

360度歩行映像の3D都市モデルへの動的なテクスチャ投影による バーチャル空間の実現とその応用

坂野 達郎[†] 武縄 瑞基[†] Leslie Woehler^{†‡} 池畑 諭^{§¶} 相澤 清晴[†]

[†]東京大学 [‡]日本学術振興会 [§]国立情報学研究所 [¶]東京工業大学

1 はじめに

現実世界のある地域をバーチャル空間として再現することで、アバターを操作して、現地に足を運ぶことなくその地域を歩き回ることや、再現した現実世界とComputer Graphics (CG) を組み合わせて、その地域での拡張現実を体験することが可能となる。従来では、このようなバーチャル空間を実現するために、専門のクリエイターが、現実世界の3Dモデルを緻密に作成する必要があったが、これには多大な労力と時間がかかるため、さまざまな地域への拡張が困難である。この課題を解決するため、本研究では、現実世界を再現したバーチャル空間を実現する新たな手法として、公的機関によって整備された現実の都市の3Dモデルと、街を歩いて撮影した360度映像を組み合わせて利用することを提案する。

現実の都市の3Dモデルに関しては、日本の場合、国土交通省が中心となり、日本全国の都市の3Dモデルを整備するPLATEAUというプロジェクトが進められており、すでに東京23区の全域を含む約10000 km²の地域の3Dモデルが公開されている[1]。このようなモデルを利用することで、バーチャル空間の実現における、幾何モデル作成の手間を削減することができる。一方で、公開されている都市の3Dモデルは、大半の地域において、テクスチャ画像を持たないシンプルな箱型のものとなっているため、そのままでは現実世界の外観情報を再現できない。

360度映像に関しては、小型の360度カメラを利用し、対象となる地域のあらゆる道路について、実際に歩きながら撮影を行うことで、その地域における外観情報のデータベースを構築することができる[2]。この撮影の実施には、専門的な知識を必要としないため、さまざ

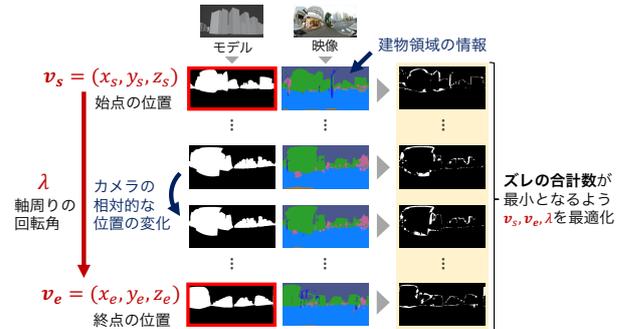


図1 モデルと映像の位置合わせアルゴリズムの概要

まな地域へと容易に拡張できる。一方で、360度映像は現実世界の3次元幾何情報を持たないので、現実世界の外観情報と、バーチャル空間中で配置したCGの3次元的に整合した合成には制限を受ける[3]。

本研究では、3D都市モデルに対して、360度映像を動的にテクスチャ投影することで、3D都市モデルの幾何情報と、360度映像の外観情報の両方を持ち、アバターで自由に歩き回ることのできるバーチャル空間を実現する。そして、このバーチャル空間の応用として、公的機関による洪水想定水位の情報をCGとして描画し、現実世界の外観の中で可視化する拡張現実システムを示す。

2 提案手法

2.1 モデルと映像の位置合わせ

まず、360度映像の各フレームが、3D都市モデル内のどの位置から撮影したかを推定するために、各道路の360度映像について、モデルとの位置合わせを実行する。この際に用いるアルゴリズムの概要を図1に示す。

はじめに、360度映像に対して、Visual SLAM [4] を実行することで、映像内でのカメラの相対的な軌跡を得る。この相対的な軌跡から、各フレームの絶対的な撮影位置を確定するためには、カメラの軌跡の始点の絶対的な位置 $v_s = (x_s, y_s, z_s)$ と終点の絶対的な位置 $v_e = (x_e, y_e, z_e)$ 、および始点と終点を結んだ直線を回転軸とした、軸周りの回転角 λ を定める必要がある。ここでは、映像撮影時に、始点と終点の大まかな位置情報が記録されていると仮定し、この情報から得た v_s, v_e, λ の

Realization of virtual world by dynamic texture projection of 360-degree walking video onto a 3D city model and its application

Tatsuro Banno[†], Mizuki Takenawa[†], Leslie Woehler^{†‡}, Satoshi Ikehata^{§¶}, and Kiyoharu Aizawa[†]

[†]The University of Tokyo

[‡]JSPS

[§]National Institute of Informatics

[¶]Tokyo Institute of Technology



図2 モデルへの映像のテクスチャ投影の概要¹

初期値を、モデルと映像のずれの評価値が最小となるように最適化していくことで、 v_s, v_e, λ の値を決定する。

モデルと映像のずれの評価値を求めるために、まず、360度映像の各フレームに対して、Semantic Segmentation [5] を行い、各フレームでの建物領域の情報を得る。次に、現在評価している v_s, v_e, λ の値によって決定された、各フレームの絶対的な撮影位置において、バーチャル空間内の仮想的な 360 度カメラから、3D 都市モデルのレンダリングを行う。そして、その結果画像における 3D 都市モデルの建物領域と、Semantic Segmentation で得られた建物領域を比較し、一方では建物領域であるが、もう一方では非建物領域であるピクセルの数を全フレームで合計し、これをモデルと映像のずれの評価値とする。この評価値が最小となるよう、 v_s, v_e, λ の値を最適化することで、モデルと映像の位置合わせを行う。

2.2 カメラ位置の決定と動的なテクスチャ投影

前項の位置合わせによって得た、各フレームの撮影位置の情報を用いて、バーチャル空間内でアバターを映すカメラの位置を決定する。また、このカメラの位置から、モデルへの映像のテクスチャ投影を行う。この手法における、テクスチャ投影の概要を図2に示す。

まず、3D 都市モデル内での現在のアバターの位置に基づき、360度映像のフレームの中から、アバターが適切に映るような位置から撮影されたフレームを1枚選択する。そして、バーチャル空間内のカメラをその位置に配置し、360度映像における、対応するフレームの画像を、その位置からモデルに対してテクスチャ投影する。こうすることで、カメラの視点から3D都市モデルを見た際に、現実世界の外観情報がそのまま再現される。また、アバターの位置が変化した際には、それに応じて、カメラが対応する位置へと追従し、また、投影する画像

¹アバターのモデルには Unity-Chan を用いた。
<https://unity-chan.com/contents/guideline/>
 ©Unity Technologies Japan/UCL



図3 バーチャル空間を応用した洪水可視化システムの様子¹

がそのカメラ位置に対応するフレームのものへと動的に変化する。このため、従来の静的なテクスチャマッピングと異なり、アバターの位置が変化しても、常に写実的な外観を持ったバーチャル空間が実現される。

3 応用

このようにして実現したバーチャル空間を応用し、その地域の洪水想定水位を可視化する、拡張現実システムを実現した。このシステムの様子を図3に示す。

このシステムは、公的機関が提供する洪水想定水位の情報を、CGとしてバーチャル空間内で描画する。バーチャル空間の幾何情報により、建物と水面が3次元的に整合した形で合成され、また、バーチャル空間の外観情報により、現実世界の外観が再現されるため、リアリティのある洪水表現が実現される。これにより、利用者の洪水災害への当事者意識の向上が期待される。

4 まとめ

本研究では、現実世界を再現したバーチャル空間を、多大な労力を要することなく実現する新たな手法として、3D都市モデルと360度映像を組み合わせる手法を提案した。また、このバーチャル空間の応用として、洪水可視化を行う拡張現実システムを示した。本研究の手法により、さまざまな地域のバーチャル空間が実現され、新たな応用が展開できる。

謝辞 本研究の一部は、JST JPMJMI21H1, SIP3「スマート防災」、JSPS KAKENHI 21H03460 の支援を受けた。

参考文献

- [1] 国土交通省. Project PLATEAU 全国 56 都市の 3D 都市モデルのオープンデータ化を完了, 2021. https://www.mlit.go.jp/report/press/toshi03_hh_000078.html.
- [2] N. Sugimoto, Y. Ebine, and K. Aizawa. Building movie map-a tool for exploring areas in a city-and its evaluations. In *ACMMM*, 2020.
- [3] M. Takenawa, N. Sugimoto, L. Wöhler, S. Ikehata, and K. Aizawa. 360rvw: Fusing real 360° videos and interactive virtual worlds. In *ACMMM*, 2023.
- [4] S. Sumikura, M. Shibuya, and K. Sakurada. Openslam: A versatile visual slam framework. In *ACMMM*, 2019.
- [5] S. Orhan and Y. Bastanlar. Semantic segmentation of outdoor panoramic images. *Springer SIVP*, Vol. 16, No. 3, pp. 643–650, 2022.