

GAN を用いた撮像人体の体形を自然に変換する画像合成手法

青木 雄亮† 田村 仁‡

日本工業大学 工学部 創造システム工学科†‡

1. 研究背景

機械学習による画像合成を応用して人間の顔や身体を変形する，人体に装飾を加える等の研究[1]が存在する．しかし人間の体形を，違和感なく変換させるものは実現されていない．例えば，スマートフォンアプリにこの変換処理を行うものが存在しているが，人体だけではなく周りの背景も変形してしまうものが多い．自然な体形変換が実現出来れば，服飾業界での個々人に合った服選びや，理想の体形を目標にボディメイクやダイエットに励む人々のモチベーション向上に一役買うことが出来る．本研究では，入力画像が学習モデルに則って変換され，目標の体形に自然に変換された画像の生成を目的とする．

2. 提案手法

本研究では，Pix2Pix という敵対的生成ネットワーク (GAN) の一種を用いる．このアルゴリズムは元画像と正解画像をペアで学習させ，学習モデルを作成する[2]．その上で入力された画像を学習モデルに従って処理を施す．したがって，本実験では大きく分けて 2 種類のデータを用いる．モデルを作成するための学習用データと，学習モデルが正しく動作するか評価する為の評価用データである．

人体を対象にした画像処理を施す研究では，データとして現実の人体を用いることが多いが，本研究では 3DCG の人体モデルを使用した．この理由は標準的な体形の人体，筋肉質な体形の人体など，あらゆる体形のモデルや，それらをあらゆる角度から撮影した画像を得ることが容易なためである．今回は Poser Pro11 という 3DCG ソフトウェアを用いた．

3. 実験

Poser Pro11 で標準的な体形の男性の画像と筋肉質な体形の男性の画像を用意した．今回使用した画像は 256×256 ピクセルのサイズにリサイ

ズしたうえで使用した．それぞれ学習用データとして元画像と正解画像をそれぞれ 12800 枚，評価用データもそれぞれ 12800 枚ずつ用意し，前面，背面，左右側面，それらの俯瞰画像，見上げた画像など様々な角度から映した画像を使用した．その画像を一对の画像になるようデータセットを作り，それを Pix2Pix を用いて学習させた．学習用データの画像と，処理の例を図 1 に示す．また，1つの訓練画像の学習回数となるエポック数[3]は今回は 100 として行った．

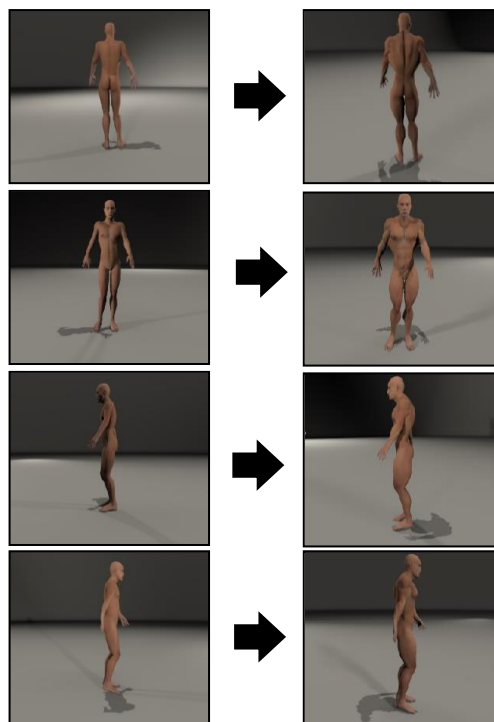


図 1： データと処理の例

4. 実験結果

この実験における理想的な結果と，今回得られた画像を図 2 と図 3 に示す．図 2 の左の画像を入力された画像，右の図を出力画像とする．本来理想とするのは図 2 のような結果が得られることである．図 3 は今回の実験で得られた画像である．出力された画像は人体のシルエットは認識できるが，想定外の着色が背景，人物になされている．

また，今回得られた結果画像と正解画像の誤

Image composition method of deform a shape of human body using GAN

†Aoki Yusuke

‡Tamura Hitoshi

‡Nippon Institute of Technology

差を数値で示すため、PSNR, MAE, MSE の 3 つの指標を用いた。PSNR(Peak Signal to Noise Ratio)とは、画像や音声などの輝度や信号の最大値とノイズの比である [4]。MAE(Mean Absolute Error)は元画像と結果画像の画素ごとの平均絶対誤差を表し [5]、MSE(Mean Squared Error)は差分 2 乗平均を表す [6]。これらの指標は [7] を参考に選択した。

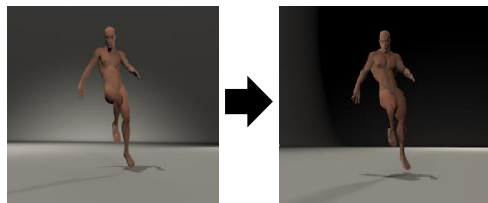


図 2: 理想的な結果

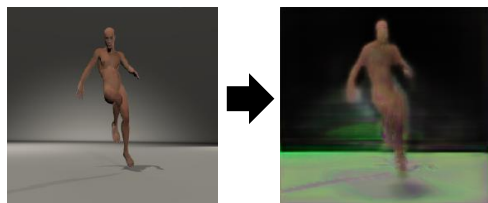


図 3: 得られた結果

本実験で PSNR の数値の最大値は 100 であり、最大値に近づくほど 2 枚の画像の誤差が少ない事を表す。例えば、同一画像では 100 になる。MAE, MSE の数値は 0 に近づくほど誤差が少ない事を表し、同一画像では 0 になる。

これら 3 つの指標を用いて、まず元画像、正解画像のペアからランダムに 100 組ずつを選び、それらの差の数値をプログラムで計算させた。得られた数値を表 1 に示す。

表 1: 体形変化による類似度

類似度指標	数値
PSNR	18.62db
MAE	51.37
MSE	28.59

次に正解画像と評価用の結果画像の差の計算結果を表 2 に示す。この表 1 と表 2 の 2 つの計算結果の差から、PSNR の値には大きな差は見られないが、MAE, MSE の 2 種類の指標の結果には大きな差が見られ、体形変化による類似度との誤差が MAE では約 130, MSE では約 45 大きい。

表 2: 正解画像と合成結果との類似度

類似度指標	数値
PSNR	14.72db
MAE	181.1
MSE	74.97

5. 考察

実験結果から、体形変化による類似度より正

解画像と評価用の合成画像との類似度はいずれも悪い。この理由は、本実験で使用したデータは元画像と正解画像で背景の明るさが異なり、照明の条件もやや異なることと考えられる。よって、その条件を一对の画像でより厳密に対応させることが出来ればより自然な処理が得られると考える。また、人体だけではなく背景も含めて変化させる対象として学習していると考えられる。このことに関しては、人体のみを認識させる手法を用いる事が解決手法として挙げられる。具体的には、姿勢推定を行う機械学習モデルである PoseNet などの技術を用いて人体の姿勢を骨格で推定したうえでその推定した骨格を基に肉付けするように、想定する体格の人体を合成するという技術を用いることが考えられる。この手法と今回の手法を組み合わせることでより精度を高めることが可能であると考えられる。

誤差の数値が大きく示された原因として着色された部分がノイズとなり、その面積が多分に含まれていることが挙げられる。よって、着色部分を減らす方法を考慮する事でより誤差を減らすことが出来ると考える。

6. まとめ

本研究では Pix2Pix を用いて人体の体形を画像上で変形させる実験を行った。実験の結果、今回の実験手法ではノイズが多くなるなどして、求める結果との誤差が大きくなった。本研究の結果から最も考慮すべき点は、いかにノイズを少なくし、人物に焦点を絞った処理を施すかという点だと考えられる。よって今後、その手法を取り入れる、より学習データを増やす、条件を改良するといった事を考慮し、精度向上を図る。

参考文献

- [1] 久保 静真, 岩澤 有祐, 鈴木 雅大, 松尾 豊 “服の領域を考慮した写真上の人物の自動着せ替えに関する研究” 情報処理学会論文誌 60 巻 3 号 pp. 870-879 発行年 2019 年 3 月 15 日
- [2]<<https://qiita.com/mine820/items/36ffc3c0aea0b98027fd>> 2019 年 6 月 5 日アクセス
- [3]<<https://www.st-hakky-blog.com/entry/2017/01/17/165137>> 2019 年 7 月 17 日アクセス
- [4]<http://www.ieice-hbkb.org/files/02/02gun_05hen_09.pdf> 2020 年 1 月 9 日アクセス
- [5]<<https://mathwords.net/rmseae>> 2020 年 1 月 9 日アクセス
- [6]<http://www.ieice-hbkb.org/files/03/03gun_05hen_01.pdf> 2020 年 1 月 9 日アクセス
- [7] 皆藤 優太, 田村 仁 “深層学習を用いた輪郭線情報に着目した画像修復” 第 81 回全国大会 2019 1 号 pp. 185-186 発行年 2019 年 2 月 28 日