RGB-D カメラを使用した 顔モデルのリアルタイム生成・合成による顔向き補正

Real-Time Face Model Reconstruction and Synthesis Using an RGB-D Camera for Face Reorientation

正会員 山崎 俊彦[†], 中村 新芽^{†*}, 正会員 相澤 清晴[†]

Toshihiko Yamasaki[†], Ibuki Nakamura^{†*} and Kiyoharu Aizawa[†]

あらまし ビデオチャットや鏡アプリなど、画面外にカメラを置いて画面上に顔を表示するシステムでは、カメラとディスプレイの位置が異なるため正面を向いた映像を撮影して表示することが難しい。本研究では 1 台の RGB-D カメラから 2.5 次元形状を表す顔モデルを生成して、その顔モデルを適切に回転させることによって正面を向いた顔モデルを取得し、映像と合成することによって映像中の顔の向きをリアルタイムに補正する手法を提案する。従来の手法と異なり、RGB-D カメラの深度情報と一般の顔モデルを相補的に使用することでシステムの利便性と合成の自然さの両立を図る。加えて、22 名による主観評価実験によってその効果を確認する。

キーワード: RGB-D カメラ, DIBR, 視線一致, 顔向き補正, 仮想視点合成

1. ま え が き

ビデオチャットにおける顔向き補正や視線の一致は長年の課題である。しかし、ビデオチャットではディスプレイ上に映る相手の目を見て会話をするのに対し、カメラは画面の隅にあるため、視線が一致しない。それが臨場感を損なう要因の 1 つとなっている。また、近年ではスマートフォンの内向きカメラを鏡代わりに利用する人が増えているが、そこでも同じ問題が生じている。

この課題に対しては長年にわたって様々なアプローチが行われてきた。古くはハーフミラーなどを使用してディスプレイとカメラの光軸を物理的に一致させる研究が報告されている^{1)~3)}。しかし、このような装置は規模が大きいため、一般家庭での利用やスマホ利用などには不適である。一方、ディスプレイの周囲に複数のカメラを設置して Image Based Rendering (IBR)⁴⁾ によって映像を合成するものも提案されている^{5)~7)}。この手法は一般的なカメラを複数台用いるだけで実現できるものの、処理速度や画像品質など技術的な課題が多く実用化に至っていない。

深度方向の情報を元に画像を合成する技術、Depth Based Image Rendering (DIBR) を用いた手法も提案されている。すでに文献 8) においては、視線一致画像生成そのものが主眼ではないものの応用例の 1 つとして提案されている。近

年の RGB-D カメラの普及により DIBR 処理は一般的になり、より実用に近い研究がなされている。文献 9) では、1 台の RGB-D カメラを用いてポイントクラウドから顔付近の 2.5 次元形状を作成して回転させることで、顔の向きが補正された映像を得ている。また、DIBR 処理と併せて様々な人の正面を向いた審美性の高い画像を事前に用意しておき、辞書を構築してそれを元に顔の一部を補正する研究例もある¹⁰⁾。この手法は視線一致と顔の美化を統一的な最適化問題として解いているが、リアルタイム処理ができない問題がある。文献 9)10) に共通する問題として合成画像に対する評価が殆どなされていないことが挙げられる*。

本論文では文献 9) と同様に DIBR によってこの問題を解決しているが、下記の 2 点で先行研究と異なる。1 つめは解像度の高い RGB-D カメラを使用しながらも、処理の高速化や一般顔モデルの利用などにより 90fps で顔の向きを補正して視線一致を実現できる顔モデルのリアルタイム生成・合成システムを実現した点である。また、一般顔モデルの利用のためには使用開始までに 30 秒程度のユーザのインタラクションを伴う前処理が必要となるが、これが不要となるようポイントクラウドモデルと一般顔モデルのシームレスな切り替えを可能とした。2 つめは 22 名による主観評価を行った点である。これにより、処理を加えない場合に比べて画像の自然さは若干劣るものの、顔の向きと視線の一致に関しては評価が大きく向上し、またビデオチャッ

2014 年 9 月、年次大会で発表

2015 年 3 月 2 日受付, 2015 年 4 月 13 日採録

[†] 東京大学大学院情報理工学系研究科電子情報学専攻

(〒113-8656 文京区本郷 7-3-1, TEL 03-5841-6761)

* 現在は富士通株式会社 に所属

* 文献 10) では、会議開催中に合成画像の自然さに関する評価結果を示していたが、クラウドソーシングを用いた完全なる他人による評価であり、処理対象を知る人による評価はなされていない。

図 1 手法の概要. (a) RGB-D カメラから得られる情報で作成した顔モデル, (b) 顔の回転角を元に a を正面に向けた顔モデル, (c) b に対して顔周辺画像を元に着色したもの, (d) RGB-D カメラで撮影した RGB 画像, (e) c と d を合成するための縫ぎ目の最適化, (f) 合成結果

図 2 作成する顔モデル. (a) ポイントクラウド元データ, (b) ポイントクラウドの点を単純につないだモデル, (c) b の縦横 5 点相当を拡張しラプラシアン平滑化をかけたモデル, (d) 個人適応していない AAM, (e) 個人適応した AAM

トとしての利用にも許容されることを確認した. さらに, 処理対象が知り合いであるかどうかで主観評価がどのように変化するかも明らかにした. 本研究の一部は文献 11)12) にて発表済みであるが, 主観評価実験の規模などが異なる.

2. 提案手法

2.1 概要

提案手法の概要を図 1 に示す. 一台の RGB-D カメラをディスプレイの上か下に設置する. なお, ディスプレイのサイズは 24 インチで, 被写体とカメラの距離は 90cm であった. また, 顔とカメラの角度はディスプレイの上, 下においたときそれぞれ +13 度と -13 度であった.

本研究では RGB-D カメラには Microsoft 社の Kinect for Windows v2 を使用した. 提案手法の流れを説明する. 1) RGB-D カメラから得られるデータから顔の 3 次元形状を作成する. この顔モデル作成手法については 2.2 にて後述する. 2) 1 で得た顔モデルに顔追跡によって得た回転角をもとにディスプレイ正面を向くように回転を与える. 3) RGB-D カメラから得た RGB 画像の顔周辺領域をもとに 2 で作成した顔モデルに着色を行う. 4) 3 で作成した顔モデルと, RGB-D カメラで撮影した RGB 画像を合成するための縫ぎ目の最適化を行う. 5) 4 で求めた縫ぎ目によって合成を行い, 合成結果を得る.

2.2 顔モデル生成

適切な出力結果を得るため, 2 つの顔モデル生成手法を状況に応じて使い分ける. 1 つ目は RGB-D カメラから得ら

れるポイントクラウドの近接点を単純に繋いだポイントクラウドによる顔モデル (図 2(c)), 2 つ目は 13) による手法を使って, 図 2(d) に示す一般顔モデル (Active Appearance Model, AAM) を個人適応したモデル (図 2(e)) である.

(1) ポイントクラウドによる顔モデル

RGB-D カメラから得られるのは単なるポイントクラウドであり (図 2(a)), メッシュ情報は付与されていない. 今回取得したポイントクラウドは RGB-D カメラから取得したことがわかっているので, 隣り合う点は面を形成するとして隣接した点同士を三角メッシュで繋いでいった. ただし, 三角メッシュの 1 辺が 1cm を超えるものについてはオクルージョンなどの領域であり面を形成していないと考え, メッシュを描画しなかった. その結果が図 2(b) となる.

この手法では空間上の点が取得できないなどポイントクラウドに穴があった場合にそこにメッシュが貼られないことや周囲の欠損部に顔画像が貼られず目立つことがあるため, 縦方向と横方向にメッシュを拡張することで穴を埋め, 顔画像を貼る領域を拡張する. まず, 顔領域の縦方向と横方向の物理的な距離をポイントクラウドの点数で割ることによって, それぞれの点の間隔が物理的にどれだけの距離があるかを求める. その間隔に基づいて, 周囲を 5 点分だけ縦横に拡張する. その際, 新しい点を追加するため, その点の距離情報は不明である. そのため, 新しく追加する点はその点から縦横に 5 点ずつの範囲, 計 121 点のうち存在する点の距離情報の平均とした. 加えて, 凸凹が目立つため全体に式 (1) を元にラプラシアン平滑化をかけた. N は v と v に近接する点の数, w_j は v と v に近接する点の一覧, $\bar{v}$ は新しい v の位置を示す. 近接する範囲はある点から縦横に 5 点ずつの範囲とする. 結果として, ポイントクラウドによる顔モデル (図 2(c)) が得られる.

$$\bar{v} = \frac{1}{N} \sum_{j=1}^N w_j \quad (1)$$

(2) AAM によって作成する個人適応顔モデル

文献 13) では AAM と呼ばれる顔モデルを提案している. AAM によって顔の位置や角度などの顔位置の追跡に加え, 目や口の開閉や顎の動きなどの表情の変化もリアルタイムで 2.5 次元形状で表すことを可能にしている. さらに, 個人適応することによって鼻の高さや目の大きさなどの個人の顔の形状も反映することが可能になっている. 個人適応する前と個人適応した後の AAM によるモデルをそれぞれ図 2(d), (e) に示す. この AAM では, 形状パラメータと呼んでいる顔の形状を個人に適応させるための値と, アニメーションパラメータと呼んでいる顔の表情を表す値によって顔のモデルを表現している. この内アニメーションパラメータは最初から取得することが可能だが, 形状パラメータの取得には個人適応するため顔の正面, 左向き, 右向き, 上向きの画像を撮影する必要がある. これは即座にビデオチャット等を始めたい人にとって不便である. 逆に, いっ

たん個人適応がおわると描画はポイントクラウドを用いたものに比べて大幅に高速化できる。そこで、システム利用中の形状パラメータが揃うまでの間は上記ポイントクラウドによる顔モデルを用い、形状パラメータが揃った段階で自動的に個人適応後の AAM 顔モデルを用いることとする。なお、個人適応を行わない顔モデルは 3 に示すとおり合成画像が不自然になるので使用に適さない。

2.3 継ぎ目の最適化と画像の合成

合成するべき正面を向いた顔画像は、2.2 で得た顔モデルを AAM の顔追跡で得た顔の回転角を元に回転させ、その顔モデルに対して RGB 画像をテクスチャマッピングすることで得ることができる。なお、顔の向きに関するパラメータは、ポイントクラウドの顔モデルを用いている場合は AAM による一般顔モデルから、AAM による個人適応後の AAM 顔モデルを用いている場合にはその顔モデルから得ることができる。しかし、合成する際に人間の顔をただ楕円形で切り取り合成を行うと継ぎ目が目立つ。文献 9) の手法を参考に、輝度値を利用したエネルギー最小化を行い目立たない継ぎ目を選択する。継ぎ目 S は次のエネルギー E_{TOTAL} を最小化するように選択する。

$$E_{TOTAL} = \sum E(p_i), \forall p_i \in S \quad (2)$$

$$E(p) = \sum \|I_s(q_i) - I_o(q_i)\|_2^2, \forall q_i \in B(p) \quad (3)$$

I_s, I_o はそれぞれ元画像と合成画像の輝度値で、 $B(p)$ は座標 S を中心とした 5×5 のブロックである。継ぎ目を推定した後、合成する境界が目立たないようにガウシアンフィルタによって境界をぼかし、アルファブレンディングによって滑らかな合成を行う。

3. 合成結果

それぞれの手法によって合成した結果は図 3 のようになった。本稿に載せた結果は画像だが、実際には動画である。画像ではわからないが、ポイントクラウドによる手法(図 3(a))では RGB-D カメラから得られるデータにはどうしてもノイズが生じるために動画にした時に少しちらつきがあり、自然な表示の妨げになっている。個人適応する前の AAM(図 3(b))では目や唇が細く表示されており顔の印象が変化してしまっている。個人適応した後の AAM(図 3(c))は他の手法と比べると優れていて、環境によっては加工していることが分かりにくいほどである。

4. 評価実験

4.1 概要

自然な顔向き補正を行えているか、評価実験を行った。本研究で求められる自然さなどの指標は機械的に判定するのは難しいため、主観評価実験を行うこととした。主観評価実験として次の 5 つを行う。加工後の画像としてはすべて個人適応した AAM による出力結果を用いた。

実験の説明における自然さとは画像・動画が自然である

図 3 顔モデルの作成手法ごとの合成結果の比較。(a) ポイントクラウドによって作成したモデルによる結果、(b) 個人適応する前の AAM によって作成した結果、(c) 個人適応した AAM によって作成した結果

ように見えるかどうか、顔向き度合いとは顔が評価者の方を向いているかどうか、視線一致度合いは視線があっているように見えるかどうか、ビデオチャット適正はビデオチャットに適しているかどうかに対する評価をそれぞれ示す。

- 実験 A: 加工後または加工前の画像 1 枚をランダムな順序で計 20 枚提示し、それぞれ自然さ、顔向き度合い、視線一致度合い、ビデオチャット適正を 4 段階で評価する。

- 実験 B: 加工後または加工前の動画 1 つをランダムな順序で計 10 本提示し、それぞれ自然さ、顔向き度合い、視線一致度合い、ビデオチャット適正を 4 段階で評価する。

- 実験 C: 加工後と加工前の画像ペア 1 組をランダムな順序で計 10 組提示し、それぞれ顔向き度合い、視線一致度合い、ビデオチャット適正を 4 段階で評価する。

- 実験 D: 加工後と加工前の動画ペア 1 組をランダムな順序で計 5 組提示し、それぞれ顔向き度合い、視線一致度合い、ビデオチャット適正を 4 段階で評価する。

- 実験 E: 同一人物を仰角 30 度から -30 度まで 5 度刻みで角度をつけて撮影し、それぞれ自然さ、顔向き度合い、視線一致度合いを 4 段階で評価する。

実験 A,C で使用する加工前画像は男女 5 人から 10 枚撮影し、RGB-D カメラをディスプレイ上に設置して撮影したものを 5 枚、ディスプレイ下に設置して撮影したものを 5 枚とする。画像の加工前後で合計 20 枚となる。実験 B,D で使用する加工前動画は男女 4 人から 5 本撮影し、RGB-D カメラをディスプレイ上に設置して撮影したものを 2 本、ディスプレイ下に設置して撮影したものを 3 本とする。動画の加工前後で合計 10 本となる。20~30 代の男女 22 名を被験者として実験はすべてウェブ上で行う。今後、簡単のために画像・映像中で加工対象となった人たちを話者、主観評価に参加した人たちを評価者と呼ぶ。なお、評価者のうち 18 名は話者らを直接的に知りまた日常的に接している人たちであり、残りの 4 名は話者らと面識がまったくない。画像・動画の提示と 4 段階の評価は専用のアンケート用ウェブサイトを用いて行った。

4.2 実験 A

静止画像を比較せず単体で見える実験である。話者らと面識がある 18 名の評価者による実験結果を図 4 に示す。すべての場合において、加工前に比べ、加工後の自然さは損なわれている。対して、ビデオチャットに適しているかど

(a) すべての平均

(b) カメラをディスプレイ上に配置したものの平均

(c) カメラをディスプレイ下に配置したものの平均

図 4 話者らと面識がある 18 名の評価者による実験 A の結果 (白が加工前, 黒が加工後, 1 が低評価, 4 が高評価). 図中, **は 1%水準の有意差, n.s. は有意差なし

図 5 話者らと面識のない 4 名の評価者による実験 A の結果 (白が加工前, 黒が加工後, 1 が低評価, 4 が高評価). すべての平均のみを示す

うかについてはそれほど変化はない. 加えて, 視線一致度合いはすべての場合において, 視線があっていないと感じる人が多い状態から視線があっていると感じる人が多い状態になっている. Welch の t 検定を行ったところ自然さは 1%水準で優位に劣化していたが, 逆に顔向き度合いと視線の一致度合いでは 1%水準で優位に改善が見られた. また, ビデオチャット適正については処理の前後で優位さはなかった. 特にカメラをディスプレイ下に配置しているものに関しては効果が大きく, 顔の向きについても顔が評価者の方を向いているように変化している. カメラをディスプレイ上部に設置した場合はビデオチャット適正が 5%水準で優位に悪化した.

話者らと面識のない 4 名の評価者による結果を図 5 に示す. 自然さ, 顔向き度合い, ビデオチャット適性の評価が, 面識のある評価者と比べてやや高くなる傾向にあることがわかる. これは, 話者らの元の顔を知らないため「顔として適切かどうか」のみを評価基準としているためと思われる.

4.3 実験 B

動画を比較せず単体で見る実験である. 話者らと面識がある 18 名の評価者による実験結果を図 6 に示す. 実験 A ほどではないが, 自然さは損なわれている. ビデオチャット適性は実験 A と同様に加工前後であまり変化していない. カメラをディスプレイ上に配置した場合において, 他の実験と比べて顔向き度合い, 視線一致度合いが大きく改善されている. Welch の t 検定で行ったところ実験 A と

(a) すべての平均

(b) カメラをディスプレイ上に配置したものの平均

(c) カメラをディスプレイ下に配置したものの平均

図 6 話者らと面識がある 18 名の評価者による実験 B の結果 (白が加工前, 黒が加工後, 1 が低評価, 4 が高評価). 図中, **は 1%水準の有意差, n.s. は有意差なし

図 7 話者らと面識のない 4 名の評価者による実験 B の結果 (白が加工前, 黒が加工後, 1 が低評価, 4 が高評価). すべての平均のみを示す

図 8 話者らと面識がある 18 名の評価者による実験 C の結果 (1 が加工後の方が評価が高い, 4 が加工前の方が評価が高い). すべての平均のみを示す

図 9 話者らと面識がある 18 名の評価者による実験 D の結果 (1 が加工後の方が評価が高い, 4 が加工前の方が評価が高い). すべての平均のみを示す

同様に自然さは 1%水準で優位に劣化, 顔向き度合いと視線の一致度合いでは 1%水準で優位に向上していた. ビデオチャット適正については処理の前後で優位さはなかった. 実験 A と異なる点はカメラをディスプレイ上部に設置した場合においてもビデオチャット適正が優位に劣化しなかった点である. これは動画にすることで静止画では気になったような細かい差異に気づきにくくなったためと思われる.

話者らと面識がない 4 名の評価者による実験結果 (図 7) は, 実験 A とほぼ同じ傾向を示している.

4.4 実験 C

静止画像の加工前後で比較する実験である. 実験結果を図 8 に示す. 自然さは加工後の評価が悪いことが予想できるため項目に入れていない. カメラをディスプレイ上に配置したものは, 平均で見ると加工前後で評価がそれほど変わっていない. 画像ごとに個別に結果を見ると, 評価の良

図 10 話者らと面識がある 18 名の評価者による実験 E の結果 (破線が加工前, 実線が加工後, 1 が低評価, 4 が高評価)

図 11 話者らと面識がない 4 名の評価者による実験 E の結果 (破線が加工前, 実線が加工後, 1 が低評価, 4 が高評価)

い画像と評価の悪い画像に大きく分かれてしまっていて、平均的な評価になってしまっているようであった。画像を見ると、顔をそらしているように感じられる人の画像は加工前の方が評価が良かった。姿勢によっては処理を別にする事で結果が向上することも考えられる。カメラをディスプレイ下に配置したものについてはビデオチャット適正は加工前後でそれほど変わらず、顔向き度合い、視線一致度合いは加工後のほうが評価が高くなっている。

4.5 実験 D

動画の加工前後で比較する実験である。実験結果を図 9 に示す。この実験はあまり動画ごとの差異は見られず、平均的に加工後のほうが項目ごとの評価が高かった。

4.6 実験 E

静止画像を単体で見て角度への頑健さをみる実験である。実験結果を図 10 に示す。まず自然さは加工前のほうが評価

図 12 いくつかの加工前後の比較画像 (すべて左が加工前, 右が加工後)

表 1 処理速度, 単位はミリ秒

処理項目	ポイントクラウド	個人適応後の AAM 顔モデル
画像の読み込み	1.81	1.85
顔認識	1.66	1.55
顔モデルの作成	3.43	—
モデルの描画	9.52	1.18
描画結果の画像化	2.13	2.25
継ぎ目の最適化	2.51	2.48
画像の合成	1.05	1.06
結果表示	0.45	0.46
合計	23.56	10.83

が良い。加工後は 5 ～ -10 度の間では良い評価を得られている。顔向き度合い、視線一致度合いは加工後のほうが良い評価を得られる範囲が広がっている。特に、0 ～ -10 度は加工後は大きく改善されている。加工前は視線が一致する範囲が 5 ～ 0 度だけだったのに対し、加工後は 10 ～ -10 度まで広がっている。自然さもふまえて考えると、0 ～ -10 度の範囲ではこの手法を使えば正面を向いて視線の合う自然な映像を出力可能になる。

話者らと面識がない 4 名の評価者による実験結果 (図 11) をみると、処理前より提案手法に対する評価が高い範囲が広がる傾向にあることがわかる。前述の通り、これは「顔として適切かどうか」のみを評価基準としているためと思われる。

4.7 様々な出力結果

主観評価実験とは別に、結果の例としていくつかの画像を図 12 に示す。図 12(a) では自然さを損なわず顔向き度合い、視線一致度合い、ビデオチャット適性のすべてが向上した。図 12(b) では自然さは自然な状態と加工した状態の中間となったが、顔向き度合い、視線一致度合いが大きく向上し、通話品質も若干良くなった。図 12(c) は失敗例で、眉毛の形が変わるなど表情が変わってしまっている上、ビデオチャット適性が悪化してしまった。図 12(d) も失敗例で、元々ビデオチャットによく適していて顔向き度合い、視線一致度合いなどすべて高評価だったが、提案手法を行うことですべて評価が低くなった。

4.8 処理速度

一般的なノートパソコン (CPU: i7-3740QM, GPU: Intel HD Graphics 4000) を用いて 2000 フレーム (約 66 秒) を撮影したときの平均処理時間を表 1 に示す。ポイントク

図 13 カメラの仰角と処理結果

ラウドを用いた顔モデルでも 30fps 以上の速度で動作し、個人適応後の AAM 顔モデルでは 90fps 程度実現できることがわかる。

4.9 考察

先行研究⁹⁾¹⁰⁾においては、カメラをディスプレイ下部に設置した場合のみを評価しているが、本実験によりカメラをディスプレイ上部に設置した場合においては処理結果に対する主観評価値が下がる傾向が示された。処理の具体例を図 13 に示す。この現象は、顔の形状モデルと取得したテクスチャ画像の対応がうまくとれず、目や口元に大きな歪みが生じるためであることがわかる。現在、パソコンやスマートフォンに内蔵されているカメラはディスプレイ上部についていることが多いが、今後 RGB-D カメラを用いた顔向き補正を行う場合はディスプレイ下部に設置した方がよいことが示唆される(ただし、ノートパソコンのようにディスプレイ上部にカメラをつけても位置が顔よりも下に来る場合はこの限りではない)。

本手法は、眼鏡をかけている場合や前髪が極端に目や眉にかかっている場合にうまく処理ができない。また、会話中に指などを顔に近づけると歪みが生じることがある。これは DIBR を用いるすべての手法に共通する問題である。

5. む す び

本論文では、1 台の RGB-D カメラから顔モデルをリアルタイム (30fps 以上) で生成し、その生成結果を RGB 映像と合成することで顔向きが補正された RGB 映像を出力するビデオチャットや鏡アプリに向けた顔向き補正手法を提案した。

22 名による主観評価実験では加工前後の画像を独立に見た場合 (実験 A)、加工前後の画像を見比べた場合 (実験 B)、加工前後の動画を独立に見た場合 (実験 C) でビデオチャットへの適性を保ったまま顔の向きを補正できたことを示した。加工前後の動画を見比べた場合 (実験 D) には加工後のほうが顔向きが補正されており、ビデオチャットへの適性も優れていることを示した。本手法が適用できる角度の範囲も実験によって示した (実験 E)。また、評価者が話者の知り合いかどうかで主観評価がどのように変化するかにについても実験的に示した。

謝 辞

顔の元画像・映像および処理後の画像・映像の利用、主観評価実験への参加においては実験参加者らに趣旨を説明した上で同意を得た上で行っている。本実験は東京大学情報理工学系研究科の倫理委員会の承認を得て行っている。本研究の一部は科学研究費助成事業 (26700008)、および Microsoft IJARC core 10 の支援プログラムを受けて行われた。

〔文 献〕

- 1) H. Ishii and M. Kobayashi. ClearBoard: a seamless medium for shared drawing and conversation with eye contact. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 525–532, 1992.
- 2) K. Okada, F. Maeda, Y. Ichikawa and Y. Matsushita. Multi-party videoconferencing at virtual social distance: Majic design. CSCW, pp. 385–93, 1994.
- 3) ディシルバ, 田原, 相澤, 羽鳥. 複数人同士の視線一致を可能とするテレビ会議システム (mpec). 電子情報通信学会技術研究報告, **95**, **181**, pp. 15–22, 1995.
- 4) H. Y. Shum, S. C. Chan, and S. B. Kang. Image-based rendering. Springer, 2007.
- 5) R. Yang and Z. Zhang. Eye gaze correction with stereovision for video-teleconferencing. ECCV, pp. 479–494, 2002.
- 6) A. Criminisi, J. Shotton, A. Blake, and P. H.S. Torr. Gaze Manipulation for One-to-one Teleconferencing. ICCV, 2009.
- 7) 味八木, 山崎, 相澤. 遠隔視線一致コミュニケーションのための実時間 IBR 処理. 画像符号化・映像メディア処理シンポジウム, P-2.7, pp. 27–28, 2004.
- 8) 日浦慎作, 佐藤宏介, 井口征士. リアルタイム距離画像センサによる動物体の多視点観測. 電子情報通信学会技術研究報告. CS, 通信方式, **15**, pp. 35–40 (1995)
- 9) C. Kuster, T. Popa, J.C. Bazin, C. Gotsman, and M. Gross. Gaze correction for home video conferencing. ACM Trans. Graph., **31**, **6**, pp. 174:1–174:6 (2012)
- 10) X. Liu, G. Cheung, D. Zhai, D. Zhao, H. Sankoh and S. Naito. Joint gaze correction and face beautification for conference video using dual sparsity prior. ICIP, pp. 4697 – 4701 (2014)
- 11) 中村新芽, 山崎俊彦, 相澤清晴. デプスカメラを用いた個人向けビデオチャットにおけるロバストな視線補正. 映像年次大会, 12-2 (2014)
- 12) 中村新芽, 山崎俊彦, 相澤清晴. RGB-D カメラを用いた顔モデルのリアルタイム生成・合成 ~ビデオチャットや鏡アプリのための顔向き補正~. 電子情報通信学会パターン認識・メディア理解研究会 (PRMU), PRMU2014-157, pp. 227–232 (2015)
- 13) N. Smolyanskiy, C. Huitema, L. Liang, and S. Eron Anderson. Real-time 3d face tracking based on active appearance model constrained by depth data. Image and Vision Computing, **32**, **11**, pp. 860–869 (2014)

やまさき としひこ
山崎 俊彦 東京大学工学部電子工学科卒業。東京大学工学系研究科電子工学専攻修士。博士 (工学)。現在、東京大学大学院情報理工学系研究科電子情報学専攻 准教授。大規模マルチメディアデータ処理、3 次元映像処理、物体認識・機械学習などの研究に従事。正会員。

なかむら いぶき
中村 新芽 東京大学工学部電子情報工学科卒業。東京大学大学院情報理工学系研究科電子情報学専攻修士課程修了。現在、富士通株式会社に所属。

あいざわ きよはる
相澤 清晴 1983 年 東京大学工学部電子工学科卒業。1988 年 同大学院博士課程修了。工博。東大工学部助手、講師、助教授を経て、2001 年より、教授。現在、大学院情報理工学系研究科教授。この間、画像・マルチメディア処理、ライフログ、3 次元映像、漫画画像処理等に関する研究に従事。学術会議連携会員。本学会フェロー。