

映像投影による他者への変容感を最大化する顔画像の最適化

羽山 大智^{*1} 渡辺 義浩^{*1}

Abstract – 近年、自身の顔を他者へと変容させる技術が様々な分野で活用されている。中でも、現実世界の顔をメイクにより他者へと変容させる手法は、特殊メイクやものまねメイクと呼ばれ、実在感と変容感を両立している。一方、映像投影により実際の顔を他者へと変容させる手法に関しては研究されており、メイクの工程を必要とせずに変容を達成している。しかし、同手法はメイクを用いた手法と比較して、色補償の不十分さや変容目標との相性により、他者への変容感が十分でない問題がある。そこで、本稿では映像投影による他者への変容感を最大化するため、投影特有の条件に基づき、StyleGAN2 の潜在空間内で変容先の顔を最適化する手法を提案する。また、本手法により生成される画像を実際に投影し、評価を行った。

Keywords : バーチャルメイク、匿名化、Facial Projection Mapping

1 はじめに

自身の顔を他者へと変容させる技術は映像制作やコミュニケーションツール、セキュリティなど様々な分野で応用されている。そのような中で、現実世界では、メイクを用いることにより、他者への変容を実現している。同手法は特殊メイクやものまねメイクと呼ばれ、メイクを用いた肌の色や立体感などの見た目の操作を行うことにより、自身のアイデンティティから離れ、他者へと近づけている。このような手法は、特殊な技能が必要であるほか、目標とする他者を設定し、注意深い観察と長い準備時間が必要となる。

一方、顔への映像投影により、見た目を変容させる Facial Projection Mapping がある [1, 2, 3]。中でも、同手法を用いてメイクを実現する研究がされている [1]。同手法は、実際のメイクをする必要がなく、瞬時に顔の見た目の変化をもたらすことが可能である。

また、ものまねメイクと同様に、特定の他者への変容を目的とした Facial Projection Mapping が提案されている。同手法は Face Swapping と呼ばれる画像生成における顔の交換手法に対して、投影対象の顔の形状を保持する制約を含めることで変容画像を求め、投影を行っている [4]。しかし、変容目標の肌の色などテクスチャ情報を直接埋め込んだ画像を生成するため、生成される顔画像によっては皮膚上で再現が難しい色となる場合がある。

同手法では、この問題を解決するため、投影目標の顔画像から実際に投影される色補償画像への計算を行う際に、投影目標の明るさを手動で調整し、可能な限り皮膚上で再現可能な色の範囲に収めている。一方、自然な明るさで調整する場合、変容目標によっては多くの画素が補償可能範囲に収まらない問題が生じる。

また、同手法は、特定の変容目標を決定し、他者への変容を行う構造のため、変容目標次第では形状が目標から大きく変わり、得られる変容感が弱い問題がある。

そこで本稿では、映像投影による他者への変容感を向上させることを目的とする。これに向けて、投影特有の条件に合わせた画像生成を行う必要がある。ここで、投影特有の条件とは、先述の投影対象の形状に合わせた画像を生成する必要があることや、投影対象の皮膚の反射率やカメラ・プロジェクタの特性により、顔上で再現可能な色が制限されていることを示す。本稿では、同条件に応じた投影画像を求めるため、StyleGAN2 の潜在空間における潜在変数の探索に投影特有の条件を損失として与え、変容目標として生成される顔画像の最適化を行う。

2 関連研究

2.1 Facial Projection Mapping

Bermano らは、顔に対して動的に追従して投影を行い、顔の見た目を変化させる手法を提案した [2]。また、同手法を応用し、画像生成モデルによる投影対象が老いた顔画像とフィードバック型の色補償を組み合わせ、年齢感を操作する手法が提案されている [3]。一方、同手法は他者への変容に対して焦点を当てていない。

そのような中で、羽山らは顔への映像投影により、特定の他者への変容を行う手法を提案した [4]。同手法は、Editing for Swapping と呼ばれる顔交換画像生成手法とメッシュ変形による位置補正を組み合わせ、投影対象の形状に合わせつつ変容目標の特徴を持つ顔画像を生成している。また、同手法により、投影対象の顔が変容目標に近づいていることを定性的に確認している。一方、先述の通り、同手法で生成される顔画像が皮膚上で色補償可能な範囲内に収まらない場合が

^{*1}東京科学大学

あり、手動で明るさを調整している。しかし、投影可能な範囲に収めるため、投影目標の明るさを大きく下げた場合はコントラストや顔の明るさの低下により不自然になる一方、明るさを十分に下げない場合は色補償が難しいトレードオフの関係がある。このため、明るさが自然かつ他者へと顔が変容して見えるよう微調整が必要となる。また、同手法は、各顔パーツの配置が固定されており、変容目標次第では似せることができず、投影対象の顔の特徴が大きく残る問題もある。

本稿では、StyleGAN2 の潜在空間内で、投影特有の条件を踏まえながら、顔類似度指標を用いてアイデンティティが投影対象から離れるように顔画像を最適化することで変容目標を求める。

2.2 顔画像生成

高品質な顔画像生成として用いられる手法は大別して 2 つあり、GAN ベースの手法と Diffusion Model ベースの手法がある。前者の GAN ベースの手法は StyleGAN など、顔画像生成において潜在空間の探索が人間に解釈しやすく直感的である一方、後者の Diffusion Model ベースの手法はテキストプロンプトでの操作がしやすく、また生成される画像の幅が広い利点がある。本稿において、我々は投影条件をベースに最適化を行うため、様々な損失による操作性が高い方が好ましい。そのため、GAN ベースの手法を採用する。

中でも、FALCO と呼ばれる StyleGAN ベースの顔画像の匿名化手法がある [5]。同手法は、与えた顔画像に対して、StyleGAN2 の潜在空間内で、顔の属性を保ちながら匿名化することが可能な顔画像を探索する。具体的には、FaRL を用いて偽の顔画像とペアリングを行い、アイデンティティに関連する 3 - 7 層の潜在変数のみを投影対象の潜在変数と交換し、顔の属性保持損失と顔類似度指標に基づく匿名化損失を用いることで、潜在変数の最適化を行う。これにより、属性を保持しつつ匿名化を達成した。

一方、StyleCLIP は StyleGAN ベースのモデルの潜在空間内で CLIP と呼ばれるテキストと画像のペアを学習した指標を基に潜在変数を最適化することで、意図した属性への画像の編集を行っている [6]。

このように、何らかの指標を用いた潜在変数の最適化により所望の顔画像を求める手法がある。これを踏まえると、投影に適した他者へと変容可能な顔画像も潜在変数の最適化によって求まる期待される。

2.3 色補償

意図した色を投影対象の表面で再現するために、投影目標から実際に投影される画像を計算する手法は色補償と呼ばれており、様々な提案が報告されている。中でも、Fujii らは、カメラとプロジェクタの色の対応を線形でモデル化し、環境光を独立させた式を提案し

た [7]。同手法は、簡潔な式で表現されており、比較的計算量が小さいため、しばしば用いられている。

また、近年では意図した画像や画風を投影面上で表現するために、画像生成モデル内で投影シミュレーションを行い、予め評価することで、投影画像を最適化する手法が提案されている [8, 9]。代表例として、Deng らはテキストプロンプトに沿った画風に近い投影を行うための手法である LAPIG を提案した [8]。同手法は、3 つのネットワークによって生成される 3 枚の画像を用いて求めた損失により、最初の変換画像を再度調整し、テキストに沿いながらも投影による再現が可能な画像を求める。

加えて、中村らは StyleTransfer に投影シミュレーションを含めることで、画風を投影対象に転送しつつ、生成される画像を色補償可能な範囲に収める手法を提案した [9]。これにより、画像の加算を用いた画風の転送を達成している。

一方、映像投影による他者への変容を目的として設計したものはない。これらの手法と同様に色補償可能性に基づく他者へと変容可能な投影画像を計算することで、投影時の変容感を最大化することが可能であると予想される。そのため、本稿では色補償のシミュレーションを最適化に含め、顔画像を生成する。

3 潜在変数の探索による変容目標顔画像の最適化

先述の通り、本稿では投影による他者への変容感の向上を目的とする。この目的の達成に向けて、色補償が達成可能であり、かつ、投影対象の形状を維持しつつ投影対象とは異なるアイデンティティを持つ顔を生成する必要がある。そこで本稿では、匿名化損失、セグメンテーション損失、色補償損失、明度損失、知覚類似度損失、CLIP 損失を用いた StyleGAN2 の潜在空間内における潜在変数の探索により、変容目標となる顔画像の最適化を行うことを提案する。同手法による投影までのパイプラインを図 1 に示す。

3.1 投影特有の条件に応じた潜在変数の探索

先述の投影特有の条件に応じた顔画像生成のため、白色投影下で撮像した顔画像を StyleGAN2 の潜在空間における潜在変数へと変換し、顔画像を最適化するために潜在変数の探索をする。本稿では、Barattin らが提案した FALCO [5] とは異なり、顔の表情や髪などの属性を保持する必要はなく、他者へとアイデンティティが変わることが重要である。そこで、特定の属性以外の属性保持はせずに最適化を行う。

同様にして、FALCO では属性保持のためにアイデンティティに関わる 3 - 7 層のみを最適化していたが、本稿ではアイデンティティだけでなく色や形状も柔軟に操作可能である方が好ましい。そのため、本稿で

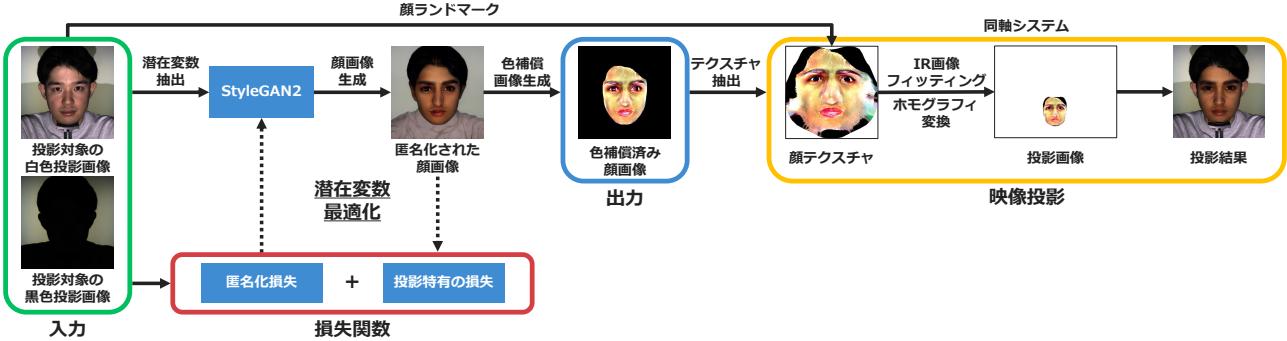


図 1 提案手法のパイプライン。
Fig. 1 Proposal method pipeline.

は探索対象となる層を拡張し, 0 - 12 層を最適化する。また, FALCO は属性保持のため, ペアリングした偽の顔画像の潜在変数を最適化の初期値としているが, 本稿では投影対象の顔画像または任意の顔画像から最適化を行う。ここで, 最適化される潜在変数は $\mathbf{w}_a^+ := \{w_a^i\}_{i=0}^{17}$ のうち $i = 0, 1, \dots, 12$ である。加えて, 式 (1) により顔画像 x_a が生成される。

$$x_a = \text{clamp}(\mathcal{G}(\mathbf{w}_a^+), 0, 255) \quad (1)$$

ここで, \mathcal{G} は StyleGAN2 のジェネレータである。

3.2 損失関数

投影特有の条件を踏まえた他者への変容を最大化する損失関数について説明する。Barattin らは, データセットの匿名化のため, 潜在変数を匿名化損失と属性保持損失を用いて最適化した [5]。しかし, 本稿では投影対象の全ての属性を保持する必要はない。一方, 投影対象の形状を保持した顔画像を生成する必要があるほか, 色補償が可能な範囲内に生成される顔画像を収める必要がある。以下で各損失について説明する。

まず, 他者の顔へと変容させる匿名化損失は FALCO と同様に, 以下の式を採用する。

$$\mathcal{L}_{\text{id}} = |\cos(\mathcal{E}_A(x_a), \mathcal{E}_A(x_p)) - m| \quad (2)$$

ここで, x_a は最適化中の顔画像, x_p は投影対象の顔画像であり, \mathcal{E}_A は ArcFace エンコーダ [10], m は匿名化強度を制御するハイパーパラメータである。 $m = 0$ のとき, 生成される顔画像は投影対象に対して特徴が直交するように生成される。

続いて, 投影対象と同じ形状を持つ顔を生成するため, セグメンテーション損失を導入する。セグメンテーションの検出は BiSeNet ベースのコードを用いる [11]。Soft mIoU を参考に以下の式で損失を定義する。

$$\mathcal{L}_{\text{seg}} = 1 - \frac{1}{|\mathcal{S}|} \sum_{c \in \mathcal{S}} \frac{\sum_{i \in \Omega} p_{i,c} q_{i,c}}{\sum_{i \in \Omega} (p_{i,c} + q_{i,c} - p_{i,c} q_{i,c}) + \epsilon} \quad (3)$$

ここで, \mathcal{S} はセグメンテーションされる顔パーツのラベル統合後の集合, Ω は顔画像の画素の集合, $p_{i,c}$, $q_{i,c}$ は最適化中の顔画像と投影対象の i 番目の画素がラベル c である確率, ϵ はゼロ除算を防ぐための微小な定数値である。なお, 本稿では色補償を行った投影をするため, 顔のパーツの配置をある程度変更することが可能であると予想される。そこで, 一部ラベルに関して統合することにより, 生成される顔のパーツの配置に自由度を与える, アイデンティティの変化を達成しやすくする。そのため, セグメンテーションラベルは, $\{[\text{背景}, \text{左耳}, \text{右耳}, \text{イヤリング}, \text{首}, \text{ネックレス}, \text{服}, \text{髪}, \text{帽子}], [\text{顔の皮膚}, \text{左眉}, \text{右眉}], [\text{左目}, \text{右目}], [\text{眼鏡}], [\text{鼻}], [\text{口内}], [\text{上唇}], [\text{下唇}]\} \in \mathcal{S}$ で統合する。

次に, 投影時の条件に関する損失について説明する。羽山らの手法 [4] と同様に, 色補償には式 (4) に示すフィードフォワード型の色補償である Fujii らの式 [7] を用いる。この式から, 内部的にシミュレーションを行い, 色補償達成可能な顔画像を生成する。

$$\mathbf{C}_i = \mathbf{A}_i(\mathbf{F}_i + \mathbf{V}\mathbf{P}_i) \quad (4)$$

ここで, \mathbf{C}_i , \mathbf{A}_i , \mathbf{P}_i , \mathbf{F}_i はそれぞれカメラの i 番目の画素値, 反射率行列, 色補償画像の画素値, 環境光, \mathbf{V} は全画素共通の色混合行列である。なお, \mathbf{V} は反射率が既知の物体から事前に求まり, \mathbf{A}_i と \mathbf{F}_i は, 白色投影と黒色投影時の顔画像から事前に求まる。

これを基に, LAPIG [8] から着想を得て, 色補償損失を以下に示す。以下の式は, 色補償画像がプロジェクタの投影可能な範囲を超えることを抑制するため, 0 と 255 の範囲外の画素値を基に損失を計算する。

$$\begin{aligned} \mathcal{L}_{\text{color}} &= \frac{1}{3|\Omega|} \sum_{i \in \Omega} \sum_{j=0}^2 \left(\frac{b_i(o_{i,j} + u_{i,j})}{255} \right), \\ o_{i,j} &= \text{ReLU}(\mathbf{P}_{i,j} - 255), \\ u_{i,j} &= \text{ReLU}(-\mathbf{P}_{i,j}), \\ \mathbf{P}_i &= \mathbf{V}^{-1}(\mathbf{A}_i^{-1}\mathbf{C}_i - \mathbf{F}_i) \end{aligned} \quad (5)$$

上記の式において, b_i は投影対象の顔内部の領域のみを 1 としたバイナリマスクの i 番目の画素の値である。また, $o_{i,j}$ は式 (5) から求まる色補償画像 \mathbf{P} の i 番目, j チャンネルの画素値で計算される 255 を超える値であり, $u_{i,j}$ は 0 を下回る値である。

また, 色補償損失のみによる操作では, 補償可能範囲内に収める方向に最適化が進み, 生成される画像の明るさが保証されず, コントラストや顔の全体的な明るさの低下が起こる。そのため, 投影対象の平均明度による拘束を導入し, 生成される画像の明るさ低下を抑制する。以下の式で, 明度損失を定義する。

$$\mathcal{L}_{\text{value}} = \left| \frac{1}{|\Omega|} \sum_{i \in \Omega} \frac{b_i (\text{Val}(x_{a,i}) - \text{Val}(x_{p,i}))}{255} \right| \quad (6)$$

上記の式の $x_{a,i}$ は最適化中の顔画像における i 番目の画素であり, $x_{p,i}$ は投影対象の顔画像における i 番目の画素である。加えて, Val は RGB 値から明度へと変換する関数である。

また, 上記の色補償損失のみでは単純な画素の計算となり, 知覚的に重要な箇所が投影によって再現されない可能性がある。そこで, 投影シミュレーションを内部で行い, 投影後の見えを計算する。この計算を基に, 投影結果が最適化された顔画像に近づくように LPIPS による知覚類似度損失 $\mathcal{L}_{\text{LPIPS}}$ を導入する [12]。投影シミュレーション画像は, Fujii らの計算式によって求めた色補償画像を基に, 顔の内部領域のみ顔への再投影を行い, その他の領域は投影対象の顔画像の画素値を合成することで得る。Fujii らの式による再投影結果を直接用いない理由は, 背景領域など顔の外において反射率計算のエラーが画像上に多く含まれるためである。以下に合成式を示す。

$$x_{sim,i} = b_i x_{reproj,i} + (1 - b_i) x_{p,i} \quad (7)$$

ここで, $x_{sim,i}$ が投影シミュレーション画像の i 番目の画素, $x_{reproj,i}$ が Fujii らの式 (4) を用いて計算される再投影画像の i 番目の画素であり, $x_{reproj,i} = \mathbf{C}_i$ となる。また, b_i が投影対象の顔内部の領域のみを 1 としたバイナリマスクである。上記によって得られた投影シミュレーション画像と最適化中の顔画像を LPIPS に適用することで知覚類似度損失を求める。

最後に, 他者へと変容させる際に性別や顔らしさなど特定の属性を保持し, 変容を制御しやすくするため CLIP を用いた損失を導入する。同損失は, StyleCLIP [6] にて定義されたテキスト特徴と画像特徴のコサイン類似度による損失 $\mathcal{L}_{\text{CLIP}}$ を利用する。

以上より, 顔画像の最適化を行う際の最終的な損失関数は $\mathcal{L}_{\text{total}}$ となる。 $\mathcal{L}_{\text{total}}$ の具体的な式を以下に

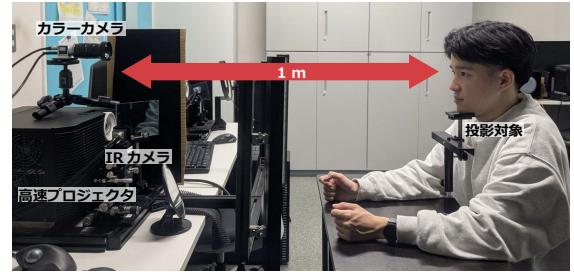


図 2 実験環境.
Fig. 2 Experimental environment.

示す。

$$\begin{aligned} \mathcal{L}_{\text{total}} = & \lambda_{\text{id}} \mathcal{L}_{\text{id}} + \lambda_{\text{seg}} \mathcal{L}_{\text{seg}} + \lambda_{\text{color}} \mathcal{L}_{\text{color}} \\ & + \lambda_{\text{value}} \mathcal{L}_{\text{value}} + \lambda_{\text{LPIPS}} \mathcal{L}_{\text{LPIPS}} \\ & + \lambda_{\text{CLIP}} \mathcal{L}_{\text{CLIP}} \end{aligned} \quad (8)$$

ここで, λ は各損失に適用する重みを示す。

3.3 顔画像の抽出・投影

次に, 最適化によって取得した顔画像を基に, 色補償済み画像の顔の内部領域のみを抜き出した $\mathbf{P}_{mask,i} := b_i \mathbf{P}_i$ を計算する。色補償画像の計算については先述の式 (5) を用いる。ここで, 投影対象の顔画像を撮像するカラーカメラとプロジェクタが同軸でなく, かつ投影対象が立体であるため, ホモグラフィでの記述は難しい。そのため, プロジェクタと同軸な IR カメラを介して実際の投影画像を求める。

具体的には, まず色補償計算後の顔画像 \mathbf{P}_{mask} に対して, 白色投影時の投影対象のランドマークから推定した顔メッシュを基にテクスチャ抽出を行う。次に, 抽出されたテクスチャを, プロジェクタと同軸な IR カメラで取得した顔画像から推定した顔メッシュに適用し, ホモグラフィ変換を行うことで投影画像を求める。

4 最適化による顔画像生成と実際の顔への投影による他者への変容効果の検証

4.1 概要

提案手法による顔画像の最適化により, 投影特有の条件を満たしつつ顔が他者へと変容しているか定性的に確認する。具体的には, 投影対象から最適化した際の顔画像とシミュレーションの検証, および実際の顔への投影を行った。

実験環境を図 2 に示す。プロジェクタと投影対象の距離は約 1 m とした。また, IR カメラとプロジェクタのホモグラフィ行列を求めた後, 標準白色板を投影対象の位置に設置し, カラーカメラとプロジェクタの応答関係を計算することで, 投影画像生成に必要なパラメータである \mathbf{V} を事前に取得した。

投影には 800 fps に設定した 1024×768 px の力

ラープロジェクタ [13]、撮像には 40 fps に設定した 1920×1200 px のカラーカメラとプロジェクタに同軸な 500 fps, 720×540 px の IR カメラを用いた。さらに、投影対象は顎台を用いて静止させた真顔とした。また、肌の鏡面反射を抑制するためフェイスパウダーおよびファンデーションを皮膚表面に少量塗布した。

各損失に与える重みは、 $\lambda_{\text{id}} = 0.8$, $\lambda_{\text{seg}} = 4.5$, $\lambda_{\text{color}} = 1.0$, $\lambda_{\text{value}} = 1.4$, $\lambda_{\text{LPIPS}} = 1.0$, $\lambda_{\text{CLIP}} = 1.0$ とした。加えて、CLIP 損失のためのテキストプロンプトには “a face” と “a male face” を用いた。さらに、潜在変数の最適化を行う 0 - 12 層の初期値に投影対象の顔画像と StyleGAN2-ADA [14] で生成した顔画像の潜在変数の 2 種類を用いた。なお最適化を行わない層に関しては、投影対象の顔画像から抽出した潜在変数を用いた。また、最適化は 300 steps 行った。

さらに、顔テクスチャ抽出時に欠落した画素は OpenCV の inpaint を用いて補完した。加えて、匿名化強度を制御するパラメータは $m = 0$ とした。

4.2 最適化画像とシミュレーション結果

投影対象の顔から直接最適化を行った実験結果を図 3 に示す。図 3 の最適化画像から、提案手法により投影対象の形状を保ちながら他者へと変容した最適化画像が生成されたことを確認できる。また、色補償画像を確認すると画素が飽和している箇所が少なく、最適化により生成される顔が概ね色補償可能範囲内に収まっている。さらに投影シミュレーションを見ると、本手法によって他者へと変容していることが定性的に確認できる。また、最適化により生成される顔画像のコントラストが保たれている。さらに、図 3 の上下段の比較により、CLIP による性別の保持の成功がわかる。

また、他者の顔を初期値として最適化を行った結果を図 4 に示す。この結果から、初期値により生成される顔画像が異なり、初期値の特徴を僅かに引き継いでいることが確認された。加えて、最適化された顔画像を見ると一部の顔画像で細かなテクスチャの崩れがわかる。また、どちらの結果も投影シミュレーションにおいて鼻の先端付近におけるずれがある。

4.3 投影結果

図 3 と図 4 の右列に投影結果を示す。どちらの初期値を用いた場合でも、他者へと変容していることが確認できる。また、概ねシミュレーションに近い投影を行えている。加えて、ラベルの統合を行った眉の形状に関しては、投影対象から形状が変化している。

一方、シミュレーションと比較し、鼻付近の僅かなずれが鼻内部に確認された。また、どの投影結果においても、投影対象の鼻付近のほくろは投影後も確認された。さらに、カメラで撮像された画像と実際に直接知覚する色の間には若干の相違があった。

5 考察

前節の結果から、提案手法により投影特有の条件に応じつつ他者へと変容可能な顔画像が生成されることを確認した。また、提案手法による投影を行うことで他者への変容が十分に達成されていると考えられる。さらに、先行研究 [4] と比較し、変容時に十分な明るさが保たれていることや、最適化を行った顔にある程度類似しており、一部の顔パーツの配置が変化していることから、他者への変容感の向上を達成したと思われる。一方、初期値の特徴を最適化画像が引き継いでいることを踏まえると、最適化中の潜在変数の動かし方が局所的であり、最適化アルゴリズムのパラメータ調整次第ではより大きな変容を達成できると考えられる。

また、鼻付近のずれが確認されたが、これはセグメンテーション損失に与える投影対象の確率マップが one-hot の形式でなかったことや画像にないラベルによるバイアスが計算に含まれていること、最適化回数が不十分であったこと、同損失による位置合わせの限界が要因と考えられる。一方、投影時に同現象が明確には見られなかった理由は、位置のずれが僅かであることや、皮膚の表面下散乱により投影画像がぼけたためであると予想される。加えて、鼻付近のほくろが投影結果から確認された。これは、最適化中に顔が僅かに動いたことや顔メッシュの推定誤差による投影ずれと、色補償範囲外であったことが原因と考えられる。

また、先行研究 [4] の生成画像で確認された皺は本手法で現れなかった。投影時の皺の付与により、アイデンティティが変化して見える例があるため、最適化を改良し、細かな特徴を持つ顔画像を生成することで、より他者へと変容して知覚される可能性がある。

6 まとめ

本稿では、投影特有の条件を踏まえた他者へと変容可能な画像生成のための顔画像最適化について提案した。本手法を適用することで、映像投影による他者への変容感と自然さが向上したことを定性的に確認した。

一方、生成される顔画像のテクスチャには未だ改良の余地があるほか、NVIDIA GeForce RTX 5090 を使用した最適化に 60 秒程度要するため高速化が求められる。また、匿名化損失や色補償の計算に用いるモデルに関して、より人間の知覚に沿った手法を採用することで、他者への変容感の向上が可能であると思われる。さらに、より高周波成分の含む顔を生成することで、さらなる変容感を得られると考えられる。

参考文献

- [1] Nao Tsurumi, et al. Rediscovering Your Own Beauty Through a Highly Realistic 3D Digital

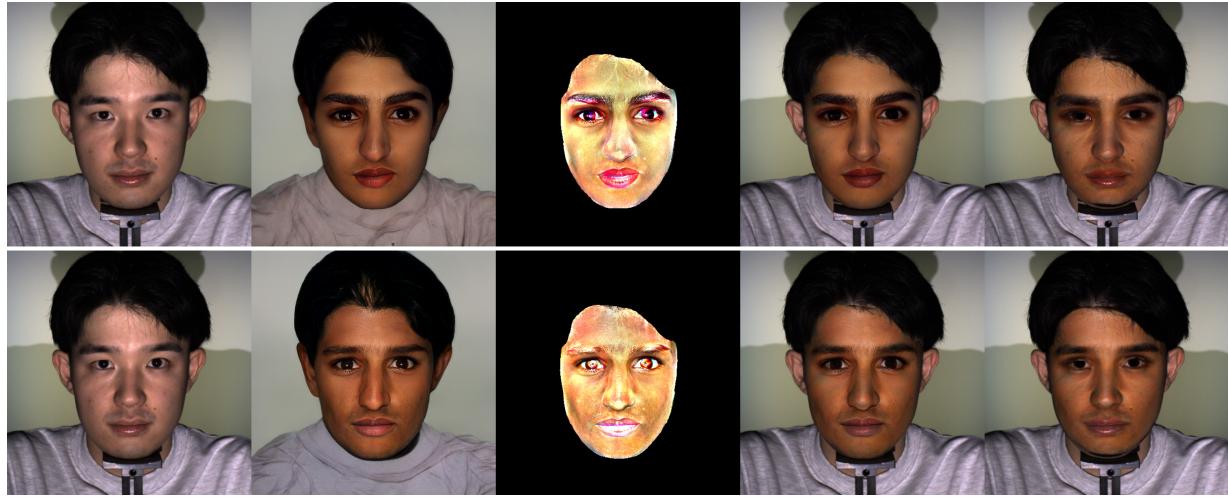


図3 自身の顔から最適化した結果。左から投影対象、最適化画像、色補償画像、投影シミュレーション、実投影結果である。また上段のテキストプロンプトは“a face”であり、下段が“a male face”である。

Fig. 3 Optimized results from projection target face. From left to right: the projection target image, the optimized image, the color compensation image, the projection simulation image, and the actual projection result. In the top row, text prompt is “a face”, and in the bottom row, text prompt is “a male face”.



図4 他者の顔から最適化した結果。左から投影対象、潜在変数の初期値を抽出する顔画像、最適化画像、投影シミュレーション、実投影結果である。またテキストプロンプトは“a face”である。

Fig. 4 Optimized results from other people’s faces. From left to right: the projection target image, the image of the initial latent code, the optimized image, the projection simulation image, and the actual projection result. Text prompt is “a face”.

Makeup System Based on Projection Mapping Technology. In *IFSCC*, 2023.

- [2] Amit H. Bermano, et al. Makeup Lamps: Live Augmentation of Human Faces via Projection. *CGF*, Vol. 36, No. 2, pp. 311–323, 2017.
- [3] 袁 璐ほか. 色補償を用いたプロジェクションマッピングによる顔の年齢変容に関する検討. 第28回日本バーチャルリアリティ学会大会, 2023.
- [4] 羽山 大智, 渡辺 義浩. 映像投影による他者の顔への

視覚的変容に向けた画像生成手法の検討. 第30回日本バーチャルリアリティ学会大会, 2025.

- [5] Simone Barattin, et al. Attribute-Preserving Face Dataset Anonymization via Latent Code Optimization. In *CVPR*, 2023.
- [6] Or Patashnik, et al. StyleCLIP: Text-Driven Manipulation of StyleGAN Imagery. In *ICCV*, 2021.
- [7] Kensaku Fujii, et al. A Projector-Camera System with Real-Time Photometric Adaptation for Dynamic Environments. In *CVPR*, 2005.
- [8] Yuchen Deng, et al. LAPIG: Language Guided Projector Image Generation with Surface Adaptation and Stylization. *TVCG*, Vol. 31, No. 5, pp. 2515–2524, 5 2025.
- [9] 中村 天翼ほか. 実シーンの外観操作のためのニューラルスタイル変換による投影画像生成. 第30回日本バーチャルリアリティ学会大会, 2025.
- [10] Jiankang Deng, et al. ArcFace: Additive Angular Margin Loss for Deep Face Recognition. In *CVPR*, 2019.
- [11] zllrunning. *face-parsing.PyTorch*: Using modified BiSeNet for face parsing in PyTorch. <https://github.com/zllrunning/face-parsing.PyTorch>, 2019. GitHub repository.
- [12] Richard Zhang, et al. The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. In *CVPR*, 2018.
- [13] Yoshihiro Watanabe and Masatoshi Ishikawa. High-Speed and High-Brightness Color Single-Chip DLP Projector Using High-Power LED-Based Light Sources. In *IDW*, 2019.
- [14] Tero Karras, et al. Training Generative Adversarial Networks with Limited Data. In *NIPS*, 2020.