

SwipeGANSpace: 潜在空間の関心次元探索によるスワイプ操作に基づく嗜好画像生成

中島 柚斗* 楊 明哲* 馬場 雪乃*

概要. GANは従来の生成モデルより高品質な画像を生成できる。GANの潜在変数を操作することで、好みの画像（嗜好画像）を生成できる。GANを利用して嗜好画像を生成するには、適切な潜在変数をGANに与える必要がある。しかし、GANの潜在空間は高次元であり、探索が困難である。既存研究では、ユーザーに複数のスライダを調整してもらい、画像を編集することで、嗜好画像を生成している。本研究では、より簡単なスワイプ操作によって嗜好画像を生成する手法を提案する。探索を効率化するために、潜在空間を主成分分析で縮小し、縮小した空間で探索を行う。ここで、潜在空間上でどの次元がユーザーにとって関心があるかを把握するのは難しい。この課題を解決するために多腕バンディットアルゴリズムを利用して、探索すべき次元を動的に決定し、ユーザーの関心次元を重点的に探索することで、さらなる探索の効率化を行う。実験の結果、提案手法がベースラインと比べて、効率的に嗜好画像を生成できることを明らかにした。

1 はじめに

高品質な画像生成が可能な Generative Adversarial Networks (GAN) [5] は、その潜在変数を操作することで生成画像を編集できる [6]。GANを利用してユーザーが好みの画像（嗜好画像）を生成するには、適切な潜在変数をGANに与える必要がある。しかし、GANの潜在空間は高次元であり、嗜好画像を生成する適切な潜在変数を探索するには、膨大な時間と労力がかかる。探索を効率化するため、ユーザーに複数のスライダを調整させる手法や画像を編集させる手法が提案されている [2]。これらの方法は、特にスマートフォンのような小さな画面では操作が難しい。また、嗜好画像を生成する方法として、自然言語から画像を生成する text-to-image の手法 [14] があるが、ユーザーにとって嗜好画像を生成する適切なプロンプトを考えることは難しい場合がある。

本研究では、ユーザーの負荷を軽減する画像生成手法を提案する。具体的には、ユーザーにスワイプ操作のみを求める。図1に提案手法のユーザーインターフェースを示す。ユーザーは画面に提示された画像を左右にスワイプすることで嗜好を表現する。スワイプ操作のみを利用して嗜好画像を生成するには効率的に潜在変数を探索する必要がある。そこで本研究では、StyleGANの潜在空間に主成分分析を適用し、画像の見た目を大きく変える主成分で構成された部分空間を作成する。この部分空間の中で、ベイズ最適化を用いて潜在変数探索の効率を高める。ここで、GANの表現能力を維持するためには、部分空間の次元を大きく保つことが必要である。しかし、部分空間の次元を大きくすると、ベイズ最適化は探索効



図1: スワイプ操作による問い合わせ例

率が悪くなるため、ベイズ最適化のみを利用して嗜好画像を生成することは困難になる。この問題に対処するために、我々は多腕バンディットアルゴリズムを採用する。このアルゴリズムによって、どの次元を探索すべきかを動的に決定し、特にユーザーに関連する次元に焦点を当てることで、探索効率を高める。

提案手法がベースラインと比べて効率的に嗜好画像を生成できるかを検証するために、シミュレーション実験とユーザー実験を行った。その結果、提案手法がベースラインと比べて、効率的に嗜好画像を生成できることを確認した。

2 関連研究

GANを用いた嗜好画像生成を行うために、ユーザーに複数のスライダ調整や画像編集を繰り返し問い合わせる手法がある [2]。この手法では、人間の嗜好を目的関数、GANの潜在変数を探索変数とした最適化問題を考え、human-in-the-loop最適化を行

うことで嗜好画像を生成した。human-in-the-loop 最適化とは、人間を計算資源とみなして最適化の反復計算に組み込む手法である。特に人間の嗜好を対象とした最適化問題では、人間が最適化の反復に入ることが有用である。近年、human-in-the-loop 用のベイズ最適化の手法の研究が進められている [10, 11, 2, 4, 8, 16]。ベイズ最適化とはブラックボックス最適化手法の 1 つである。ベイズ最適化は他のブラックボックス最適化手法と比べて、目的関数の評価回数を抑えることができるため、人間の評価を必要とする目的関数の最適化に適している。このアプローチでは、特に人間にとって評価が容易な問い合わせを行うことや、問い合わせ回数を少なくすることが重要である。一般的な次元数のパラメータ空間を扱う場合のこのアプローチでは、スライダ調整 [11] や N 対比較 [10, 16] など、様々な手法が提案されている。特に、既存研究 [2] は GAN の潜在空間が高次元であるため、複数のスライダ調整や画像編集など複雑な問い合わせをしているが、これは問い合わせの難易度が高く、ユーザの負担が大きい。

3 提案手法

3.1 画像生成のためのスワイプ操作

本研究では、嗜好画像を簡単なスワイプ操作のみで生成できるように設計した新しいユーザインタフェースを提案する。このインタフェースでは、スワイプ操作を繰り返し行うことで、ユーザの好みを捉えて嗜好画像を生成する。インタフェースは、Tinder のようなマッチングアプリを参考に、図 1 のように一度に 1 枚の画像を表示する。ユーザは、現在表示されている画像と以前のものとを比較する一対比較を求められる。右にスワイプすることで現在の画像に対する好意を示し、左にスワイプすることで以前の画像に対する好意を表現する。このフィードバックに応じて、次の画像が生成され、ユーザに提示される。このプロセスは、嗜好画像が生成されるまで繰り返される。

3.2 効率的な潜在空間探索のための提案手法

3.2.1 画像生成モデル

本研究では GAN の一種、StyleGAN [7] を用いて画像を生成する。StyleGAN では、ランダムベクトル z を Mapping Network に入力し、中間潜在変数 w を得る。この w は年齢や性別といったスタイル情報を反映し、 w を Synthesis network 入力することで画像が生成される。これにより、生成画像のスタイルの制御が可能であり、本研究ではこの利点を活用する。

3.2.2 目的関数

嗜好画像を生成するためには、最適な潜在変数を探索する必要がある。そこで本研究では、人間の嗜好を目的関数、StyleGAN の中間潜在変数を探索変数とした最適化問題を考え、これを最適化する中間潜在変数 $w \in W$ を探索する。StyleGAN の Synthesis network を $s: W \rightarrow X$ とする。ここで、 $W \subset \mathbb{R}^d$ は $d = 512$ の高次元空間であり、 X は画像データ空間である。ユーザはこの生成モデル s を利用して画像データ $x \in X$ を生成しようとしているが、制御できるのは中間潜在変数 w のみである。画像データ x への嗜好度は関数 $g: X \rightarrow \mathbb{R}$ によって決定されるが、これはシステムにとって未知である。最終的な目標は、以下の最適化問題を解き、嗜好度の高い画像を生成することである。

$$w^* = \arg \max_{w \in W} g(s(w)) \quad (1)$$

ここで、 $g(s(w))$ は任意の点 w に対して計算可能であるが、 $g(\cdot)$ の計算には人間への問い合わせが必要であるため $g(\cdot)$ は評価コストが高く、評価回数できるだけ少なくすることが重要である。

3.2.3 一対比較データを利用するベイズ最適化

本研究では、式 (1) を解くために、目的関数の評価に人間の入力を利用したベイズ最適化を用いる。 g と s の合成写像 $f = g \circ s$ を定義し、 f を目的関数、 w を探索変数としたベイズ最適化を考える。提案手法は、ユーザから直接 f の値を受け取る代わりに、ユーザの一対比較結果を受け取る。Chu らの手法 [3] を利用して、ベイズ推定の枠組みをもとに、一対比較の結果を実数値に変換し、 f を計算する。以降は既存のベイズ最適化と同様、ガウス過程帰帰と獲得関数の最適化を行い、評価候補点を決定し、ユーザに問い合わせる目的関数の評価を行う。ユーザに問い合わせる際には、StyleGAN の Synthesis network を利用して、 w を画像 x に変換する。

3.2.4 次元削減

一対比較では、各反復で得られるユーザの嗜好情報が少なく、高次元空間で適切な潜在変数を探索するには、ユーザへの問い合わせ回数が増加するという課題がある。これを解決するために、探索空間を小さくすることで、問い合わせ回数を削減する。Elena ら [13] は、高次元空間でベイズ最適化を行う場合、探索空間に主成分分析 (PCA) を適用し、実探索空間を縮小することで、効率的な探索が可能であることを示した。GANSpace [6] では、StyleGAN の潜在空間に PCA を適用し、画像の見た目を大きく変化させる主成分を特定している。これらから着想を得て、StyleGAN の潜在空間に PCA を適用し、画像の見た目を大きく変える主成分のみで構成される部分空

間での探索を行うことで、潜在変数の探索を効率化する。

具体的な手続きは以下の通りである。まず、標準正規分布から N 個のベクトル $z \in \mathbb{R}^d$ をランダムサンプリングする。次に、サンプリングした z を StyleGAN の Mapping network に入力することで、中間潜在変数 $w \in W \subset \mathbb{R}^d$ を獲得する。獲得した中間潜在変数 w を利用して、主成分分析を行い、寄与率上位 d' 個の主成分で部分空間 $W' \subset \mathbb{R}^{d'}$ を作成する。この部分空間 W' 内でベイズ最適化を行う。 W' 内でのベイズ最適化で提案される評価候補点 $w' \in W'$ の次元数は、 $d' < d$ であるため、 w' を直接 StyleGAN の Synthesis network に入力することはできない。そこで、縮小された次元を戻すために主成分分析の逆変換を行うことで、 w' を StyleGAN の潜在空間 W 上に写像し、 Synthesis network に入力する。 Synthesis network へ入力して得られた生成画像 x をユーザーに提示し、直前の生成画像との一対比較結果を獲得する。

3.2.5 重要な次元の探索

GAN の表現能力を維持するためには、部分空間の大きさを大きく保つ必要がある。しかし、部分空間の次元数を大きくすると、高次元空間ではベイズ最適化の性能が低くなり、ベイズ最適化だけでは嗜好画像の生成が難しくなる。そこで本研究では、探索効率を高めるために、多腕バンディットアルゴリズムを利用して、探索すべき次元を動的に決定する方法を提案する。具体的には、あらかじめ各次元に対してベイズ最適化モデルを用意し、多腕バンディットアルゴリズムを用いて探索する次元を選択し、その次元に対してベイズ最適化モデルを用いて候補点を決定する。そして、候補点を画像化し、ユーザーのフィードバックを得る。多腕バンディットモデルは、ユーザーからの一対比較結果を用いて更新される。このアプローチを用いることで、部分空間に存在する多様な特性を持つ次元から、最も価値のある次元を重点的に選択し、効率的に探索することができる。

本研究では、多腕バンディットアルゴリズムの中で基本的なアルゴリズムである、UCB アルゴリズムを利用する [1]。UCB アルゴリズムでは、イテレーションごとに各次元の UCB スコアを計算し、最も UCB スコアが高い次元を探索する。 i^* を探索する次元とし、次元 i に関する UCB スコア U_i とすると、

$$i^* = \arg \max_{i=1}^{d'} U_i \quad (2)$$

となる。次元 i に関する探索回数を N_i 、報酬の推定値を \hat{r}_i とすると、多腕バンディットモデルを以下のように更新する。

$$\hat{r}_i = \frac{\sum_{j=1}^{N_i} r_{i,j}}{N_i} \quad (3)$$

ここで、 $r_{i,j} \in \{0,1\}$ は i 次元目が j 回目に選ばれた時の報酬である。直前の生成画像が選ばれたら $r_{i,j} = 0$ 、現在の生成画像が選ばれたら $r_{i,j} = 1$ とする。選ばれた次元 i^* に関して、ベイズ最適化を利用して候補点の i^* 成分を決定する。獲得関数を A 、候補点の i^* 成分を w_{i^*} とすると、

$$w_{i^*} = \arg \max_{w \in R} A(w) \quad (4)$$

i^* 以外の次元については、過去の観測結果の各次元でのベイズ最適化モデルの最良値を割り当てる。各次元での最良値を $w_0, w_1, \dots, w_{d'}$ とすると、

$$w'_t = [w_0, w_1, \dots, w_{i^*}, \dots, w_{d'}] \quad (5)$$

w'_t を主成分分析の逆変換と Synthesis network を利用して画像に変換し、ユーザーに問い合わせる。

4 シミュレーション実験

4.1 実験設定

ユーザーが安定した回答を提供できると仮定した場合に、提案手法は効率的に嗜好画像を生成できるかを検証するために、シミュレーション実験を行った。我々は、提案手法からバンディットアルゴリズムを除いた対抗手法 (SimpleBO) と提案手法 (BanditBO) を比較した。目標画像を事前に設定し、生成画像と目標画像の類似度を計算し、1つ前のイテレーション時と比較することで、一対比較の結果を作成した。この一対比較結果をシステムに入力し、ベイズ最適化を行う。部分空間 W' の次元数は、 $d' \in \{4, 8, 16\}$ と設定した。10枚の目標画像を、探索空間からランダムサンプリングして用意した。本実験では、Flickr-Faces-HQ というデータセットで学習された、StyleGAN の事前学習済みモデルを利用した [7]。また、2つの画像の類似度を計算するために FaceNet [15] を使用した。FaceNet は、顔画像を埋め込みベクトルに変換する顔認識用の CNN モデルである。FaceNet を用いて、生成画像と目標画像の埋め込みベクトルを求め、埋め込みベクトル間のコサイン類似度を計算する。

4.2 実験結果

図2に、 d' ごとの、類似度の移動平均を示す。 $d' = 4$ の場合は、BanditBO と SimpleBO の差は大きく見られなかった。 $d' = 8$ の場合は、 $d' = 4$ の時と比べて、BanditBO が SimpleBO より効率的に目標画像に近くようになった。 $d' = 16$ の場合は、さらに BanditBO が SimpleBO より効率的に目標画像に近づいた。以上から、探索空間が高次元の場合に、提案手法は対抗手法と比べて効率的に目標画像に近づくことがわかった。

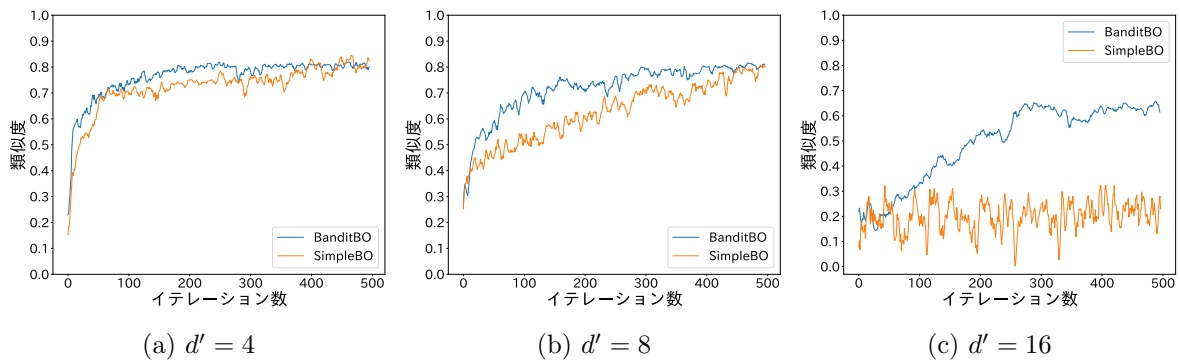


図 2: 各 d' に関する, 生成画像と目標画像の類似度の推移 (10 枚の目標画像での移動平均の推移)

5 ユーザ実験

5.1 実験設定

ユーザの実評価を用いたときに, 提案手法がベースラインと比較して効率的に嗜好画像を生成できるかどうかを検証するために, ユーザ実験を行った. 被験者には, 指定したシナリオに沿った好みのアバターを生成する実験を行った. 指定するシナリオは6個(弁護士, 教師, スポーツインストラクター, 受付, 図書館のスタッフ, 保育士)用意した. 被験者は14人(男性13人, 女性1人, 年齢:18歳から27歳)であった. 被験者ごとに実施されるシナリオの順番はランダムである. 各シナリオに対して, BanditBO, SimpleBO, 探索空間から一様分布で中間潜在変数をランダムサンプリングする手法(Random)の3つの手法と, 2つの部分空間の次元数 $d' = 4, 16$ の組み合わせをランダムに割り当てた. 実験中のユーザの嗜好を確認するために, 各シナリオ開始前に, 事前の嗜好画像のイメージを尋ねた. 被験者には, 事前アンケートの内容に固執しないように, 与えられたシナリオのみに焦点を当てて画像の比較を行うように指示した. 各シナリオでの回答終了後に, 画像の提示・比較プロセスおよび最終アバターの満足度に関するアンケートを実施した. 全てのシナリオ終了後, 実験全体を通して感じたことをアンケートで尋ねた.

5.2 実験結果

図3に, 「目的の画像まで効率的に辿り着くことができましたか?」というアンケートの結果を示す. これを見ると, 提案手法がベースラインと比べて, 目的の画像まで効率的に辿り着くことができたという回答した人が多かったことがわかる. 提案手法についての自由記述では, 「微調整しながら, 途中で大幅に違うものと比較しては元の路線に戻ってくることを繰り返す, 良い生成をしていた」, 「適当に画像を出してきて, 特に調整をしないシナリオや, ずっと微調整だけを繰り返すシナリオなどさまざまなシナリオがあったが, それらを混ぜ合わせたような, 微調

整をしつつ, 途中で一回大幅にずれてみて必要に応じてそっちに移るといったシナリオが最も良かった」といった自由記述が複数みられた. これらから, 提案手法は細かい変化と大きな変化を適切に組み合わせることで, より効率的に嗜好画像を生成できることがわかった. SimpleBOの4次元のアンケート結果も良い評価を得ている部分があるが, 「画像は大抵の場合ほとんど変化せず, 変化した場合も目的と無関係の変化が多かったので, 画像の選択が難しかった」などの自由記述があり, ユーザの関心領域以外が変化してしまっていた可能性がある.

図4に, 「提示される画像は徐々に好む画像に変化しましたか?」に関するアンケート結果を示す. これを見ると, 提案手法がベースラインよりも好意的に評価されており, 嗜好画像に変化した事例が多いことがわかる. 図3, 図4をみると, SimpleBOの4次元の結果は分散が大きいことがわかる. 自由記述を見ると, 「ほとんど画像の変化がなく, 比較が難しかった」, 「想定していた顔に非常に近い顔になったため」といった自由記述がみられ, 人によって受け取る印象に差があることがわかった.

図5に, 各シナリオ後の「毎回の画像の比較は簡単でしたか?」というアンケートに対する結果を示す. これをみると, BanditBOとSimpleBOの4次元の場合の画像の比較が難しかったことがわかる. アンケートには, 「同じ画像が生成され続け, 選びにくい」という自由記述が多く見受けられた. ユーザから有用なフィードバックを得るためには, 類似した画像同士の比較は避けた方が良かった.

6 議論

6.1 生成画像の多様性とユーザの満足度の関係

本節では, ユーザ実験を通じて得られた知見について議論する. 最終アンケートから, 「効率が悪かったシナリオでは, 互いにほぼ関係がない画像が次々と出力され, 比較しても手応えがないように感じた. 効率が良かったシナリオでは, 少ない種類の画像を少しずつ変えていくようになり, 選んだものが反映

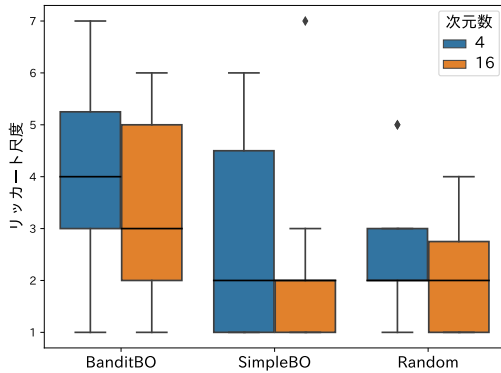


図 3: 「目的の画像まで効率的に辿り着くことができましたか?」に関するアンケート結果

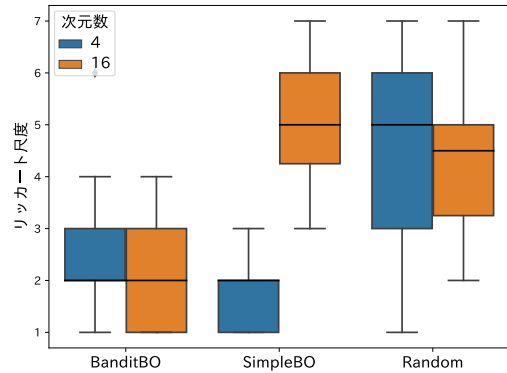


図 5: 「毎回の画像の比較は簡単でしたか?」に関するアンケート結果

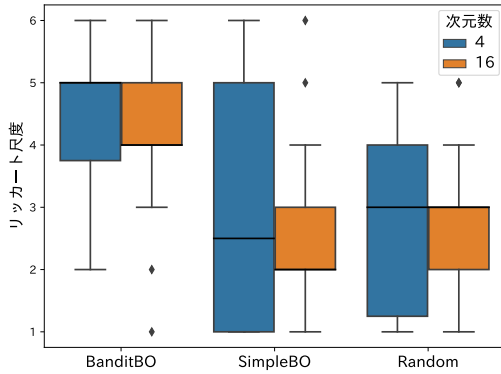


図 4: 「提示される画像は徐々に好む画像に変わりましたか?」に関するアンケート結果

されていくように感じたが、最初に選んだ画像でその後の出力がほとんど決まってしまう、局所的に最適な画像に落ちてしまった気がする」といった自由記述が見られた。提案手法は Random と比べると生成画像の多様性に欠けている。

異なる次元数 d' における生成画像の多様性を確認するために、 d' ごとに、 w' をランダムサンプリングして複数枚の画像を生成した。図 6 は、その結果である。 $d' = 8, 16$ の場合は、多様な画像が生成できていることがわかる。しかし、 $d' = 4$ の場合は、類似した画像が生成されており、多様な画像の生成ができていない。StyleGAN の潜在空間を主成分分析によって次元削減することで、StyleGAN の画像生成の表現力が小さくなっていくことが定量的に確認できた。また、4 次元空間の Random の自由記述で、「今回は過去に見た画像がまた出てくる事が多かった」といった自由記述が見られた。これらから、4 次元空間では 16 次元空間と比べて生成画像の多様性に欠けていることがわかる。しかし、両方の次元に関するアンケートから、「事前のアンケートで答えたイメージに近いアバターが生成できた」や「最

初のイメージとは異なるけど自分が納得する画像を生成できた」といったようなコメントが見受けられた。このことから、次元の違いは生成される画像の多様性と表現力に影響を与えるが、ユーザが満足する画像を生成する能力には直接的な影響を与えない可能性がある。

6.2 画像の比較による嗜好の変化

アンケートの中に、「欲しいアバターの職業によっては性別のイメージが画像生成途中で変わることがよくあった。また、アバターを生成する中で明らかにイメージと異なる画像が生成されることも多かったが、従来と全く異なる系統の画像が表示されることで一種のインスピレーションを受けてアバターの好みが変わったこともある」といった自由記述が多く見受けられた。これにより、人間の嗜好は画像の比較を行うことで変化することがわかった。このような嗜好の変動を捉えるため、提案手法は多腕バンディットアルゴリズムを用いて各次元が現在のユーザの嗜好にどれだけ寄り添っているかを動的に評価しており、ユーザの嗜好の変動に柔軟に対応し、適切な画像を生成することができる。

6.3 提案手法による多様な画像の探索支援

既存研究では、創造的なアイデアが独創性と有用性と関連しているとの議論が行われてきた [12]。ベイズ最適化の中の探索と活用は、それぞれ独創性と有用性という考え方に関連しており、これにより創造性とのつながりについて考察できる可能性があることが既存研究で提案されている [9]。BanditBO に関する自由記述で、「微調整しながら、途中で大幅に違うものと比較しては元の路線に戻ってくることを繰り返す、良い生成をしていた」といった自由記述がいくつか見られた。このことから、ベイズ最適化の探索と活用は、独創性と有用性の考え方に類似しており、提案手法はユーザの多様な画像の探索を

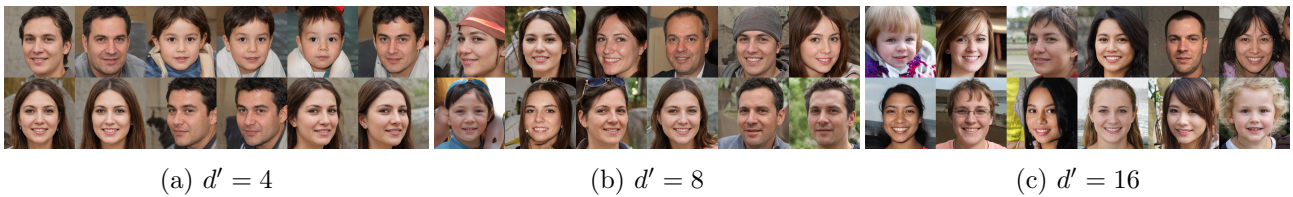


図 6: 部分空間の次元数 d' による表現力の違いの検証

支援できることがわかった。特に、BanditBO は有用性を追求しながらも独創性を失わないように作られている。

7 ユーザビリティ向上のための将来展望

7.1 比較画像の検討

本節では、嗜好画像生成のための人間との適切なインタラクションに関して議論する。アンケートの中で、「自分の選ばなかった方の画像が次の比較対象となると、自分の好む画像から遠ざかっているような感じがして、少し苛立ちを覚えた」や「前回の比較で優位だった方を次の比較に引き継ぐ方が、より比較しやすく良い方法だと感じた」といった自由記述が複数みられた。これらから、ユーザは一対比較を行う際に前回選択した良い方と今回の生成画像の比較をシステムに求めていることがわかった。ただし、本研究ではスマートフォン操作を想定しており、画面の小ささから、Tinder のようなアプリを参考に 1 画面に 1 画像のユーザインタフェースを採用した。この場合、1 つ前の生成画像と比較することが自然だと考えている。スマートフォンで一対比較を行うための最適なユーザインタフェースを考える必要がある。

7.2 画像の変更箇所の可視化

アンケートの中で、「微調整のところでは両者がどのように違うのかがわかりづらいものも多く、もし、特微量ベースでやっているのなら、2 つがどのような点で異なるかを文で教えてくれると助かると思った」という自由記述が見られた。ここから、新しい生成画像と前回の画像の変更箇所の可視化が求められていることがわかった。

7.3 探索と活用のバランスを表示

アンケートの中で、「画像の選択を繰り返すにあたって、大まかに絞り込んでいるのか、詳細な特徴を詰めているのかが提示されると選択中の不安感が減るように感じた」という自由記述が見られた。ユーザの不安感を軽減するために、現在のベイズ最適化の探索と活用のバランスをユーザに示してあげることが有用であることがわかった。ユーザがベイズ最適化の探索と活用のバランスを調整できるようにす

ることも今後の展望になる。

7.4 ユーザ主体のシステム

アンケートの中に、「眼鏡の部分を変えたいと思いながら生成を繰り返していたが、眼鏡以外の部分がどんどん変化しており、なかなか思い通りの画像に近付かなかった」、「選びたくないのに 2 択だから仕方なく選んだ画像によってだいぶ進路がブレたので、選ばないという選択肢があればいいと思った」など、ユーザが思い通りに操作したい内容の自由記述が多く見られた。これらから、ユーザは自分が主体性を持つシステムを好むことがわかった。既存研究で、ベイズ最適化によるデザイン提案システムの中で、人間のデザイナーが完全な主導権を持つ新しい概念が提案されている [9]。しかし、このコンセプトでの実証実験は行われていなく、ユーザから実際にユーザ主体のシステムが求められているかはわからなかった。これに対して、本研究のアンケートの結果から、ユーザ主体のシステム設計が実際に求められていることが明らかとなった。本研究はスマートフォン操作を想定しており、複数のスライド調整などは操作性が低く難しいという制約がある。実際のシステム設計では、この制約を考慮してユーザ主体のシステム設計を検討する必要がある。

8 結論

本研究では、一対比較結果のみを利用する状況下で、提案手法がベースラインより効率的に嗜好画像を生成できるかを検証した。StyleGAN の潜在空間に対して主成分分析を適用して次元を削減し、ユーザの関心次元を重点的に探索することで、ベースラインより効率的に嗜好画像を生成できることを明らかにした。本研究により、スマートフォン上でスワイプ操作を用いて嗜好画像を生成する新しいアプローチを提案することができた。今後は、探索次元が大きい場合においても、少ないイテレーション数で嗜好画像を生成できるように、効率的な探索方法の開発に取り組む。具体的には、画像提示からスワイプまでの経過時間など、ユーザの行動から得られる情報を利用することで、嗜好画像生成までに必要なイテレーション数を下げることを目指す。また、7 節で述べたような、ユーザビリティの高いシステムにすることに取り組む。

謝辞

本研究は、株式会社メルカリ R4D とインクルーシブ工学連携研究機構との共同研究である価値交換工学の成果の一部である。

参考文献

- [1] P. Auer. Using Confidence Bounds for Exploitation-Exploration Trade-Offs. *J. Mach. Learn. Res.*, 3:397–422, 2003.
- [2] T. Chong, I.-C. Shen, I. Sato, and T. Igarashi. Interactive Optimization of Generative Image Modelling using Sequential Subspace Search and Content-based Guidance. *Computer Graphics Forum*, 40:279–292, 2021.
- [3] W. Chu and Z. Ghahramani. Preference Learning with Gaussian Processes. In *Proceedings of the International Conference on Machine Learning*, Vol. 119, pp. 137–144, 2005.
- [4] J. Dudley, J. Jacques, and P. O. Kristensson. Crowdsourcing Interface Feature Design with Bayesian Optimization. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2019.
- [5] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative Adversarial Nets. In *Proceedings of the Advances in Neural Information Processing Systems*, Vol. 27, 2014.
- [6] E. Härkönen, A. Hertzmann, J. Lehtinen, and S. Paris. GANSpace: Discovering Interpretable GAN Controls. In *Proceedings of the Advances in Neural Information Processing Systems*, Vol. 33, pp. 9841–9850, 2020.
- [7] T. Karras, S. Laine, and T. Aila. A Style-Based Generator Architecture for Generative Adversarial Networks. *IEEE transactions on pattern analysis and machine intelligence*, 43(12):4217–4228, 2021.
- [8] M. Khajah, B. Roads, R. Lindsey, Y.-E. Liu, and M. Mozer. Designing Engaging Games Using Bayesian Optimization. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp. 5571–5582, 2016.
- [9] Y. Koyama and M. Goto. BO as Assistant: Using Bayesian Optimization for Asynchronously Generating Design Suggestions. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, pp. 1–14, 2022.
- [10] Y. Koyama, I. Sato, and M. Goto. Sequential Gallery for Interactive Visual Design Optimization. *ACM Trans. Graph.*, 39:1–12, 2020.
- [11] Y. Koyama, I. Sato, D. Sakamoto, and T. Igarashi. Sequential Line Search for Efficient Visual Design Optimization by Crowds. *ACM Trans. Graph.*, 36(4):1–11, 2017.
- [12] A. Kozbelt, R. A. Beghetto, and M. A. Runco. *Theories of Creativity*, pp. 20–47. Cambridge University Press, 2010.
- [13] E. Raponi, H. Wang, M. Bujny, S. Boria, and C. Doerr. High Dimensional Bayesian Optimization Assisted by Principal Component Analysis. In *Proceedings of the International Conference on Parallel Problem Solving from Nature*, pp. 169–183, 2020.
- [14] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer. High-Resolution Image Synthesis With Latent Diffusion Models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10684–10695, 2022.
- [15] F. Schroff, D. Kalenichenko, and J. Philbin. FaceNet: A Unified Embedding for Face Recognition and Clustering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 815–823, 2015.
- [16] Y. Zhou, Y. Koyama, M. Goto, and T. Igarashi. Generative Melody Composition with Human-in-the-Loop Bayesian Optimization. In *Proceedings of the Joint Conference on AI Music Creativity*, 2020.

未来ビジョン

本研究は、スマートフォンのスワイプ操作によって直感的に嗜好画像を生成する新しいフレームワークを提案し、そのようなインタラクションが可能な未来を開拓します。近年、テキストから画像を生成する「text-to-image」の技術が急速に進化し、質の高い画像生成が可能になっています。しかし、この手法では適切なプロンプトを考え出す作業が煩雑であり、ユーザにとっては一定の負荷が伴います。またその性質上、偶発的な発見が限られています。

一方で、近年マッチングアプリの急速な普及によって、スワイプ操作は日常的なインタラクションとなっています。この現状を踏まえ、本研究ではスワイプ操作を通じて気軽かつ効率的に嗜好に合った画像を生成できる方法を開発します。また、予想外の美しい画像に出会え

るような手法も探求します。これにより、複雑なプロンプトを考える必要がなく、手軽に高品質な画像を生成する未来が期待されます。

また、本研究の発展として産業連携を期待します。ゲーム産業では、ユーザが直感的にゲーム内のキャラクターや背景をカスタマイズできるようにになれば、より深いエンゲージメントとユーザ体験が実現します。ファッション・デザイン業界では、ユーザの嗜好に合わせたアパレルやアクセサリのデザイン案を即座に生成することができ、既存の商品開発プロセスを高速化し、多様化することができます。広告業界では、ユーザがスワイプ操作で気に入る画像やスタイルを生成できるという特性を利用して、将来的にはパーソナライズされた広告コンテンツの生成が可能になります。