

StyleGAN2 に基づく表情差分付き ゲームキャラクターの顔画像の自動生成

陳 晴† 阿倍 博信†

東京電機大学†

1. はじめに

キャラクターを動かすゲームの開発において、表情や顔向きの異なる表情差分の付いたゲームキャラクターの顔画素材が大量に必要な。通常は、ネットで探す、デザイナーに依頼する、自分で作る、などの方法が考えられるが、オリジナル素材を使用したい場合、デザイナーに依頼するか自ら制作する必要がある。このような背景から、オリジナルの顔画素材の準備作業の効率化を目的として、ゲーム制作支援のための表情差分付きの顔画像の自動生成を提案する。本研究にて提案するゲーム制作支援のための表情差分付きのゲームキャラクターの顔画像の自動生成により、ゲームの開発現場においてオリジナル素材の利用と作業効率の向上を図る。

本研究では、StyleGAN2[1]を用いた顔画素材の自動生成を提案する。StyleGAN2 は画像生成を目的とした GAN(Generative Adversarial Network)のひとつである StyleGAN[2]の改良版であり、主にスタイルを適用した画像生成に使用されている。本研究では StyleGAN2 を使用した表情差分付きのゲームキャラクターの顔画像の自動生成について、ゲームキャラクターの顔画像のデータセットの作成、StyleGAN2 によるモデル構築、潜在空間の分析、表情差分付きゲームキャラクターの顔画像の生成、有効性の評価について報告する。

2. StyleGAN2 における潜在空間の分析

StyleGAN2 では、指定のレイヤーの潜在変数をミックスすることによって、新たな画像を生成する Style Mixing という機能があり、本研究においても Style Mixing を用いて表情差分付きのゲームキャラクターの顔画像を生成する。

特定の画像を事前に学習した GAN モデルの潜在空間に反転させ、ジェネレータによって反転されたコードから画像を再構築することを目的とした GAN Inversion[3] という技術が注目されており、StyleGAN2 においても事前学習済みの GAN モデルを画像編集アプリケーションで使用する際に重要な役割を果たす。StyleGAN2 に対応した潜在空間の分析手法は、主成分分析に基づくもの[4]、潜在セマンティック解析に基づくもの[5]があるが、本研究では、顔の向きや表情などの表情差分生成に必要な特

徴ベクトルを発見するために、StyleGAN2 で構築した潜在空間において、各特徴の重心を求める手法を用いて探索し、見つけた重心を元に、StyleGAN2 に基づく表情差分付きゲームキャラクターの顔画像の自動生成を行う方針とする。

3. 提案手法

3.1 方針

同じパラメータの画像を大量集めて、指定のレイヤーの潜在変数の重心を求める手法で探索し、構築した潜在空間から各特徴の重心を求める。見つけた重心を元に顔の向きや表情などの表情差分付きのゲームキャラクタの自動生成を可能にする。

3.2 提案手法の概要

アニメの顔画像を学習済みの StyleGAN2 モデル[6]に対して、二次利用可能なゲームキャラクターの顔画像データ[7]をデータ拡張したデータセットを転移学習させたモデルを構築する。構築したモデルの潜在空間から各特徴の重心を求める手法を用いて探索し、見つけた重心を元に顔の向きや表情などの表情差分付きのゲームキャラクターの顔画像の自動生成を行う。

3.3 提案内容

(1) データセットの作成

ガイドラインに従う形で二次利用が許諾されているスマホ RPG ゲームのゲームキャラクターの顔画像データ[7]に対し、モンスターなどの人物以外の顔画像データを削除し、別途準備した顔画像データ(130 枚)を追加したデータを準備した。その後、データ拡張として回転(15°, 30°, 45°)、左右反転処理を行った結果、2,976 枚となった。今回作成したデータセットの各パラメータについて、口があいている笑顔を Smile1 に、口が閉まっている笑顔を Smile2 とし、表情に関するパラメータを表 1 に、顔向きに関するパラメータを表 2 に示す。

表 1. 顔画像データセットにおける表情パラメータ

Parameter	Smile1	Smile2	Other	Total
Page	1096	912	968	2976
%	37%	31%	32%	100%

表 2. 顔画像データセットにおける顔向きパラメータ

Parameter	Left	Right	Other	Total
Page	2392	424	160	2976
%	80%	14%	6%	100%

Facial image generation of game character with facial expression difference based on StyleGAN2

† Qing Chen, Hironobu Abe: Tokyo Denki University

(2) StyleGAN2 によるモデル構築

アニメの顔画像を学習済みの StyleGAN2 モデル[6]に対して、(1)でデータ準備したデータセットを学習データとして転移学習を行う形で、ゲームキャラクタの顔画像を生成する StyleGAN2 のモデル構築を実施した。学習時間は23時間58分37秒となった。

(3) 潜在空間の分析

次に、構築した StyleGAN2 のモデルの潜在空間を下記のステップに基づき分析を行った。

Step1: まず、(2)で構築した StyleGAN2 のモデルを使用して1,000枚の顔画像を生成する。生成した1,000枚の画像を Smile1, Smile2, Left, Right の観点から目視で確認する形で、各パラメータ(Smile1, Smile2, Left, Right)の重心の初期値として、4つの画像(Smile1, Smile2, Left, Right)の重心を抽出する。次に、同じモデルを使用して、新たに200枚の顔画像を生成する。生成した顔画像に対して、4つのパラメータ(Smile1, Smile2, Left, Right)の重心に向けて潜在方向の変更処理を行う。目視により、画像の確認を行い、成功画像は残し、失敗画像は削除する。最後に、成功画像として残った各パラメータ(Smile1, Smile2, Left, Right)の重心を算出して保存する。

Step2: 同じモデルを使用して、更に新たに200枚の顔画像を生成する。生成した顔画像に対して、保存した各パラメータ(Smile1, Smile2, Left, Right)の重心に向けて潜在方向の変更処理を行う。目視により、画像の確認を行い、成功画像は残し、失敗画像は削除する。最後に、成功画像として残った各パラメータ(Smile1, Smile2, Left, Right)の重心を算出して保存する。

Step3: Step2を繰り返す。

4. 評価

各ステップに、200枚の画像を指定の潜在方向へ変更させた成功率に関する結果を表3に示す。

表3. 各ステップにおける Style Mixing の結果

Parameter	Smile1	Smile2	Left	Right
Step1	37%	34%	25%	30.5%
Step2	30.5%	36%	30.5%	30.5%
Step3	49%	38%	48%	36.5%

また、Step2 から求めた潜在空間の重心で Style Mixing した結果を図1に示す。



©2014 CloverLab., Inc.

図1. Step2における Style Mixing の結果

5. 考察

潜在方向に関しては、潜在空間の重心を求める手法で特定できることが確認できた。表3の結果から、一枚の画像の潜在方向と複数枚画像の潜在空間の重心とを比較し、一枚の画像から、複数枚画像から求めた潜在方向の方が指定の潜在方向へ変更させた成功率が高いことが分かった。大量のデータから最も適切な潜在方向を特定し、指定の潜在方向へ変更する際の成功率の向上が可能と考察する。

しかし、指定の潜在方向に変更する前に、元画像の状態を確認する注意点がある。例えば、図1の結果から、seed0000のような元画像の口が開いている状態で、Smile1に変更させた場合、逆に口が閉まる状態になる場合があった。顔の向きは予想通りに変更させることができたが、あわせて髪の毛の形も変わってしまう場合があった。そのため、元画像の状態を確認しておかないと、予想通りの画像が得られない可能性がある。

6. おわりに

本研究では、StyleGAN2 を用いて構築したモデルの潜在空間から、各特徴の重心を求める手法を用いて探索を行い、算出した重心を元に顔の向きや表情などの表情差分付きのゲームキャラクタの顔画像の自動生成を行う方法を提案した。Smile 以外の他の種類の表情差分付きのゲームキャラクタの顔画像に対応するためには、学習データに対応した表情差分付きの顔画像の追加が必要である。

今後の課題として、他の表情差分の生成に必要な特徴ベクトルを発見する手法の検討、画像生成及び Style Mixing 機能や生成画像の編集機能を持つ Web アプリケーションの構築、などがある。

参考文献

- [1] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, Timo Aila: Analyzing and improving the image quality of stylegan, <https://arxiv.org/abs/1912.04958> (2019).
- [2] Tero Karras, Samuli Laine, Timo Aila: A style-based generator architecture for generative adversarial networks, <https://arxiv.org/abs/1812.04948> (2018).
- [3] Weihao Xia, Yulun Zhang, Yujiu Yang, Jing-Hao Xue, Bolei Zhou, Ming-Hsuan Yang: GAN Inversion: A Survey, <https://arxiv.org/abs/2101.05278> (2021).
- [4] Erik Härkönen, Aaron Hertzmann, Jaakko Lehtinen, Sylvain Paris: GANSpace: Discovering Interpretable GAN Controls, <https://arxiv.org/abs/2004.02546> (2020).
- [5] Yujun Shen, Bolei Zhou: Closed-Form Factorization of Latent Semantics in GANs, <https://arxiv.org/abs/2007.06600> (2020).
- [6] Gwern Branwen: Making Anime Faces With StyleGAN, <https://www.gwern.net/Faces>.
- [7] クローバーラボ: ゆるドラシル RPG ツクール素材提供サイト, <http://yurudora.com/tkool/>.