



# モバイル AR における前面カメラ画像の外挿を用いた映り込みの再現

Reproduction of Environment Reflections  
using Extrapolation of Front Camera Images in Mobile AR

小田嶋駿, 小室孝

Shun ODAJIMA, Takashi KOMURO

埼玉大学 理工学研究科 (〒338-8570 埼玉県さいたま市桜区大字下大久保 255, odajima@is.ics.saitama-u.ac.jp)

**概要:** 本研究では, モバイル端末を用いた AR において, 端末の前面カメラで撮影した画像のみを用いて仮想物体への実シーンの映り込みを再現する手法を提案する. 前面カメラで端末手前のシーンを画像として取得し, 映り込みの一部分を再現する. また, pix2pixHD を用いて画像の外挿を行うことでカメラの視野外の領域を含むシーン全体を推定する. 訓練には約 18,000 枚の全方位画像を使用し, 前面カメラでの撮影を想定した  $128 \times 128$  ピクセル画像を  $512 \times 256$  ピクセル画像まで外挿した. この画像を用いて銀素材の仮想物体への環境マッピングを行い, 外挿画像で自然な映り込みが再現されることを確認した.

**キーワード:** 拡張・複合現実, 光学的整合性, 画像の外挿, 環境マッピング

## 1. はじめに

拡張現実感(AR)において現実世界と仮想空間上の光学的整合性を保つ研究が行われている. 撮影画像から得られる影や, 鏡面球を設置して撮影した画像から光源推定を行い, 仮想物体への映り込みを再現する研究が存在する[1][2]. これらの研究では, 事前に画像を撮影する必要があったり, 鏡面球を用いたりといった, 追加の準備や装置が必要になる.

そこで, モバイル端末の前面カメラで撮影した画像のみを用いて, 仮想物体への実シーンの映り込みを再現する方法を提案する. モバイル端末の前面カメラで端末の手前のシーンを画像として取得し, その画像を用いて映り込みの一部を再現する. また, 前面カメラから取得できるシーンはカメラの前方のシーンのみであり, 実シーン全体の映り込みを再現するには不十分であるが, 取得したシーンの周辺部を外挿することにより, 映り込みを再現するのに十分な範囲のシーンを推定する.

## 2. 画像の外挿を用いた映り込みの再現

提案手法ではモバイル AR を用いて実シーンの仮想物体への反射を再現する. その概要を図 1 に示す. 取得したシーンを外挿し, 外挿したシーンを環境マップに変換し, 環境マッピングを用いて物体をレンダリングすることに

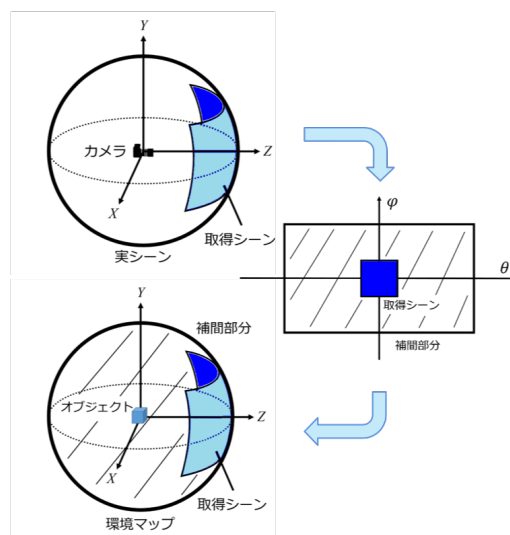


図 1: 提案手法の概要

より, 実シーンの映り込みを再現することができる.

### 2.1 周囲部分の外挿

提案方法では, 前面カメラで取得した画像からシーンの周辺部分を外挿する. 画像の外挿には Wang らが提案した pix2pixHD[3]を用いた. これは Isola らが提案した pix2pix[4]による画像補間の技術をより大きな画像に用いるように改良したものである. pix2pix は Generative

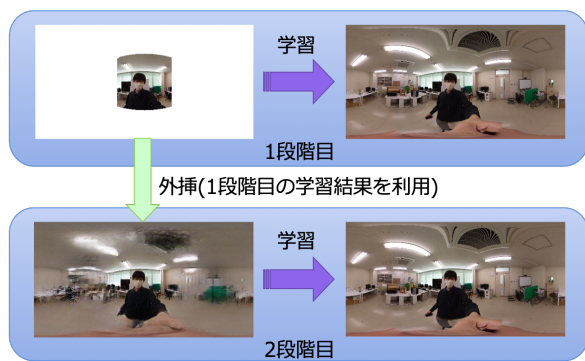


図 2: 2段階学習

Adverse Network (GAN) を用いた手法であり、二つの画像のペアを与えることでそれらの画像間の変換を学習し、学習したデータを用いて画像の補間をすることができるものである。pix2pixを用いてKimuraらは、没入感拡張システムのための周辺視野映像生成を行う手法[5]を提案した。本研究では、全方位画像を用いて学習を行う。全方位画像のうち前面カメラで取得可能な部分を切り取って入力とし、元の全方位画像を正解として学習に用いる。

学習は2回に分けて行う。1段階目と2段階目それぞれの例を図2に示す。まず全方位画像のうち前面カメラで取得可能な範囲から全方位画像全体への学習を行う。これを用いて外挿を行ったものでは入力部分と比較して外挿部分がぼやけてしまう。そこで1段階目で学習したデータを用いて生成された全方位画像から正解となる全方位画像への学習を行う。これにより、入力部分と外挿部分における鮮明さの差を減らすことで、外挿結果の違和感を減らすことができる。

## 2.2 映り込みの再現

外挿した画像を元に環境マッピングを用いて映り込みの再現を行う。本研究では、球を用いた環境マッピングであるスフィアマッピングを使用し、緯度経度の図を環境マップとして与える。そのため、取得した $(x,y)$ 座標系の2次元画像を $(\theta,\varphi)$ 座標系の画像に変換し、環境マップとして利用する。

## 3. 外挿と映り込みの再現結果

今回の実験では、室内で全方位カメラを使用して撮影した約18,000枚の全方位画像を使用した。モバイルARの前面カメラの視野角は垂直方向 $94^\circ$ 、水平方向 $110^\circ$ を、得られる入力画像のサイズは $128 \times 128$ ピクセルを想定し、入力画像を用いて $512 \times 256$ ピクセルまで外挿した。図3にpix2pixHDによる画像の外挿結果の一例を示す。入力画像には、訓練画像同様に全方位カメラで撮影した画像のうち、前面カメラで取得可能な範囲のみを保持した画像を使用した。また、入力画像と訓練画像にはそれぞれ別の部屋で撮影したものを使用した。外挿結果の1段階目では外挿を行った部分がぼやけているのに対し、2段階目ではより鮮明な結果が得られた。Ground truth と比べると外挿部分が異なっているが、外挿の目的は正確な外挿を得ることで



外挿前



外挿後(1段階目)



外挿後(2段階目)



Ground truth

図 3: 外挿結果の例

はなく、自然な映り込みを再現することである。そのため外挿結果は必ずしも正解に近いものではないが、外挿後の画像が自然に見える場合に映り込みの再現に使用しても問題はない。

仮想物体のレンダリングにはUnityを使用し、等方性Wardモデルを使用した[6]。物体は、拡散アルベド、鏡面アルベドおよび粗さのパラメータをそれぞれ $\rho_d=0.1$ 、 $\rho_s=0.33$ および $\alpha=0.16$ に設定した銀材料で作成した。図4に映り込みを再現した結果を示す。外挿画像を用いた場合でも、Ground truthを用いた場合と遜色なくらいに自然に見える映り込みを再現できているように考える。



外挿前画像



外挿後画像



Ground truth

図 4: 映り込みの再現結果

#### 4. モバイル AR による再現結果

実際にモバイル AR を用いて取得した画像による映り込みを再現した。その様子を図 5 に示す。モバイル AR を実現するための装置にはタブレット PC である Microsoft Surface Pro 6 を使用した。前面カメラからフレームごとに画像を取得し、一定時間ごとにカメラ画像を外挿する。環境マップのうち前面カメラ画像を使用する部分はリアルタイムに更新し、外挿画像を使用する部分は一定時間ごとに更新する。画像取得は前節の実験と同様に訓練画像を取得した部屋とは別の部屋で行った。

取得した画像を用いて外挿と映り込みの再現を行った結果を図 6 に示す。入力画像は取得画像を  $128 \times 128$  ピクセルに落としたものを、前節の実験における外挿前画像と同様の形になるように座標系の変換を行ったものを使用した。結果は多少鮮明さに欠ける部分があるが、映り込みに用いるには問題がない程度である。映り込みの再現に用いた仮想物体のパラメータは前節の実験と同様のものを使用し、環境マップは前面カメラの座標系に位置合わせを行った。結果として自然に感じられる程度の映り込みが再現できていると考える。

#### 5. おわりに

本研究では、モバイル AR における前面カメラ画像を用いた外挿を行い、仮想物体への実シーンの映り込みの再現をするための手法を提案した。また、提案手法を用いて実験を行った。

今後の課題として、外挿を行った画像を用いた仮想物体への映り込みが自然なものとして再現されているかどうかを調べるために、被験者を集めて主観評価実験を行うことなどが挙げられる。

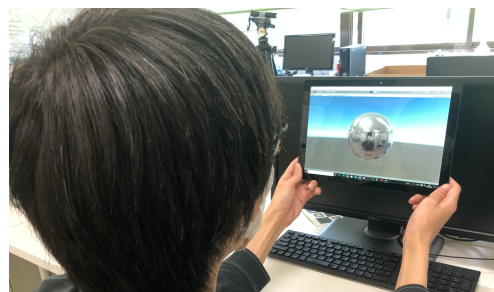


図 5: モバイル AR による再現



外挿結果



映り込み結果

図 6: モバイル AR による再現結果

#### 参考文献

- [1] 仁科勇作, 奥村文洋, 神原誠之, 横矢直和. 適応的なハイダイナミックレンジ画像合成による拡張現実感のための光源環境推定. 情報処理学会研究報告コンピュータビジョンとイメージメディア(CVIM), Vol. 2008, No. 3 (2008-CVIM-161), pp. 185-190, 2008.
- [2] 佐賀野正行, 安室喜弘, 眞鍋佳嗣, 千原國宏. 鏡面球を用いた拡張現実感における映り込み表現. 第 50 回システム制御情報学会研究発表講演会講演論文集, p. 331, 2006.
- [3] Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Andrew Tao, Jan Kautz, and Bryan Catanzaro. High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 8798-8807, 2018.
- [4] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.5967-5976, 2017.
- [5] Naoki Kimura and Jun Rekimoto. Extvision:

Augmentation of visual experiences with generation of context images for a peripheral vision using deep neural network. In Proceedings of the CHI Conference on Human Factors in Computing Systems, Paper No. 427, 2018.

- [6] Gregory J. Ward. Measuring and Modeling Anisotropic Reflection. In Proceedings of the 19th Annual Conference on Computer Graphics and Interactive Techniques, pp. 265-272, 1992.