

肖像画からの写実的な顔画像生成手法

中村 優文† 山口 周悟† 福里 司† 森島 繁生‡

† 早稲田大学

‡ 早稲田大学理工学術院

E-mail: † m-nakamu@ruri.waseda.jp, ‡ shigeo@waseda.jp

1 はじめに

本稿では、入力 of 肖像画の陰影や形状は保持しつつ、実写の肌テクスチャと色味を参照することで、肖像画の実写化を行った。肖像画は顔の詳細な特徴を視覚的に伝える手段として用いられ、主に人物の顔を記録、または顔の印象を視覚化するために描かれている。肖像画には大きく二種類ある。モデルや個人の特徴の情報に対して忠実に描かれた肖像画と、モデルの特徴的な箇所や印象が誇張して描かれた肖像画（似顔絵）の二つである。これら二種類は状況によって使い分けられているが、本研究では前者の肖像画に注目した。モデルや個人の特徴の情報に対して忠実に描かれた肖像画の特徴は顔の陰影や形状は写実的であるのに対し、肌の質感（テクスチャ）や色味は油彩・水彩・鉛筆といった描画材料での表現により非写実的な点である。この肖像画が具体的に使用される場面の例として、犯罪捜査等の人物検索、歴史的偉人の顔の記録などがあげられる。人物検索においては対象の人物の印象や記憶から肖像画が描かれ、それをもとに検索をする。しかしながら、肖像画では人のリアルな肌のテクスチャや自然な色味を表現するのに限界があり、肖像画から実際の人物を特定することは困難である。また、同様の理由から歴史人物の肖像画をもとに、本人の印象を推定することも依然として困難である。

そこで、本手法によって、入力 of 肖像画の陰影や形状は保持しつつ、肌のテクスチャと色味を実写真で表現することで、肖像画から実際の人物を特定する補助になり、人物検索への貢献や、歴史人物の顔情報の付与による歴史学・考古学への貢献が期待される。

2 関連研究

2.1 肖像画から写実的な顔画像を生成する研究

肖像画から写実的な画像を生成する既存手法として、実写画像のデータベースや統計モデルを利用した手法がこれまでに提案されてきた。その主な理由として、肖像画と実写の顔画像は色味や肌質といった見た目の特徴が大きく異なるため、事前に特徴量ベースの対応関係を取得する必要があることが挙げられる。Sucontphunt ら [1] は人種、肌色、性別ごとに実写顔画像から得られた顔特徴点に主成分分析を行うことで、ユーザが描いたスケッチ画から得られ

た顔特徴点から実写顔画像を生成する対話的なシステムを提案した。しかし、この手法は疎な顔特徴点のみを用いるため、肖像画のような細部まで描き込まれた形状や陰影の情報を反映することが困難である。そこで、Wu[2] らは実写画像のデータベースと肖像画の顔の部位（目、鼻、口など）の形状と色味の類似度に着目し、類似度の高い実写パーツを選択する合成手法を提案した。また、部位毎の対応関係を用いる手法は実写の写真から肖像画（スケッチ画）を生成するような応用も検討されている [3, 4]、しかし、これらの手法は顔の部位単位での位置・形状・色味の調整に過ぎず、生成結果がデータベースに大きく依存する課題があった。また、パッチ単位での肖像画とデータベースの顔画像の対応付け手法として Klare ら [5] の手法がある。この手法では、同じ作者の描いた肖像画とモデルとなった顔写真を、パッチ単位で特徴量を計算・対応付けを行い、事前に学習データを構築することによって、学習データと同じ作者の描き入力画像と最も類似しているデータベース上の写真を提示している。この手法により、肖像画に描かれる細部の情報を考慮した対応関係を取得することができるが、肖像画の作風やユーザごとに学習データを大量に用意する必要がある。そこで、溝川ら [6] は肖像画とデータベースの顔画像を矩形領域（以下、パッチ）単位に分割し、陰影情報に基づく対応付けを行うことで、実写画像の生成を実現した。この手法では肖像画とデータベースのパッチ単位の対応関係を取得するために、事前に用意した平均顔形状（二次元のメッシュモデル）にフィットするように顔画像を変形させる必要がある（正規化処理）。しかし、この処理によって、1) 変形処理に伴い、局所的な陰影情報を再現できない点と 2) 斜め顔のように三次元的な動きを含む画像への適用が困難であることが課題として挙げられる。肖像画の多くは、斜め顔で描かれることが多いことから、肖像画を実写真で表現するには、斜め顔にも対応可能な手法を検討する必要がある。

2.2 色味及びテクスチャ情報を転写する研究

白黒画像に色味 (RGB) 情報を付与する手法として、Levin ら [7] はユーザが入力した色味と白黒画像の勾配情報を基づく最適化手法を考案した。この手法は形状や陰影情報の保持が容易である一方、あくまでも色味の変換に過ぎないため、肖像画に描かれたテクスチャ情報を実写の顔画像

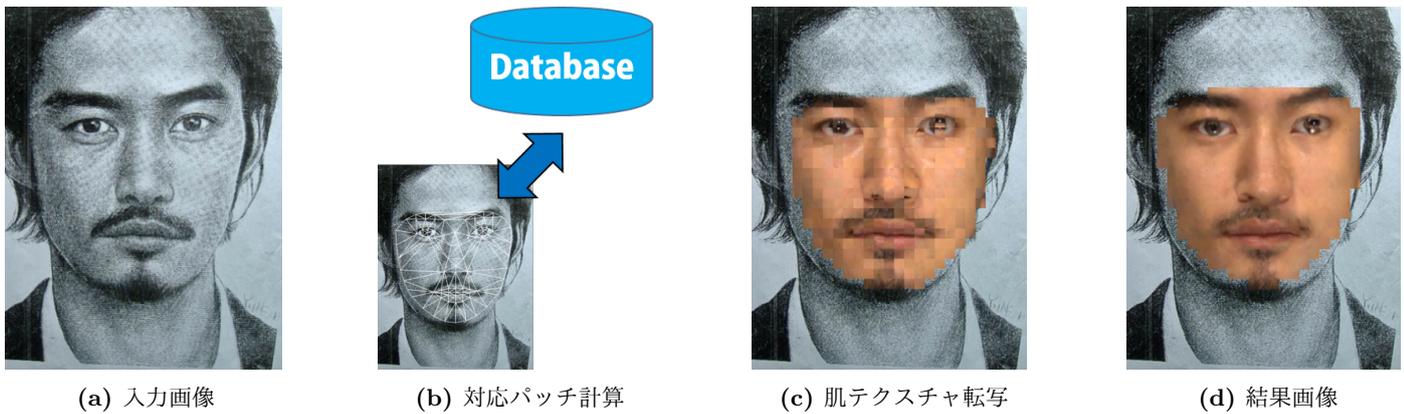


図 1: 本手法のワークフロー

のようなテクスチャに変換することは困難である。また色味に関しても、ユーザの主観に大きく依存してしまう。また、テクスチャ情報に着目し入力画像に異なる画風や作風を転写する研究がこれまでに複数提案されている。Kwatraら [8] は二枚の画像間を密なパッチ単位の対応関係を取得することで、テクスチャ情報の転写を実現している。しかしながら、これらの手法では本稿で挙げるような「肖像画」と「実写顔画像」のように見た目の特徴が大きく異なる画像の場合、パッチ単位の対応関係が十分に取得できず、テクスチャ情報が破綻してしまう点や、出力結果がぼけてしまう点が課題として挙げられる。

3 提案手法

本稿では、肖像画から実写の顔画像を生成する手法として Mohammed ら [9] や溝川ら [6] と同様、パッチ単位の画像合成手法を用いる。Mohammed らの手法はデータベースの顔画像をパッチに分割し再配置することによる、新たな顔画像を生成するものである。モンタージュ写真のように顔のパーツ単位で合成した場合、生成結果がデータベースに大きく依存するが、パッチ単位で合成を行うことで、データベース中の顔画像と見た目の異なる画像の生成が可能となる。Mohammed の手法は色味成分を直接用いるが、本稿では、入力とする肖像画は絵の具などの実写とは大きく異なる染料・色味で描かれるため、Mohammed らと同様の対応付けルールを適用することはできない。そこで、我々はパッチの類似度計算のために、輝度値の分布を考慮する制約を定めた。また、入力画像（肖像画）とデータベース上の画像を直接パッチ単位に分割した場合、顔の各部位の位置が異なるため、正しいパッチ探索を行うことが困難である。そこで、入力画像とデータベースの顔画像の顔形状の対応関係を取得する必要がある。溝川らは変形処理を行うことでこの問題を解決したが、2.1 章で述べたように、いくつか問題点が生じる。そこで、本研究では変形処理は

行わず、顔画像を特徴点に基づく二次元メッシュで分割し、対応するメッシュ上で対応座標を計算する。既存手法の問題点についてまとめると以下の通りになる。

- 変形処理に伴う入力画像（肖像画）の陰影・形状情報の損出
- 斜め向きの肖像画に非対応

また、本研究のコントリビューションについてまとめると以下の通りになる。

- 変形処理を行わない処理による、入力画像の陰影・形状情報の保持
- 斜め向き肖像画への対応

本手法では3つの工程に分けられ、1) 画像間の対応座標計算 (3.2 章)、2) 最適パッチ選択・置換 (3.3 章)、3) パッチ間境界線除去 (3.4 章) で構成されている。本章では各項目の詳細について述べる。

3.1 データベース構築

データベースとして、本稿では二次元顔画像を性別、年齢、顔向きの3つの要素別に構築した。ここで、年齢は20代から70代まで、顔向きは0度から左右45度まで(5度間隔)を用いた。但し、実写化処理を行う際には性別と年齢はユーザが指定し、入力となる肖像画から顔特徴点を検出する際に顔向きを自動判定することで、同じ顔向き角度のデータベースを選択する。画像サイズは入力画像の眉間から顎の長さで正規化しておく。

3.2 画像間の対応座標の計算

入力画像（肖像画）とデータベース画像間での対応座標計算の手順としては、入力画像と各データベース画像の顔

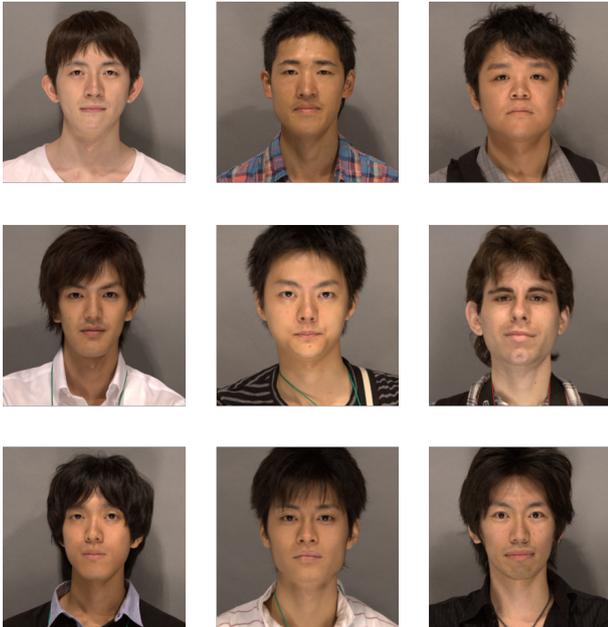


図 2: データベース画像例 (正面)



図 3: データベース画像例 (斜め)

特徴点 (86 点) をそれぞれ検出する. 次に二次元上の三角形メッシュを用意し, 得られた特徴点情報を基に, 入力形状とデータベースの各形状にフィッティングする. これにより, フィッティングされたメッシュモデルの三角形単位の対応関係を取得する. 次に, 入力画像をパッチ単位に分割し, パッチの中心座標からデータベース画像上での座標を算出することで, 探索すべきパッチ位置を決定する. 対応点座標計算の模式図を図 4 に示す. 図 4 において, 入力画像メッシュとデータベース画像メッシュは入力画像上のパッチ探索点 (パッチ中心座標) q が属する入力画像のメッシュと, メッシュ生成で対応付けられた任意のデータベース画像上のメッシュである. また, p'_1, p'_2, p'_3 は入力画像の特徴点 p_1, p_2, p_3 に対応するデータベース画像の特徴点である. 入力画像の特徴点 p_1 を原点とした, パッチの中心 q のベクトル \vec{q} は重み s, t と p_1 を原点とした p_2, p_3 の

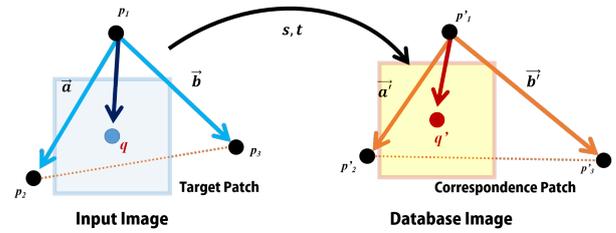


図 4: 対応座標計算の概要

ベクトル \vec{p}_2, \vec{p}_3 を用いて以下の式 (1) で表される.

$$\vec{q} = s \cdot \vec{p}_2 + t \cdot \vec{p}_3 \quad (1)$$

q に対応するデータベース画像の座標 q' のベクトル \vec{q}' は式 1 で算出した重み s, t と p'_1 を原点とした p'_2, p'_3 のベクトル \vec{p}'_2, \vec{p}'_3 を用いて以下の式 (2)

$$\vec{q}' = s \cdot \vec{p}'_2 + t \cdot \vec{p}'_3 \quad (2)$$

で求められ, \vec{q}' より, 対応パッチ中心座標 q' が求められる.

3.3 最適パッチ選択・転写

入力画像とデータベース画のパッチ対応関係を基に, 入力画像とデータベース画像の輝度値分布の比較を行うことで, 最適なパッチを選択する. ただし, 肖像画は作品ごとに色味が大きく異なるため, データベース上の実写顔画像と直接比較することが出来ない. そこで, Kawai ら [10] の手法を参考に, 各データベース画像のパッチ内輝度値分布の平均値を比較対象となる肖像画像パッチ内輝度値の平均に統一させたものを用いた. 具体的にはデータベース上の顔画像から得られるパッチ内の輝度値の平均値を, 入力画像の対応するパッチの輝度値の平均値に足し合わせることで補正を行った. パッチ内輝度値の補正後, 式 (3) を定義し, 誤差が最小となるようなパッチを選択する.

$$\mathbf{L}_{dst}(q) = \arg \min_{i \in N} \|\mathbf{L}_{IN}(q) - \mathbf{L}_i(q')\|^2 \quad (3)$$

ここで $\mathbf{L}_{dst}, \mathbf{L}_{IN}, \mathbf{L}_i$ はそれぞれ出力画像上の座標 q , 入力画像上の座標 q 及び, データベース画像上の対応座標 q' におけるパッチの輝度値である. ただし, i はデータベースの番号である ($i \in \{1, \dots, N\}$ N : データベースの総数). 選択された最適なパッチの RGB 値を転写することで, 実写真画像のテクスチャの転写を行う. パッチの転写結果画像を図 1c に表す.

3.4 境界線除去

パッチ合成後画像図 1c では, データベース上の人物顔画像のパッチを直接用いるため, パッチ間が不連続な合成結果となる. そこで, パッチ間の不連続性を取り除くため

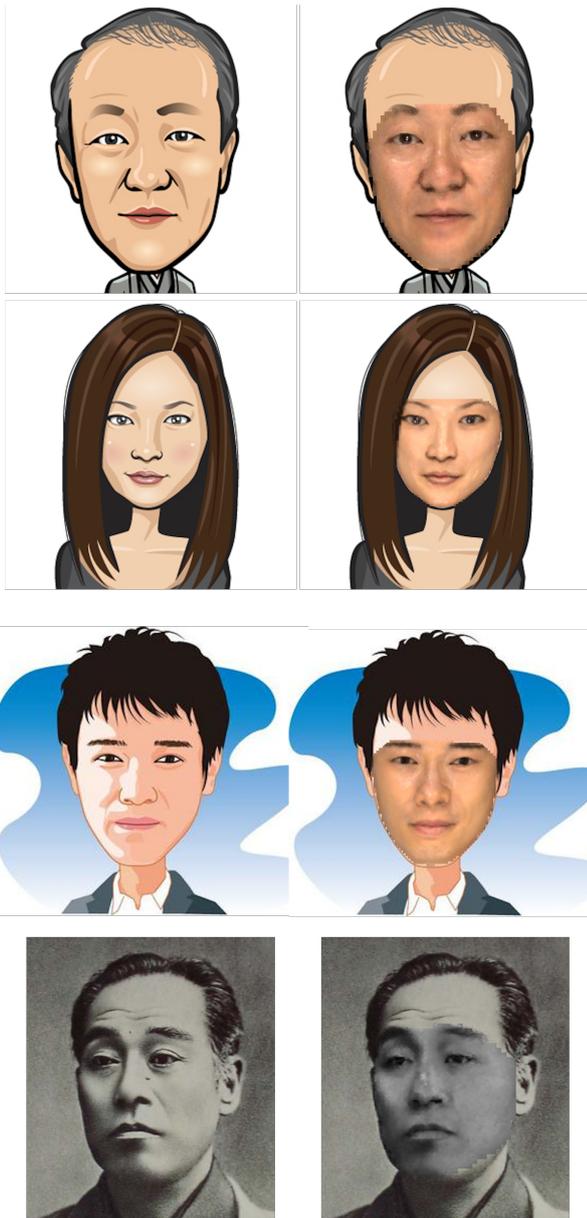


図 5: 生成結果

に, Perez ら [11] の勾配情報を用いる合成手法を用いる. Perez らの手法は二枚の画像を勾配情報は保持し, 境界条件を揃えることで自然に合成する手法である. 本手法では隣接するパッチ同士を 1[pixel] 分重複させ, 重ねた個所の平均色 (RGB 値) を境界条件として設定した. これにより, パッチ合成後の勾配を保持され, 全体の色味が統一された自然な顔画像が生成される.

4 生成結果

生成結果を図 5 に表す. 図の左の画像が入力画像であり, 右が本手法によって得られた実写化された画像である. 本手法によって, 入力画像の顔の各部位の形状や陰影を保持しつつ, 色味やテクスチャが実写真のように表現できていることが確認できる. 特に, 頬の光沢や, 目頭の奥行など

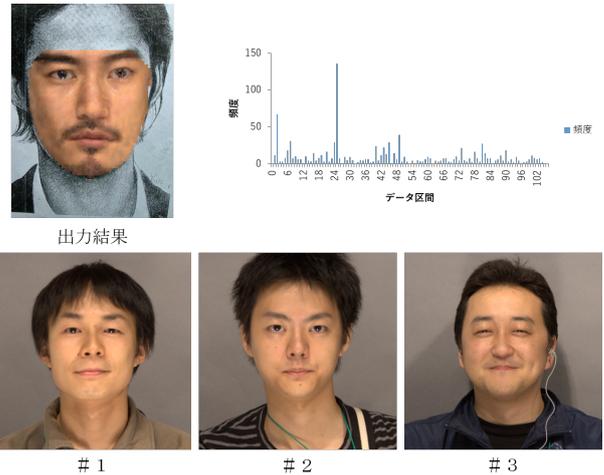


図 6: 生成結果で使用されたデータベース画像のヒストグラム

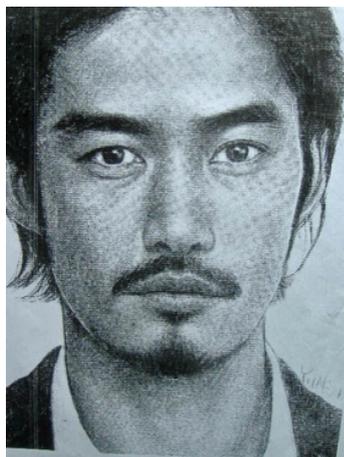
細部にわたって実写化ができていることが見てわかる. また, 生成結果においてデータベースの度の画像が使用されているかをヒストグラムで表したものの一例を図 6 に示す. ここで # 1, # 2, # 3 は使用された頻度の上位 3 位以内のデータベース画像であり, それぞれ 14%, 7%, 4% の割合で試用された. 上位一位の画像はよく参照されているが, 全体的に同じように参照されていることがわかる.

4.1 溝川らの手法との比較

本手法の有用性を検証するために, 既存手法 (溝川ら) の手法との比較検証を行った. 正面画像 (0 度) の比較結果を図 7, 8 に示す. ここで, 提案手法, 既存手法ともに, 図 7 では 30 代男性 105 枚, 図 8 では 20 代女性 95 枚のデータベースを使用した. パッチサイズは図 7, 図 8 共に 20[pixel]×20[pixel] とした. その結果, 従来手法と比べ, 入力となる肖像画の形状と陰影を保持し, 肌テクスチャを転写した実写らしい顔画像を自動生成できたことが確認できる. また, 従来手法で生成が困難であった斜め顔の肖像画を実写化した結果を図 9 で示す. データベースは角度左 15 度, 20~30 代の男女 80 枚の画像を用い, パッチサイズは 12[pixel]×12[pixel] とした. 本手法を用いることで, 斜め向き画像に対しても実写化できたことが確認できる. ただし, 出力結果の目や口の詳細部分がぼけてしまい, 肖像画の見た目の印象と異なってしまいう課題も確認できる.

5 まとめと今後の課題

本稿では, 入力の肖像画の形状や陰影を保持したパッチベースの実写画像生成手法を提案した. 肖像画から取得した顔特徴点を基に生成した三角形構造を用いたパッチ単位の対応付けにより, 従来手法の課題であった顔向きへの対応を可能とした. 今後の課題として, 眉上部や輪郭部等へ



(a) 入力画像



(b) 提案手法



(c) 溝川ら.[2015]

図 7: 正面向き男性肖像画の比較結果結果



(a) 入力画像



(b) 提案手法



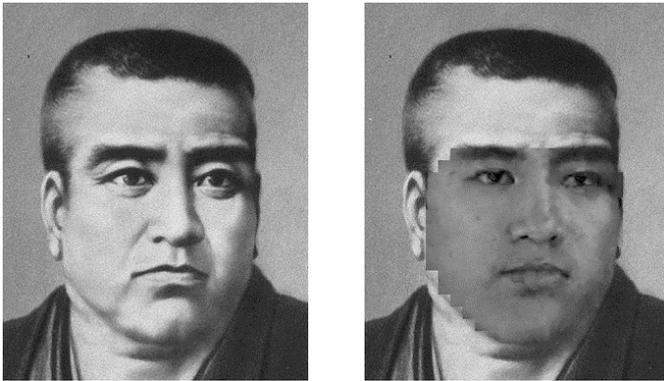
(c) 溝川ら.[2015]

図 8: 正面向き女性肖像画の比較結果

の拡張と、斜め向き肖像画の生成結果の質の向上が挙げられる。本手法は顔特徴点に割り当てた構成された二次元メッシュを利用するため、上記の領域は対応付けが行われなかった。そこで、より精細なテンプレートモデルを事前に用意することや、インペインティング処理 [12] による領域拡張を検討している。また、斜め向き肖像画において、出力結果の質が低かった理由として、肖像画の詳細を再現するようなパッチを正しく選択できなかったことが原因として挙げられる。そこで実写画像データベースの増築やパッチの周辺探索によるパッチ選択の候補を増やすことを検討している。

参考文献

- [1] Tanasai Sucontphunt, Borom Tunwattanapong, Zhigang Deng, and Ulrich Neumann. Crafting 3d faces using free form portrait sketching and plausible texture inference. In *Proceedings of Graphics Interface 2010*, pp. 209–216. Canadian Information Processing Society, 2010.
- [2] Di Wu and Qionghai Dai. Sketch realizing: Lifelike portrait synthesis from sketch. In *Proceedings of the 2009 Computer Graphics International Conference, CGI '09*, 2009.
- [3] Patrick Tresset and F Fol Leymarie. Generative portrait sketching. In *Proceedings of VSMM*, Vol. 5, pp. 739–748, 2005.



(a) 入力画像

(b) 提案手法

図 9: 斜め向き肖像画の実写化結果

- [12] Denis Simakov, Yaron Caspi, Eli Shechtman, and Michal Irani. Summarizing visual data using bidirectional similarity. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pp. 1–8. IEEE, 2008.

- [4] Long Zhu and Alan Yuille. A hierarchical compositional system for rapid object detection. *Department of Statistics, UCLA*, 2011.
- [5] Brendan F Klare, Zhifeng Li, and Anil K Jain. Matching forensic sketches to mug shot photos. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 33, No. 3, pp. 639–646, 2011.
- [6] 溝川あい, 森島繁生. “似顔絵からのリアルな顔画像生成”. *visual computing/グラフィックスとCAD 合同シンポジウム*, 発表番号 01, 2015.
- [7] Anat Levin, Dani Lischinski, and Yair Weiss. Colorization using optimization. In *ACM Transactions on Graphics (TOG)*, Vol. 23, pp. 689–694. ACM, 2004.
- [8] Vivek Kwatra, Irfan Essa, Aaron Bobick, and Nipun Kwatra. Texture optimization for example-based synthesis. In *ACM Transactions on Graphics (TOG)*, Vol. 24, pp. 795–802. ACM, 2005.
- [9] Umar Mohammed, Simon JD Prince, and Jan Kautz. Visio-lization: generating novel facial images. *ACM Transactions on Graphics (TOG)*, Vol. 28, No. 3, p. 57, 2009.
- [10] Masahide Kawai and Shigeo Morishima. Focusing patch: Automatic photorealistic deblurring for facial images by patch-based color transfer. In *MultiMedia Modeling*, pp. 155–166. Springer, 2015.
- [11] Patrick Pérez, Michel Gangnet, and Andrew Blake. Poisson image editing. In *ACM Transactions on Graphics (TOG)*, Vol. 22, pp. 313–318. ACM, 2003.