

## GAN を用いた隠消現実感 –補完機能の改善–

## Diminished Reality using GAN - Improving complementary functions -

D-11

小林 海人

Kaito KOBAYASHI

芝浦工業大学

Shibaura Institute of Technology

高橋 正信

Masanobu TAKAHASHI

システム理工学部

College of Systems Engineering and Science

## 1. 背景

隠消現実感(Diminished Reality 以下 DR)とは、現実物体に背景画像を上書きして表示することにより、現実空間から対象の物体が削除されたように見せる技術である。我々は、スマホのカメラを用いた DR アプリケーションを開発している[1]。これは、GAN (Generative Adversarial Networks)をベースとした深層学習を用いて対象物体を隠すための背景画像(以下補完画像)を生成する手法を用いている。しかし、DR が行える背景は単一な床面のみに限定されており、補完精度も十分とはいえず、原理実証の段階であった。そこで、本研究では補完機能に着目し、その改善を図った。

## 2. 目的

DR アプリケーション[1]の使われているネットワーク構成を改善することで、ネットワークが生成する対象物体を隠すための補完画像の精度を向上する。

## 3. 従来手法

処理の流れを図1に示す。従来手法[1]では、カメラ画像から推定した現実平面に仮想平面を構築する。その後、削除したい物体を含むように仮想円柱を配置し、円柱内部の3D空間を削除領域とする。次に、カメラパラメータを元に撮影画像中の2Dの削除領域を求める。削除領域は、画像補完機能で補完する領域(以下補完領域)であり、補完領域はその周辺領域を元に補完される。画像補完には生成器、局所識別器、大域識別器の3つのネットワークで構成されたGANに基づく画像補完手法[2]を利用している。

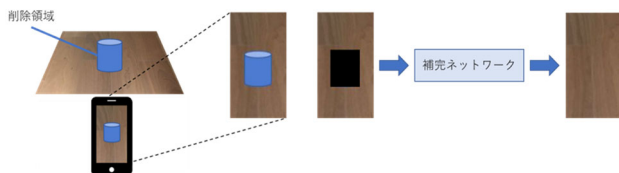


図1 従来手法の処理の概要[1]

## 4. ネットワーク構成の改善

DR アプリケーション[1]の補完機能を改善する方策として、ネットワーク構成の改善を行った。

新たな2種類のネットワークによる改善をし、処理速度と補完精度の観点から比較した。ネットワークはそれぞれ、補完する際にどの特徴量チャンネルが重要かを重み付けする SE-Net を採用したネットワーク[3]と、補完領域内の各ピクセルが補完領域外のどこのピクセルと相関があるかを学習し、画像生成に活用する Contextual Attention を採用したネットワーク[4]である。

## 5. 実験

## 5.1 データセット

従来手法[1]では床面画像を利用したが、本研究ではより多様な室内の画像[5][6][7][8]を利用した。室内画像は画素数を128×128画素に変換し、学習用に約16,000枚、評価用に約1,800枚用意した。

## 5.2 ネットワークの学習

学習用画像中のランダムな位置に16~64×16~64のランダムな形、大きさの四角形を補完領域として設定した画像を入力画像とし、補完領域の元画像を

正しい補完結果として学習した。

## 5.3 評価方法

補完精度は補完領域の大きさに依存するため、評価時の補完領域は13種類の大きさ(16×16, 20×20, ..., 60×60, 64×64)の正方形とし、それぞれの大きさと精度を評価した。評価指標としてはPSNRとSSIMを用いた。また、リアルタイム性の評価のため、補完画像の生成速度[fps]を求めた。

## 5.4 結果

図2は、元の画像と補完領域を設定した入力画像、および各ネットワークで補完した画像例である。



図2 各ネットワークの画像補完結果例

評価用画像に対する従来のネットワークとの平均評価値の差を図3に示す。新たに導入した2つの手法の評価値は何れも従来手法より優れていた。2つの手法の中では、補完領域が小さいと Contextual Attention を採用したネットワークが、領域が大きいと SE-Net を採用したネットワークが優れていた。Contextual Attention は、領域が大きいとき、ピクセル同士の相関付けがうまく学習できていないことが考えられる。

生成速度の計測結果を表2に示す。実行環境は、CPU: AMD Ryzen 9 3900, GPU: NVIDIA GeForce RTX 3080 である。生成速度の面では、階層数が少ない SE-Net を採用したネットワークが従来と同等以上に速かった。したがって、総合的には SE-Net を採用したネットワークが DR アプリケーションの補完機能に適していると考えられる。

今後は、生成速度を可能な限り落とさずに補完精度のさらなる向上に取り組みたい。

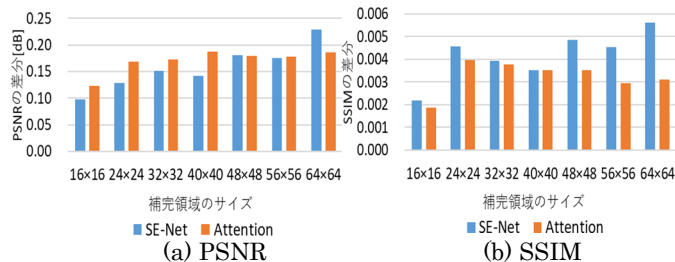


図3 補完領域の精度(一部のサイズ)

表2 補完画像の生成速度[fps]

	従来	SE-Net	Attention
	22.8	23.1	19.4

## 【参考文献】

- [1] 澤田悠暉, 他: 2019年度電子情報通信学会東京支部学生会研究発表会, 60, 2020.
- [2] Iizuka, et al.: ACM Trans. on Graphics, Vol. 36, No. 4, pp.1-14, 2017.
- [3] Chen, et al.: The Visual Computer, Vol.37, issue.7, pp.1691-1705, 2021.
- [4] J. Yu, Z. Lin, et al.: arXiv:1801.07892, 2018.
- [5] Quattoni, et al.: IEEE CVPR 2009, pp.413-420, 2009.
- [6] "Pixabay", <https://pixabay.com/ja/>, 2021年9月22日アクセス.
- [7] "Unsplash", <https://unsplash.com/>, 2021年10月18日アクセス.
- [8] "GAHAG", <https://gahag.net/>, 2021年10月18日アクセス.