

3D 仮想眼鏡試着のためのリアルタイム眼鏡取替映像生成法

小林 巧[†] 杉浦 裕太[†] 斎藤 英雄[†](正会員) 上間 裕二^{††}[†]慶応義塾大学, ^{††}株式会社ジンズ

Real-Time Eyeglasses Replacement Method for 3D Virtual Try-on

Takumi KOBAYASHI[†], Yuta SUGIURA[†], Hideo SAITO[†](*Member*), Yuji UEMA^{††}[†] Keio University, ^{††} JINS Inc.

〈あらまし〉 本論文では、ユーザー自身が矯正眼鏡を着用した状態のままでも利用可能な新しい仮想眼鏡試着の映像生成手法を提案する。研究背景として、眼鏡店を訪れる消費者は鏡を使って試着用の眼鏡が自分に似合っているかどうか試すが、視力が悪いために矯正眼鏡を掛けた状態でないと鏡に映った自身の様子を確認できないというジレンマが存在する。スマートフォンのカメラで撮影したユーザーの顔画像に仮想的に眼鏡を掛けた映像を生成するアプリケーションも存在するが、それらほぼ全てのシステムは眼鏡を外した顔の画像または動画を撮影して仮想眼鏡を重ねることを前提としているため、視力の悪いユーザーが画面をはっきりと見ることができないという課題は解決されていない。そこで本手法では、カメラで撮影した動画中の実眼鏡領域を本人の皮膚テクスチャで上書きすることで消去し、さらに別の仮想眼鏡を重ねた複合現実感映像を生成する。これにより、眼鏡試着における障壁を取り除くことが可能な3次元仮想試着を実現する。また、ユーザーの顔の動きに合わせてリアルタイムに3D仮想眼鏡を表示するために、高精度で高速に顔追跡可能な顔解析ツールを応用する。さらに、矯正レンズの歪みによる変化と光の屈折・反射も考慮して仮想レンズを描画することで、通常の試着用眼鏡では確認できないようなより実際の眼鏡を掛けた状態に近い映像を生成することが可能となる。実験を通して、眼鏡を掛けた状態のままでも違和感の生じない試着映像がリアルタイムで生成されたことを示し、その有効性を確認する。

キーワード: 3D 仮想眼鏡試着, 眼鏡取替, 複合現実感, 拡張現実感

<Summary> This paper presents a 3D virtual eyeglasses try-on system for practical use. For fitting eyeglasses in a shop, consumers wish to look at themselves in a mirror while trying on various eyeglass styles. However, for people who need to wear eyeglasses for correcting problems with eyesight, it is impossible for them to clearly observe their face in the mirror without wearing eyeglasses. This makes fitting them for new eyeglasses difficult. Many studies on virtual glasses try-on have been conducted, and some eyewear brands have already put the system into practical use, but none of these mixed reality systems considered processing in case users wear their own glasses. This research proposes a virtual try-on system that can be used while wearing eyeglasses. We replace the user's eyeglasses in the input video with new eyeglasses virtually. Also, a fast and accurate face tracking tool enables our system to automatically display 3D virtual glasses following a user's head motion. Moreover, taking into account the change due to distortion of the correction lens and the refraction and reflection of light, it is possible to generate more realistic glasses by drawing a virtual lens. Experimental results demonstrate that the proposed method can render virtual glasses naturally while the user is wearing real eyeglasses.

Keywords: 3D virtual eyeglasses try-on, eyeglasses replacement, mixed reality, augmented reality

1. はじめに

一般的に、人は眼鏡を購入する際は、実際に眼鏡店へ行き試着用に用意された眼鏡を鏡の前で掛けることで見た目を確かめる。しかし、眼鏡の試着に関しては多くの課題が存在し

ている。まずはじめに、消費者が眼鏡を試着する際に感じる最大の不満点として店舗で試着用に用意されている眼鏡には矯正用レンズが挿入されていないため視界がぼやけてしまうという問題が挙げられる。人によって近視や遠視の度合いが異なるため、店舗での試着では度の入っていないレンズが

挿入された眼鏡しか通常試すことができない。だが、眼鏡を購入しようとする消費者は当然視力の悪い人が多く、矯正眼鏡を外して試着用眼鏡を掛けても鏡に映った様子をはっきりと確認することが困難である。したがって、本人の代わりに他人に眼鏡を掛けた様子を見てもらって意見を聞くか、写真を撮って後で確認するしか術がなかった。加えて、何度も様々な種類の眼鏡を掛けて外すという作業を繰り返すことは非常に煩わしく、いくつもの試着結果を覚えておくことは非常に困難であるといえる。さらに、眼鏡を掛けている状態のユーザーの見た目は、矯正レンズの挿入前と挿入後で光の屈折の影響によって目の大きさなどの見た目が変わってしまうという懸念点もある。この点においても試着用眼鏡では再現することができない課題の一つであるといえる。

一方で、複合現実感と呼ばれる仮想空間と現実空間を混同させる映像生成技術の進歩により、ファッション業界でも仮想試着サービスの実応用化が加速している。近年では、実際に撮影された人物や空間に仮想的に衣服や家具を重畳することで、仮想空間上で商品を試すことができる技術も盛んに開発されている。このようなシステムは実際に店舗に行かなくても自宅などで試着することが可能になるため、インターネット販売の売上促進に繋がる。眼鏡に関しても、仮想試着システムによって場所を問わず手軽に眼鏡を試すことができるようになるといえる。実際にいくつかの眼鏡ブランドでは、スマートフォンのカメラで撮影したユーザーの顔画像に仮想的に眼鏡を掛けた映像を生成するアプリケーションも実用化されつつある。しかし、それらほぼ全てのシステムは、眼鏡を外した顔の画像または動画を撮影して仮想眼鏡を重畳することを前提としているため、視力の悪いユーザーが画面をはっきりと見ることができないという課題は解決されていない。特にリアルタイムで仮想試着映像を生成するサービスにおいては、仮想眼鏡を掛けている様子を直接目で確認できるように、矯正眼鏡を外す必要のないシステムであることが求められる。

そこで本研究では、眼鏡試着における課題点を解決するために、ユーザー自身が掛けている矯正眼鏡を掛けた状態のままでも利用可能な新たな仮想眼鏡試着映像生成法を提案する。ここで、本研究における前提条件を整理する。

- 様々な状況下で簡易的に利用可能な仮想試着システムを実現するために、距離カメラや特別なセンサなどは使用しない。また、スマートフォンやPCでも作動できるような簡易的なシステムを想定し、単眼のカメラだけを使用する。
- 仮想試着を開始する前の準備として、ユーザー本人の目周辺の皮膚テクスチャを取得するため、眼鏡を外した顔画像を1枚だけ撮影する。さらに、背景の画像も事前に1枚撮影する。カメラは固定されているものとし、背景は大きく変化しない。

また、本研究では、顔追跡と顔器官検出のための手法として Convolutional Experts Constrained Local Model (CE-CLM)^{1),2)}を用いた顔解析ツール「OpenFace³⁾」を使用する。OpenFaceにより、顔の特徴となるランドマーク点の3次元位置と2次元位置を高速で取得することが可能になり、3次元的なユーザーの顔の動きに合わせてリアルタイムに仮想眼鏡を重畳することができることが特徴である。さらに、入力動画中の人物の実眼鏡が存在すると推定される目周辺のマスク画像を取得し、その領域に顔皮膚テクスチャのマッピングを行うことにより、実眼鏡の領域推定を行わずに消去することが可能になる。次に、仮想眼鏡のレンズとフレームの鏡面反射光と、レンズの表面における光の屈折を考慮することで、より再現性の高い仮想眼鏡を描画できることを明らかにする。

2. 関連研究

2.1 3D 眼鏡仮想試着システム

近年では、眼鏡仮想試着システムに関する研究が多くなされており、眼鏡ブランドの中では既に実用化しているサービスを提供しているものもある。3次元的な顔の動きに合わせて忠実に仮想眼鏡が重畳される複合現実感を実現するには顔の3次元形状の取得する必要があるため、仮想眼鏡試着に関する多くの研究では距離カメラを使用している^{4),5)}。また、Yuan⁶⁾の仮想眼鏡試着システムのように通常のカメラを利用している研究もあるが、鏡とスクリーンを組み合わせた大規模な装置が必要であり、事前に手作業による操作も求められるため簡易的なシステムであるとはいえない。一方で、眼鏡のレンズにかかる光の反射と屈折、フレームの影まで忠実に描画することで、仮想眼鏡の現実感を高めることを目的とした研究も行われている⁷⁾。レイ・トレーシング法によって仮想眼鏡の描画を行っているためリアリティは非常に向上するが、計算コストも増大してしまうためリアルタイム性を求めるシステムには向いていない。

そして、これらの仮想試着システムは実眼鏡を掛けたユーザーが利用する場合を想定しておらず、眼鏡試着における潜在的な問題は残存している。当然であるが、眼鏡を着用したまま従来の仮想試着を行った場合、図1(a)のように実眼鏡と仮想眼鏡の二つの眼鏡が重なって表示された不自然な結果が得られてしまう。なお、図1(b)は改善例である。(本提案方法を利用)

2.2 顔画像中の眼鏡消去

画像中から眼鏡を消去する手法については、主に学習ベースのアプローチをとっている研究が広く行われている。顔画像のデータセットを主成分分析することで入力顔画像に類似した目付近の皮膚画像を再構築する手法^{8),9)}や、深層畳み込みニューラルネットワーク(DCNN)によって眼鏡を掛けた顔画像から眼鏡を外した画像に変換する手法¹⁰⁾がある。さらに、敵対的生成ネットワーク(GAN)¹¹⁾と呼ばれる学習ベ-

スの画像生成モデルによって、眼鏡を掛けていない顔画像を再現する研究も行われている。しかし、これらの手法は顔認識を行う際に精度低下の原因になりうるオクルージョン処理を主な目的としている場合がほとんどであり、入力として動画ではなく静止画を扱っている。

2019年11月にJINS社が公開した「Megane On Megane¹²⁾」では、本研究と同様に眼鏡の試着映像を生成することを目的として動画中の眼鏡消去を行っている。この手法ではディープラーニングによって顔画像データを学習することで、眼鏡を着用した顔画像から眼鏡を外した状態の顔画像に変換

して画像生成を行っている。これに対して本手法では、眼鏡を装着しない状態で撮影した本人の顔表面テクスチャを、眼鏡装着時の眼鏡領域にマッピングことで眼鏡の無い画像を生成しているため、学習ベースの手法で必要となる大量の顔画像や眼鏡あるいは眼鏡を掛けた顔画像のデータベースを作成する必要が無く、また学習と認識時に高性能なGPUも用いる必要も無い。

3. 提案手法

本研究では、ユーザーの眼鏡を別の眼鏡に自動的に取り替える仮想試着映像生成法を提案する。図2に提案手法の流れを示す。

まず、仮想試着を開始する前に背景画像1枚と眼鏡を外した素顔画像1枚を撮影する。背景画像は仮想レンズを通して映る背景部分のテクスチャとして、素顔画像は仮想レンズの描画と眼鏡消去で必要になる皮膚テクスチャとして用いる。この事前準備を行った後、実眼鏡を掛けたユーザーを動画で撮影して仮想試着へと移る。本手法の構成は大きく四つの段階に分けられる。まず、顔解析ツール「OpenFace」を利用し、以降の処理に必要な顔器官点の検出・追跡と顔姿勢推定を行う。顔器官検出は入力動画だけでなく素顔画像に対しても行い、顔の皮膚のテクスチャ座標を取得する。そして、OpenFaceによって得られた入力動画中の顔の3次元位置姿勢に合わせて、仮想の矯正レンズと眼鏡、事前に取得した顔皮膚テクスチャを張り付けた顔モデルをそれぞれ3次元空間上に配置する。次に、生成した顔皮膚テクスチャと矯正レンズをカメラ視点から見た画像中に投影し、入力フレームに合成することで実眼鏡領域を消去する。最後に、オクルージョ

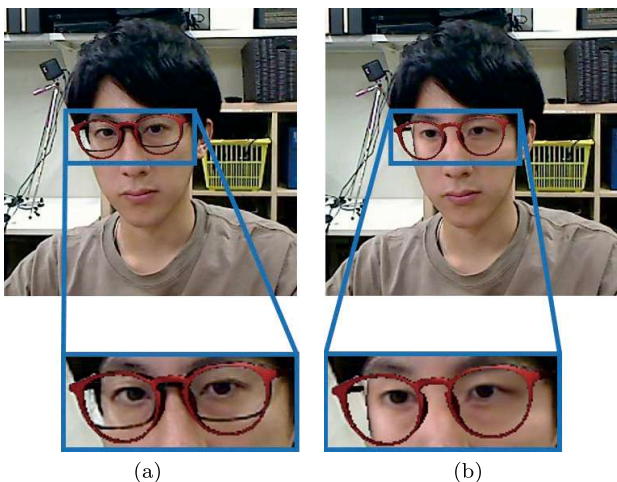


図1 実眼鏡着用時の眼鏡切替の例 ((a)従来の仮想眼鏡試着システムを利用, (b)提案手法を利用)

Fig.1 Example of eyeglasses switching while wearing real eyeglasses; (a) Result by a conventional virtual reality try-on and (b) Result of the proposed method)

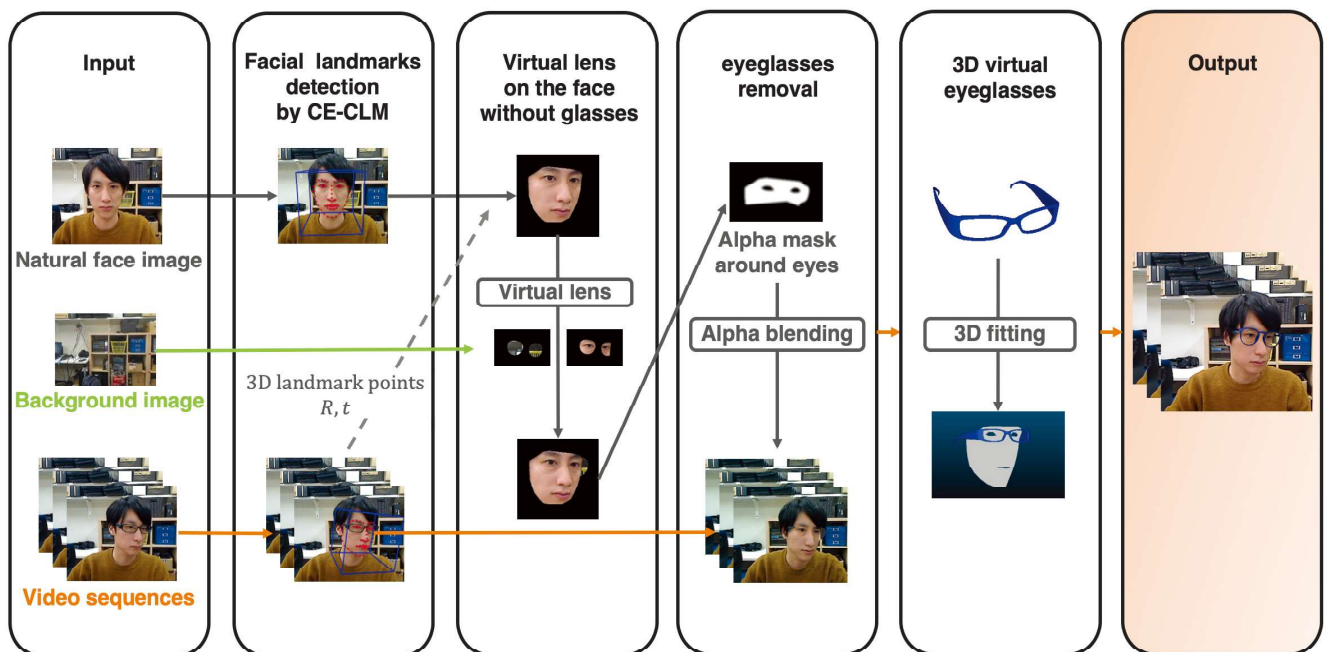


図2 提案手法の流れ

Fig.2 Overview of the pipeline for our 3D virtual try-on system

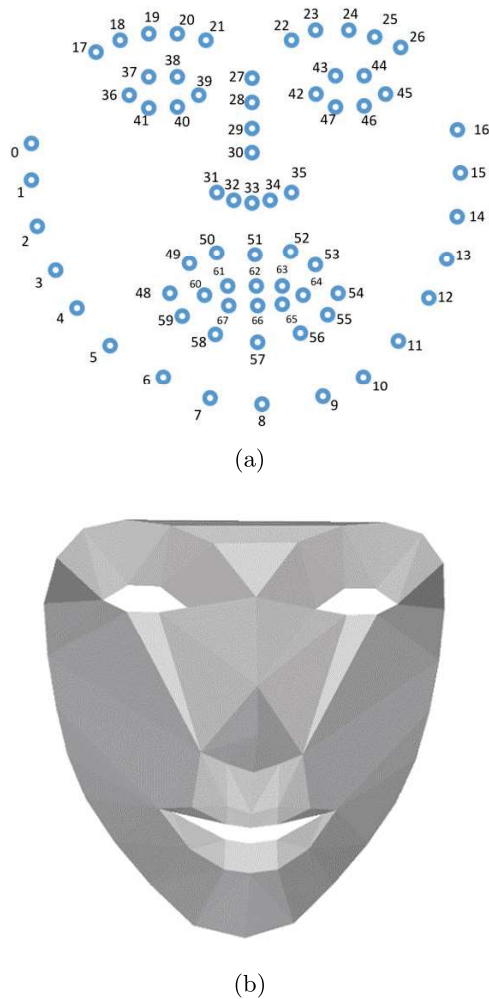


図3 Point Distribution Model (PDM) ((a) PDM で表現される顔器官 68 点, (b) PDM 顔モデルの各頂点を結ぶ三角メッシュで表現された 3 次元顔形状)

Fig. 3 Point Distribution Model (PDM); (a) Facial landmark points of PDM and (b) 3D face shape reconstructed by triangle mesh

ン処理を施した仮想眼鏡を重畳することで、別の眼鏡に取り替えたような画像を生成する。これらの処理を毎フレーム行うことで複合現実感映像を生成し、眼鏡を着用した状態でも利用可能な仮想試着システムを実現する。

3.1 顔器官点検出と顔姿勢推定

本手法では、OpenFace を利用して動画中のユーザーの顔器官点検出を行う。OpenFace では、ディープラーニングによる顔検出器 CE-CLM によって、Point Distribution Model (PDM)¹³⁾で表現される顔器官 68 点に対応する 3 次元座標と 2 次元座標を取得することが可能である。PDM の顔モデルに含まれる顔器官点 68 点を図 3 (a) に示す。そして、内部的に各点の 3 次元座標と 2 次元座標の対応関係から PnP(Perspective-n-Points) 問題を解くことで、顔の位置姿勢推定までを行う。ゆえに、取得された顔器官点座標と顔位置姿勢推定の結果を利用することで、皮膚テクスチャのマッピングや仮想眼鏡を重畳することが可能になる。ここで、事前撮影された素顔画像を顔解析し、検出された顔器官点の 2 次元座標を u^{skin} ,

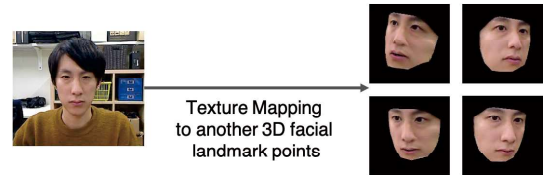


図4 別の 3 次元顔器官点に合わせてテクスチャマッピングした顔モデルの例

Fig. 4 Example of face model by texture mapping to another 3D facial landmarks



図5 レンズ表面での屈折の例

Fig. 5 Example of refraction on the surface of prescription lens

入力動画フレーム中の顔器官点の 3 次元座標を v とおく。

$$u_i^{skin} = (x_i, y_i), \quad (i = 0, 1, \dots, 67) \quad (1)$$

$$v_i = (x_i, y_i, z_i), \quad (i = 0, 1, \dots, 67) \quad (2)$$

動画中のユーザーの顔の各顔器官点の 3 次元位置 v_i に皮膚テクスチャの u_i^{skin} が対応しており、図 3 (b) で表される三角メッシュを構成する面上にテクスチャマッピングすることで、最初に撮影した顔画像を入力動画中の顔の向きや表情の変化に合わせて変形させたような素顔の 3 次元顔形状を再現することが可能になる。皮膚テクスチャを別の顔形状にテクスチャマッピングすることで生成された 3 次元顔モデルの例を図 4 に示す。

3.2 矯正レンズの描画

矯正レンズには歪みがあるため、図 5 のように近視レンズを通すと光の屈折により輪郭が内側に狭まったように映ることや目が小さく見えるなどの変化が生じる。一般的に矯正レンズは単純な凹凸形状になっているのではなく、中心部と外側の厚みを変化させるなど様々な形状が存在する。コンピュータグラフィックスで仮想レンズの屈折による影響を正確に表現するためには、レンズの表面だけでなくレンズ内を通過した裏側の面の法線も考慮した屈折光の方向ベクトルと、その先にある物体との距離を把握する必要がある。しかし、リアルタイムで描画する際にこれら全てを計算しようとする

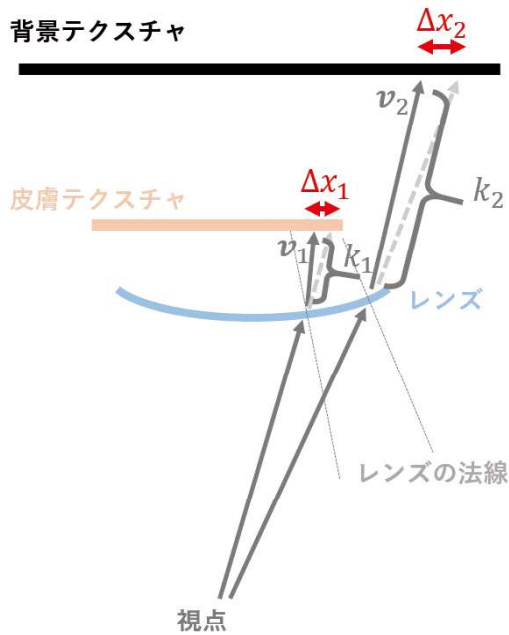


図6 屈折光の光路
Fig. 6 Refraction light path

と処理が非常に重くなってしまう。そこで本手法では、図6で示すように単純にレンズの表面の屈折のみを考慮し、レンズを通過した屈折光 v を一定の長さ k で変化させることで、レンズとの距離が大きく異なる顔と背景の歪みの違いを表現する。まず、顔皮膚画像と背景画像が映る仮想レンズを二種類生成することを考える。ここで、レンズを通り皮膚テクスチャと背景テクスチャに届く屈折光の単位ベクトル v_1, v_2 の長さをそれぞれ k_1, k_2 とおく。屈折光は入射角を θ 、反射角を ϕ とすると、スネルの法則 $\sin \theta = n \sin \phi$ (n は屈折率の比) により求められる。そして、テクスチャ画素の移動量である $\Delta x, \Delta y$ は、次の近似式によって計算し、レンズ上にマッピングする。

$$\Delta x \approx kv_x \sin(\theta - \phi) \quad (3)$$

$$\Delta y \approx kv_y \sin(\theta - \phi) \quad (4)$$

仮想試着中に入力フレームに仮想レンズを重ねていく際には、皮膚画像を映したレンズだけでは埋めきれない顔輪郭外の部分を仮想レンズを通した背景テクスチャで合成するという処理を行う。

また、レンズには屈折光だけでなく鏡面反射光も考慮する必要がある。忠実に反射光を再現するためには、全方位カメラを使って環境マップと光源位置を取得する必要があるが、本手法では単純にカメラ上方に光源が存在すると仮定する。そして、仮想眼鏡の向きによって、レンズの法線方向を軸として視点の反射方向に光源が来る場合に光が反射して見えるように描画する。これによって、顔の動きに合わせてレンズの色味も変化するようになり、仮想レンズの再現性を高められることが期待される。レンズとフレーム部分の反射光に

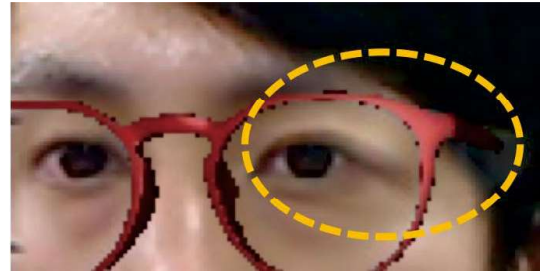


図7 仮想眼鏡の鏡面反射光を反映させた例
Fig. 7 Example of specular reflection on virtual glasses

よる色情報の変化を加味した描画は OpenGL による Phong シェーディング¹⁴⁾によって実装した。仮想眼鏡に反射光の影響を反映させた様子の例を図7に示す。

3.3 眼鏡消去

次に、実眼鏡着用時にも利用可能なシステムを実現する上で不可欠な処理である、顔画像中の眼鏡消去を行う。画像中から不要な領域を消去する方法としては、近傍画素などから推定される別の画素値によって上書きする Inpainting^{15),16)} という手法がある。眼鏡の場合であれば近傍の皮膚の画素を用いて消去する方法が考えられるが、Inpainting によって境界部分を中心に滑らかに皮膚の画素を生成することは困難であり半透明な眼鏡が表示されているような違和感のある結果になってしまう。さらに、動画中の実眼鏡領域は細く小さいため、マスク画像となる眼鏡領域を正確に抽出することも難しい。したがって、本手法では動画中の実眼鏡を追跡・検出するのではなく、顔位置推定の結果を利用する。眼鏡は常に目周辺に存在すると想定し、事前に眼鏡を外した素顔画像の目周辺領域を動画中の対応位置に合成することで実眼鏡領域を消去する。この方法により、ユーザーが掛けている眼鏡のサイズや大小に依存することなく処理することが可能となる。

まず、3.1 節で生成した素顔の3次元顔形状モデルを顔姿勢推定の結果を用いて2次元画像上に投影する。顔全体の皮膚テクスチャを置き換えるのではなく、目周辺のみの領域を切り抜いて合成することを考える。本手法で用いる PDM によって表現される顔モデルの各頂点にはそれぞれ ID が振られているため、目周辺の領域を判断することができる。これにより、対象領域を抽出してマスク画像を生成することが可能となる。このとき、目の内部は動画中の目を使用するため、マスク画像には含まないようにする。目の内部は取残すことで、ユーザーの視線の動きだけでなく、実際に掛けている矯正眼鏡レンズを通して見える目の大きさや形状を維持することが可能になり自然な合成結果を得られる。

ただし、このまま合成処理を行ってしまった場合、合成した境界部分が目立ってしまい望ましい結果を得ることができない。よって、合成割合を境界部分を中心に段階的に変化させるアルファブレンディングを行う。まず、合成領域を表すマスク画像を二値画像で表す。そして、マスク画像に対して平滑化フィルタを掛け、境界部分が徐々に0から1の値にな



図8 アルファマスク画像
Fig.8 Alpha mask image

るように周辺画素の平均値を計算していく。このマスク画像の画素値を α とおき、 α の値に従って二つの画像の合成割合を決定する。すなわち、各画素値の値は次の式 (5) によって求められる。

$$I = \alpha S + (1 - \alpha)D \quad (5)$$

I は合成後の各画素値であり、 S が合成元となる皮膚テクスチャの画素値、 D が合成先の入力フレームの画素値を表す。これにより得られるアルファマスク画像の例を図8に示す。

そして、3.2節で生成した仮想レンズを重ねる。仮想レンズは境界部分を滑らかにする必要がないためアルファブレンディングは行わずに合成する。

3.4 3D 仮想眼鏡の重畳

3D 仮想眼鏡の重畳処理について述べる。複合現実感を表現するために、ユーザーの顔の動きに追従して仮想眼鏡を移動・回転させながら表示する必要がある。顔姿勢推定の結果から仮想眼鏡モデルを剛体変換させることができるが、そのまま2次元投影してしまうと眼鏡のテンプル(つる)の部分などの姿勢によっては顔で隠れるはずの部分がそのまま表示されてしまう。オクルージョン処理がされていない例を図9(a)に示す。そこで、入力フレームの3次元顔位置推定の結果から頭部の3次元形状を再構成し、仮想眼鏡が頭部よりも内側にありオクルージョンとなるべき領域を判断して、手前に見えるべき領域だけを適切に表示するように描画処理を行う。具体的には、顔の中心から正面方向に z 軸をとり、顔輪郭を表す線分を通り z 軸方向に平行な面よりも内側に存在する仮想眼鏡領域を非表示にし、外側に存在する領域のみを表示した。図9(b)に適切にオクルージョン処理を行った結果を示す。

4. 実験結果と考察

実眼鏡を着用したユーザーに対して、提案手法によって仮想試着を行った実験結果を述べる。実験で用いたPCのOSはUbuntu 18.04 LTS, CPUはInter Core i7-7700K @ 3.60 GHz, GPUはNVIDIA GeForce GTX 1070である。GPUについては仮想眼鏡のCG処理のみに用いた。撮影に用いるカメラは、Logicool社製のUSBカメラ「Logicool Webcam Pro 9000」を使用した。リアルタイムで処理するために撮影画像の解像度は 640×480 とし、仮想眼鏡をレンダリングした合成元画像も全て同じ解像度に合わせた。

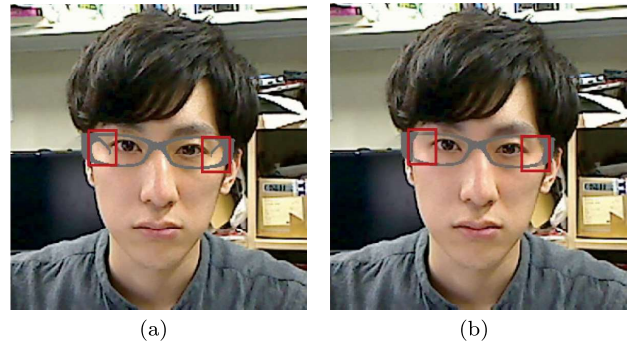


図9 オクルージョン問題とその処理例 ((a) オクルージョン問題, (b) 処理結果)

Fig.9 Example of occlusion problem and its handling; (a) Occlusion problem and (b) The result after occlusion handling

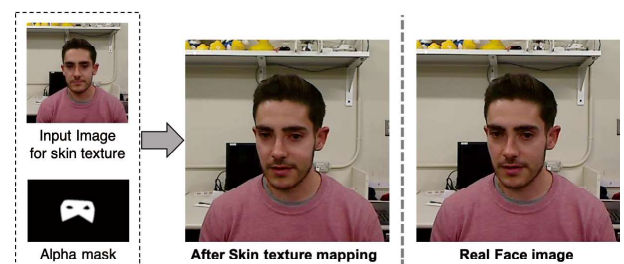


図10 顔皮膚テクスチャをマッピングした結果と実際の顔画像

Fig.10 Result of facial skin texture mapping and real face image

実験の流れとしては、まずパソコンにUSBカメラを設置し、事前に背景となる画像を1枚撮影しておく。そして、ユーザーはパソコンの前に座り、仮想試着を行った。最初に、画面の指示によってユーザーは眼鏡を外し、正面を向いた画像を1枚撮影した。そして、動画撮影が開始された後にユーザーが自身の眼鏡を着用し、頭を回転させて様々な視点と角度を変えながら見た目の確認を行った。提案手法による仮想試着映像生成の実行速度は12~15fpsであった。

4.1 顔皮膚テクスチャマッピング

まず、実眼鏡領域を消去するために行う顔皮膚テクスチャマッピングの結果を図10に示す。再現性を確認するために、眼鏡を着用していない実際の顔画像と比較を行う。提案手法では目周辺の3次元顔器官点が一致するようにマッピングを行っているため、使用する皮膚テクスチャの元となる顔画像が別方向を向いていたとしても、定性的に大差ない再現結果が得られている。また、図10のマスク画像を用いてアルファブレンディングを行ったため、境界部分が目立たないように合成されていることを確認した。

ここで、顔皮膚テクスチャマッピングを行った結果が実際の顔画像を復元していることを定量的に評価するために、これらの画像の類似度を計算した。ここでは類似度の指標として、画像の構造的類似性を表すSSIM(Structural similarity)を用い、眼鏡を掛けていない3人の被験者それぞれに対して

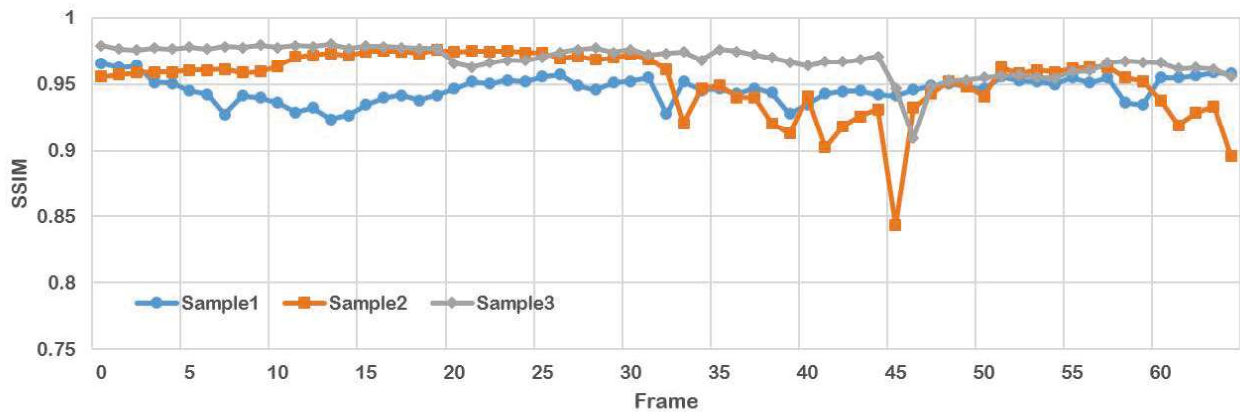


図 11 顔皮膚テクスチャマッピング結果と実際の顔画像の類似度 (SSIM)

Fig. 11 Similarity calculation between each result of facial skin texture mapping and real face image in terms of SSIM

約 5 秒間撮影した動画に目周辺の顔皮膚テクスチャを合成した結果の 65 フレーム分を比較した。顔の動きに依らず違和感なく合成されていることを確認するため、被験者は顔を上下左右に回転させるなどして動画を撮影した。また、目的の顔領域のみの類似度を計算するために、後処理として全フレームにおいて顔の輪郭が収まるようにトリミングし、それぞれ 120×120 の画像サイズに調整してから SSIM 値を求めた。図 11 に計算した SSIM の値をグラフで示す。SSIM の値は 1 に近いほど 2 枚の画像が類似しており、一般的に 0.9 以上であれば人間の感覚では差がないように見えるとされている。今回の実験結果では被験者が大きく横を向いた場合に SSIM 値が低下したが、0.9 を下回っているのは Sample2 の 2 フレームのみであった。この結果、使用するユーザーやその姿勢に依らず類似性の高い復元結果を得られていることを確認した。

4.2 実験結果と本手法の有効性の検証

図 12 (a)～(d) に、3 人の被験者に提案手法による仮想試着映像の一部の結果画像を示す。背景画像についても二つのパターンで実験を行った。図 12(b) は眼鏡消去を行った結果を表しており、図 12(a) で顔上に存在していた実眼鏡領域が皮膚テクスチャで上書きされていることが確認できた。また、合成された境界部分も滑らかに合成されていることが分かる。図 12(c) は、図 12(b) に仮想レンズを描画した結果である。この結果により、仮想レンズ内の部分の実眼鏡領域も消去されていることを確認した。さらに、顔の部分と背景の部分で屈折による影響が異なるように描画されていることが確認できる。最後に、図 12(d) に仮想眼鏡を重畳した最終的な生成画像を示す。図 12(a) と比較すると、図 12(a) に存在していた実眼鏡が仮想眼鏡に取り替えられた画像が生成されていることが分かる。この結果から、視力の悪い人であっても自身の眼鏡を着用したまま利用できるため、容易に試着の様子を確認することができた。

次に、本手法の有効性を検証するために、従来の仮想眼鏡試着システムのように仮想眼鏡の重畳処理のみを行った場合の結果と、提案手法により生成した仮想試着結果を比較した。図 13 に異なる 2 パターンの背景画像で比較した実験結果を示す。図 13(a) は、入力フレーム、図 13(b) は、眼鏡消去処理を行わずに仮想眼鏡のフレームのみ重畳した結果であり、図 13(c) が本手法による結果である。図 13(b) では、ユーザーが実際に着用している眼鏡と仮想眼鏡が二重に重なって見え、新しい眼鏡の形状や色が見にくいことが分かる。一方で、図 13(c) では実眼鏡が本人の皮膚テクスチャによって消去されているため、仮想眼鏡だけを着用した試着映像になっており、眼鏡を着用したまま仮想試着を行う場合に本手法が有効であることを確認できる。

ただし、本手法では実眼鏡領域自体の検出は行っていないため、実眼鏡のサイズや形状が極端に大きく顔や仮想眼鏡よりも外側に実眼鏡が存在している場合などには完全に消去することができない。このような場合に、背景の画像で埋め合わせる方法も考えられるが、実眼鏡の裏側に髪の毛や頭が存在する場合もあるため望ましい結果を得られない。これらの実眼鏡領域については消去されていなくても、仮想試着の結果においては大きな影響がなかったため本手法では対処しないこととした。また、本手法では実際の照明環境を取得していないため仮想眼鏡の影のレンダリングは行っていない。顔皮膚テクスチャを合成する際にユーザーが実際に掛けている眼鏡の影が一部消えてしまうという課題があるものの、架空の光源のみで違和感なく仮想眼鏡の影を描画することが困難であった。よって、仮想眼鏡の影は再現されていないが、陰影処理や反射光を反映させることでユーザーが現実感を得られたということを実験を通して確認した。

5. む す び

本論文では、単眼カメラのみで作動する、ユーザーが眼鏡を着用したまま利用できる仮想眼鏡試着のための映像生成

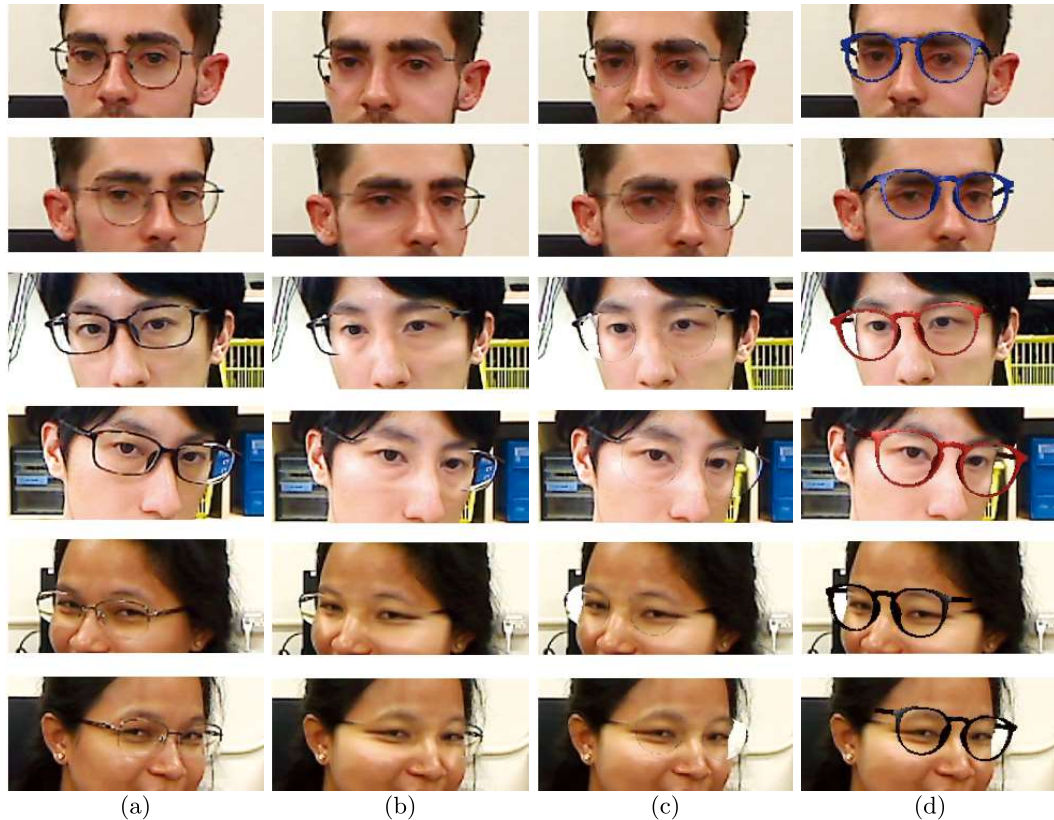


図 12 提案手法による画像処理例 ((a) 入力フレーム, (b) 眼鏡消去後画像, (c) 仮想レンズ描画後画像, (d) 仮想眼鏡重畳画像)

Fig. 12 Examples of the images processed by the proposed method; (a) input image, (b) Image after eyeglasses removal, (c) Image with virtual lens rendering, and (d) Image with virtual glasses rendering)

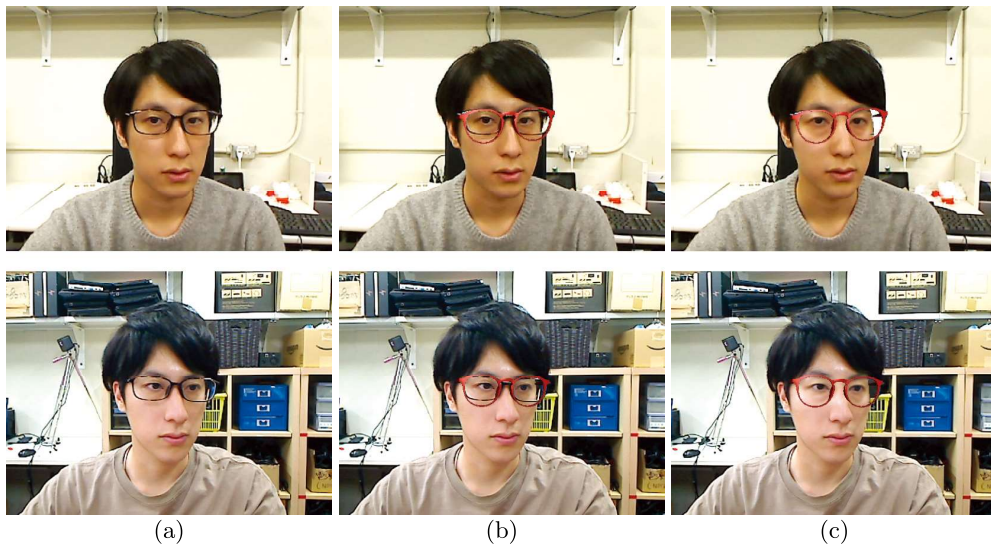


図 13 本手法と従来の眼鏡試着画像の比較 ((a) 入力フレーム, (b) 従来の仮想眼鏡試着, (c) 提案手法)

Fig. 13 Comparison between result images obtained by the proposed method and the previous virtual fitting; (a) input image, (b) usual virtual fitting and (c) proposed method

手法を提案した。提案手法では、眼鏡を着用しない状態で撮影した 1 枚の画像を使用して、眼鏡を着用した状態で試着中の動画像中の眼鏡のみを消去し、試着する仮想眼鏡を顔の向きに応じて重畳表示することにより、眼鏡を着用した状態での仮想試着を実現した。実験結果を通して、提案手法で生成された画像において違和感なく皮膚テクスチャと仮想眼鏡が

合成されていることを示した。また、本手法では素顔の顔画像と背景画像を 1 枚撮影するだけなので、多くの眼鏡消去手法に用いられる顔画像データセットや学習に必要な GPU も不要である。

今後の課題として、顔や仮想眼鏡よりも外側に存在する実眼鏡領域の処理が挙げられる。実眼鏡領域の検出法を検討し、

周辺画素などで完全に消去することができればより理想的な結果が得られると考えられる。また、本手法では実際の照明環境を取得していないため仮想眼鏡の影の部分については考慮していない。しかしながら、複合現実感における3次元CGにおいて仮想物体の影は接触している物体同士の構図を理解するために重要な要素の一つである。したがって、本手法においても全方位カメラを用いて実際の照明環境を取得することでシャドウマッピングを行うが可能になり、試着結果の再現性が向上すると考えられる。

謝 辞

本研究は、JST AIP-PRISM JPMJCR18Y2の支援を受けたものです。

参考文献

- 1) A. Zadeh, T. Baltrušaitis, L.-P. Morency: "Convolutional Experts Constrained Local Model for Facial Landmark Detection", Proc. of IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp.2519–2528 (2017).
- 2) T. Baltrušaitis, P. Robinson, L.-P. Morency: "Constrained Local Neural Fields for Robust Facial Landmark Detection in The Wild", Proc. of IEEE International Conference on Computer Vision Workshops, pp.354–361 (2013).
- 3) T. Baltrušaitis, A. Zadeh, Y. C. Lim, L.-P. Morency: "OpenFace 2.0: Facial Behavior Analysis Toolkit", Proc. of IEEE International Conference on Automatic Face and Gesture Recognition, (2018).
- 4) D. Tang, J. Zhang, K. Tang, L. Xu, L. Fang: "Making 3D Eyeglasses Try-On Practical", Proc. of IEEE International Conference on Multimedia and Expo Workshops (ICMEW), pp.1–9 (2014).
- 5) P. Azevedo, T. O. Dos Santos, E. D. Aguiar: "An Augmented Reality Virtual Glasses Try-On System", Proc. of XVIII Symposium on Virtual and Augmented Reality (SVR), pp.1–9 (2016).
- 6) M. Yuan, I. R. Khan, F. Farbiz, A. Niswar, Z. Huang: "A Mixed Reality System for Virtual Glasses Try-on", Proc. of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry", pp.363–366 (2011).
- 7) Q. Zhang, Y. Guo, P.-Y. Laffont, T. Martin, M. Gross: "A Virtual Try-On System for Prescription Eyeglasses", Journal of IEEE Computer Graphics and Applications, Vol. 37, No.4, pp.84–93 (2017).
- 8) C. Wu, C. Liu, H.-Y. Shum, Y.-Q. Xu, Z. Zhang: "Automatic Eyeglasses Removal from Face Images", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 26, No. 3, pp.359–363 (2004).
- 9) G. Pei, S. Fei: "Enhanced PCA Reconstruction Method for Eyeglass Frame Auto-Removal", Proc. of the 4th IEEE International Conference on Network Infrastructure and Digital Content, pp.322–336 (2014).
- 10) M. Liang, Y. Xue, K. Xue, A. Yang: "Deep Convolution Neural Networks for Automatic Eyeglasses Removal", DEStech Trans. on Computer Science and Engineering (AIEA) (2017).
- 11) W. Shen, R. Liu: "Learning Residual Images for Face Attribute Manipulation", Proc. of Computer Vision and Pattern Recognition (CVPR), pp.8789–8797 (2017).
- 12) JINS. Inc, Megane On Megane, <https://jinsholdings.com/jp/ja/news/shibuya-parco> (2019).

- 13) T. Cootes, C. Taylor, D. Coopers, J. Graham: "Training Models of Shape from Sets of Examples", Proc. of British Machine Vision Conference, Vol. 9, pp.9–18 (1992).
- 14) F. Hill, S. Kelley, Computer Graphics Using OpenGL, 3rd ed., Pearson (2007).
- 15) M. Bertalmio, A. L. Bertozzi, G. Sapiro: "Navier-Stokes, Fluid Dynamics, and Image and Video Inpainting", Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Vol. 1 (2001).
- 16) A. Telea: "An Image Inpainting Technique Based on The Fast Marching Method", Journal of Graphics Tools, Vol. 9. No.1, pp.23–34 (2004).

(2019年12月9日 受付)

(2020年2月13日 再受付)



小林 巧

2018年 慶應義塾大学理工学部情報工学科卒業。現在、同大学院理工学研究科開放環境科学専攻博士前期課程に在学中。コンピュータビジョンやその複合現実感への応用に関する研究に従事。



杉浦 裕太

2013年 慶應義塾大学大学院メディアデザイン研究科博士後期課程修了。慶應義塾大学大学院メディアデザイン研究科特任助教、産業技術総合研究所特別研究員、慶應義塾大学情報工学科助教、2018年4月より、同専任講師。ヒューマンコンピュータインタラクションに関する研究に従事。博士(メディアデザイン学)。



斎藤 英雄 (正会員)

1992年 慶應義塾大学大学院理工学研究科電気工学専攻博士課程修了。同年 同大助手。その後、専任講師、助教授を経て2006年 慶應義塾大学理工学部情報工学科教授。この間、1997–1999年まで学術振興会海外特別研究員として、米国カーネギーメロン大学ロボティクス研究所に滞在し主に Virtualized Reality の研究に従事。2000年–2003年 JST さきがけ研究「情報と知」領域研究員兼務。2006–2012年 JST CREST 研究者代表。現在、主としてコンピュータビジョンとそのVR応用に関する研究に従事。博士(工学)。電子情報通信学会、日本バーチャルリアリティ学会フェロー、情報処理学会、IEEE シニア会員、映像情報メディア学会、計測自動制御学会会員。



上間 裕二

2008年 慶應義塾大学理工学部物理情報工学科卒業。2010年 同大学院メディアデザイン研究科前期博士課程修了。2013年 同大学院メディアデザイン研究科後期博士課程単位取得退学。2013–2015年 慶應義塾大学大学院研究員。2015年 株式会社ジェイアイエヌ(現:株式会社ジンス)入社。スマートアイウェアの研究開発に従事。