかな背景領域を決定できることを確認した．

今後の課題として，機械学習による深度推定と画像処理を組み合わせることで，背景領域の決定や深度推定の実行時間，精度の向上が考えられる．実行時間の観点では，リアルタイムのアプリケーションを実現するため，実装システムの GPU による高性能化や分散化も考えられる．

謝辞 本研究の一部は G-7 奨学財団研究開発助成事業および福井大学研究育成経費，JSPS 科研費 18K11316 の助成による成果である．

参考文献

[1] 姜 京日：本当の夢を探し続けよう，感性工学， Vol. 19, No. 2, p. 87 (2021).
[2] 仲野潤一，大澤荘平，鳴海拓志，谷川智洋，廣瀬通孝：領域型バーチャルタイムマシンを用いたまち歩きイベント，日本バーチャルリアリティ学会論文誌，Vol. 22, No. 2, pp. 241–250 (2017).
[3] 森 尚平，一刈良介，柴田史久，木村朝子，田村秀行：隠消現実感の技術的枠組と諸問題: 現実世界に実在する物体を視覚的に隠蔽・消去・透視する技術について，日本バーチャルリアリティ学会論文誌，Vol. 16, No. 2, pp. 239–250 (2011).
[4] 李 金霞，斉藤純哉，森 尚平，池田 聖，柴田史久，木村朝子，田村秀行：隠消現実感技術を用いた映画制作支援システムの開発と運用，日本バーチャルリアリティ学会論文誌，Vol. 21, No. 3, pp. 451–462 (2016).
[5] Laina, I., Rupprecht, C., Belagiannis, V., Tombari, F. and Navab, N.: Deeper Depth Prediction with Fully Convolutional Residual Networks, *Proceedings of the 4th International Conference on 3D Vision (3DV 2016)*, pp. 239–248 (2016).
[6] OpenCV Team: OpenCV. Available at `https://opencv.org/` (accessed Sept. 19, 2021).

©2021 Information Processing Society of Japan