

テレビ会議話者間の視線一致知覚範囲を考慮した 目領域画像合成型視線補正法

井上 卓弥[†] 高橋 友和[†] 平山 高嗣[†] 川西 康友[†] 出口 大輔^{††}
井手 一郎[†] 村瀬 洋[†] 黒住 隆行^{†††} 柏野 邦夫^{†††}

[†] 名古屋大学 大学院情報科学研究科 〒464-8601 愛知県名古屋市千種区不老町

^{††} 名古屋大学 情報戦略室 〒464-8601 愛知県名古屋市千種区不老町

^{†††} 日本電信電話株式会社 NTT コミュニケーション科学基礎研究所
〒243-0198 神奈川県厚木市森の里若宮 3-1

E-mail: †inouet@murase.m.is.nagoya-u.ac.jp, ††takahashi@gifu.shotoku.ac.jp, ††ddeguchi@nagoya-u.jp,
†{hirayama,kawanishi,ide,murase}@is.nagoya-u.ac.jp, †††{kurozumi.takayuki,kashino.kunio}@lab.ntt.co.jp

あらまし 近年、ウェブカメラなどの普及により、テレビ会議を容易に行なうことができるようになった。しかし、一般的にカメラはディスプレイの外部に設置されており、ユーザはディスプレイ上の相手のユーザを見るため、視線が一致しないという問題が発生する。この問題を解決するために、様々な手法が提案されているが、それらの手法の多くは特定のセンサを必要とする。そのため、提案手法では、単一のカメラのみを用いて、対話相手が視線一致を知覚する範囲（視線一致知覚範囲）をユーザが注視している場合にのみ、ユーザの目領域をカメラを直視している画像に変換することで視線の一致した画像を生成する。

キーワード テレビ会議, 視線一致, 注視点分類

A method for gaze correction between video conference participants that synthesizes eye areas image within the perceptual range of eye contact

Takuya INOUE[†], Tomokazu TAKAHASHI[†], Takatsugu HIRAYAMA[†],

Yasutomo KAWANISHI[†], Daisuke DEGUCHI^{††}, Ichiro IDE[†], Hiroshi MURASE[†], Takayuki

KUROZUMI^{†††}, and Kunio KASHINO^{†††}

[†] Graduate School of Informaion Science, Nagoya University
Furo-cho, Chikusa-ku, Nagoya-shi, Aichi, 464-8601 Japan

^{††} Information Strategy Office, Nagoya University
Furo-cho, Chikusa-ku, Nagoya-shi, Aichi, 464-8601 Japan

^{†††} NTT Communication Science Laboratories, Nippon Telegraph and Telephone Corporation
3-1 Morinosato-Wakamiya, Atsugi-shi, Kanagawa, 243-0198 Japan

E-mail: †inouet@murase.m.is.nagoya-u.ac.jp, ††takahashi@gifu.shotoku.ac.jp, ††ddeguchi@nagoya-u.jp,
†{hirayama,kawanishi,ide,murase}@is.nagoya-u.ac.jp, †††{kurozumi.takayuki,kashino.kunio}@lab.ntt.co.jp

Abstract Recently, the spread of Web cameras has facilitated video-conferencing. Since a Web camera is usually located outside the display while the user looks at his/her partner in it, there is a problem that they cannot establish eye contact with each other. Many researchers have proposed methods to solve this problem, but most of them required specific sensors. We proposed a method which uses only one camera and transforms the user's eye areas in an image to his/her eye image with a straight gaze to the camera only when the user's gaze falls in a range that the partner would perceive eye contact, which is called *perceptual range of eye contact*.

Key words Video conferencing, eye-contact, gaze classification

1. まえがき

近年、ウェブカメラなどの普及により、テレビ会議を容易に行なうことができるようになってきた。しかし、多くのユーザは対話相手と互いに視線が一致しないため不自然さを感じている。これは、カメラを対話相手の目と同じ位置に設置することが困難であるためである。そのため、テレビ会議における視線一致の重要性について調査研究がなされており [1]、自然なコミュニケーションを可能にするために視線を補正する必要がある。

一般にテレビ会議を行なう際に、ユーザは常に対話相手の目を注視しているわけではない。ユーザが視線を一致させようと対話相手の目を注視している場合のみ、視線を一致させるように視線を補正することで、実際の対話と同じタイミングで視線一致を感じることができ、自然なコミュニケーションが可能になると考えられる。

また、近年、ノートパソコンや携帯電話でテレビ会議を行なうようになってきた。しかし、それらには一般的に単一の可視光カメラしか装備されていないため、我々は単一の可視光カメラのみを用いる手法を提案する。具体的には、対話相手が視線一致を知覚する範囲をユーザが注視している場合のみ、カメラを直視するようにユーザの目領域の画像を変換することで視線が一致した画像を生成する(図 1)。本手法では、対話相手が視線一致を知覚する範囲を視線一致知覚範囲と呼ぶ。

魚野らは視線一致知覚範囲がおおよそ 4 度であることを示した [2]。また、Anstis によると、対面で頭部を回転させないという条件において、対話相手の注視方向が自己の目の位置から 10 度以上離れている場合には、実際の注視方向よりもさらに大きな角度方向を相手が注視していると知覚する [3]。一方、注視方向が 4 度以下である場合には、実際の注視方向よりも小さな角度を注視していると知覚することが知られている。このことから、人間は自分の目の周辺への相手の視線には敏感であるが、それ以外の視線方向に対してあまり敏感ではないことがわかる。したがって、提案手法では、ユーザが視線一致知覚範囲を注視している場合のみ目領域を変換し、それ以外の場合には変換を行わず元画像を送信する。そのために 2 つの技術の実現が必要である。

- (1) 画像から視線一致知覚範囲の注視を検出する技術
 - (2) 目領域を変換し、視線が一致した顔画像を生成する技術
- 想定する利用状況は図 2 に示す通りであり、ユーザと対話相手の双方で目領域を変換された画像を送信することで視線一致が実現できる。

以下、2 節で関連研究について述べ、3 節で提案手法について述べる。4 節で提案手法の有効性について調査した実験と結果について述べ、5 節で視線が一致しているか否かを判定する処理の精度を評価した実験と結果を述べる。最後に、6 節でまとめと今後の課題について述べる。

2. 関連研究

テレビ会議における視線が一致しない問題を解決する方法と

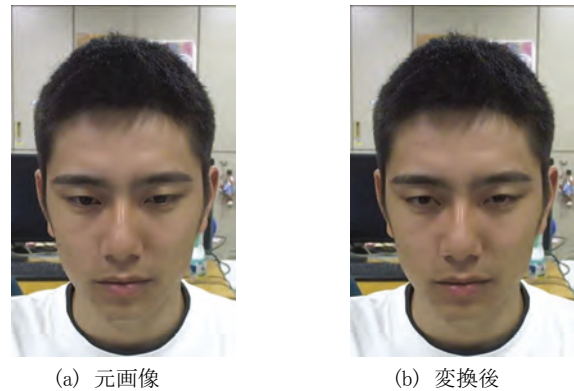


図 1 提案手法で変換した画像の例

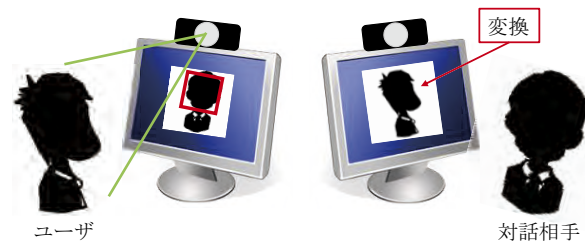


図 2 想定する利用状況

してハードウェアによる手法とソフトウェアによる手法が提案されている。ハードウェアによる手法として、Kollarists らはハーフミラースクリーンを利用する手法を提案している [4]。しかし、この装置はとて大きく、設置に時間がかかるという問題がある。ソフトウェアによる手法として、複数のカメラを利用する手法と単一のカメラを利用する手法が存在する。

Yang らは 2 台の可視光カメラから得られた顔画像を特徴点を用いて密に対応付けを行ない、View Morphing を用いて合成することで正面から撮影したような顔画像を生成する手法を提案している [5]。この手法は、2 台のカメラから得られた顔画像間の特徴点の密な対応付けのために、様々なアピアランスの変化に頑健で高精度な特徴点抽出が必要である。また、Kuster らは可視光カメラと深度カメラの 2 台のカメラを利用した手法を提案している [6]。この手法では、顔全体を適切に 3 次元変換することで視線が一致した画像を生成する。しかし、顔画像が顎の位置を基準に合成されるため、額の大きさが変わってしまう可能性がある。

一方、Giger らは単一カメラと 3 次元顔モデルを併用した手法を提案している [7]。この手法では、3 次元顔モデルを生成する際に、深度カメラが必要である。Yip はアフィン変換と目のモデルを利用して視線を一致させる手法を提案している [8]。この手法では、単一のカメラのみを使用しているが、準備段階でカメラをディスプレイの中心に設置する必要がある。これらの手法では、テレビ会議を行う前に、ユーザが他のセンサを利用したり、特定の位置にカメラを移動させる必要があり、ノートパソコンや携帯電話で使用する事が困難であると考えられる。それに対して、Solina らは単一のカメラのみを用いて視線が一致するように画像全体を水平軸のまわりに射影変換する手法を

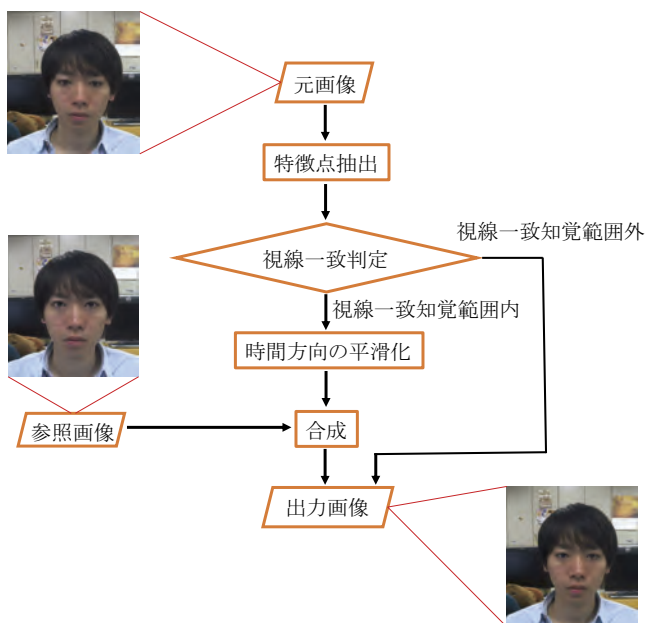


図3 提案手法の処理の流れ

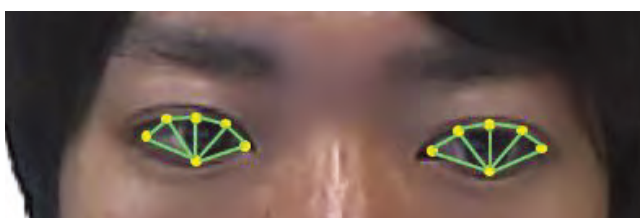


図4 抽出された特徴点と3角形パッチ分割

提案している [9]. この手法では、3次元構造を考慮せずに画像全体を回転させているため、顔画像が歪んだり、視線一致を感じる適切な顔画像を生成できない場合がある。

3. 視線一致知覚範囲を考慮した目領域画像合成型視線補正法

図3に提案手法の処理の流れを示す。まず、元画像から特徴点を抽出する。次に、ユーザが視線一致知覚範囲を注視しているか否か元画像から判定する。もしユーザが範囲内を注視していると判定された場合、参照画像を用いて目領域を変換する。

提案手法では、ユーザがカメラを直接注視した参照画像を1枚、および、視線一致判定に利用する複数の学習用画像を撮影する必要がある。

3.1 特徴点抽出

提案手法では、Saragihらの顔追跡手法[10]を利用して特徴点を抽出する。この手法では、顔全体から65点の特徴点を抽出することが可能であり、本研究では、図4に示すように、左右の目の輪郭の6点の特徴点を利用する。この顔追跡手法では、顔の同じ位置から安定して特徴点を抽出できない。この抽出精度が目領域の変換に影響を与え、動画にした際に不自然に感じる場合がある。この問題を避けるために、平滑化を行なう。具体的には、連続するフレーム間の特徴点の距離がしきい値未満の場合、前フレームと特徴点の位置がほとんど変化していな

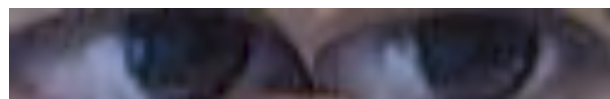


図5 目領域画像の例

いとみなし、現在のフレームの特徴点の代わりに前フレームの特徴点の位置を使用する。連続するフレーム間の特徴点の距離の2乗和 d は以下のように定義される。

$$d = \sum_{i=1}^6 \|\mathbf{x}_i^{(t-1)} - \mathbf{x}_i^{(t)}\|^2 \quad (1)$$

ここで $\mathbf{x}_i^{(t)}$ は t 番目のフレームの i 番目の特徴点の位置である。

3.2 視線一致判定

視線一致判定はユーザが視線一致知覚範囲を注視しているか否かを判定する処理である。範囲内を注視している場合、視線の補正を行ない、それ以外の場合には元画像をそのまま送信する。

我々が提案する視線一致判定は学習段階と識別段階の2つの段階からなる。学習段階では、学習用画像を用いて識別器を構築し、識別段階では、ユーザが視線一致知覚範囲を注視しているか否かを判定する。学習用画像は各ユーザに対して事前に撮影しておく。具体的には、視線一致知覚範囲内やそれ以外の領域にまんべんなく点を表示し、ユーザがその点を注視している画像を撮影する。

3.2.1 学習段階

学習段階では、学習用画像から目領域を切り出し、画像特徴を抽出し、識別器を構築する。

● 目領域画像の切り出し

まず、視線一致知覚範囲を注視している画像 (positive) とそれ以外を注視している画像 (negative) を収集する。次に、6つの特徴点の外接矩形で目領域を切り出し、左右の領域を連結して、図5に示すような目領域画像とする。この際、平行移動や拡大縮小などを加えて目領域画像の学習パターンを増やす。この処理をすべての学習用画像に対して行ない、各目領域画像のサイズをすべての目領域画像の外接矩形の平均の大きさに正規化する。

● 特徴抽出と識別器の構築

黒目の輪郭部分に最も特徴が表れやすいと考えられるため、目領域画像からエッジ特徴であるHOG (Histograms of Oriented Gradients) [11] を抽出する。照明変動に頑健にするため、特徴を抽出する前にヒストグラム均等化を行なう。識別器には、一般的に2クラス識別問題に対して高い性能を持つとされるSVM (Support Vector Machine) 識別器を用いる。

3.2.2 識別段階

入力画像に対し、学習段階と同様の特徴抽出を行ない、識別器を用いて視線一致知覚範囲を見ているか否かを判定する。その際、前処理として特徴点間の距離を利用して目の開閉度を計算し、しきい値よりも小さい場合は、予め見ていないと判定する。これは、まばたきにより、目が完全に開ききっていない状態では、ユーザは特定の場所を注視していないと考えられるた

めである。

3.3 時間方向の平滑化

提案手法では、視線一致判定をそれぞれのフレームに対して独立に行なうため、視線一致判定が時間方向に対して不安定である可能性がある。そこで、範囲内と範囲外の2値の判定に対して時間方向の平滑化を適用する。具体的には、現在のフレームと前後2フレームの5フレームを用いて多数決を行い、現在のフレームの視線一致判定とする。

3.4 合成

入力画像の目領域に対して参照画像の目領域を合成する。参照画像と入力画像とで目の形状が多少異なる場合でも、その変形に対応するために、合成には3角形パッチ分割を利用する。3角形パッチ分割は、前の処理で抽出した6点の特徴点を使用する。まず、6点の特徴点を用いて、図4に示すように入力画像と参照画像中の目領域を4つの3角形に分割する。次に、分割した各3角形において、参照画像の3角形をアフィン変換により変形し、入力画像中の対応する3角形の形状にあわせて合成する。この際、合成した画像がより自然に見えるように、アルファブレンディングを適用する。

4. 視線補正の評価実験

提案手法の評価のためには、提案手法により変換した映像に対して、対話相手が視線一致を感じることができると否かを評価することや、適切なタイミングで視線を補正できるかを評価する必要がある。そのため、4.1節では、提案手法や比較手法で変換された様々な映像を被験者が評価することにより視線一致の度合いを評価する。4.2節では、4種類の映像を1枚ずつ注意深く見てもらうことで、被験者が視線一致を感じるタイミングを評価する。

4.1 視線一致の度合いの評価

提案手法で変換された映像に対して被験者が視線一致を感じるかどうかを評価した。1,920×1,200ピクセルの解像度をもつ24インチのディスプレイ上部にカメラを設置し、1,280×980ピクセルの解像度で映像を撮影した。実験に先立って、5人の被験者が視線一致知覚範囲内や、それ以外の場所を注視する映像を撮影した。その映像に対して、提案手法と射影変換を利用した比較手法[9]を適用した。比較手法に関して、水平軸周りの回転角は20度に設定した。別の12人の被験者は、元の映像、提案手法と比較手法を適用することにより変換された映像の3つの映像を、“1”が視線が一致していない、“2”がやや視線が一致していない、“3”が視線が一致しているかしていないかわからない、“4”がやや視線が一致している、“5”が視線が一致しているという、5段階で評価した。

表1に被験者によって5段階評価された視線一致の評価値を示す。提案手法は元の映像と比較手法よりも高い値を示した。また、この結果に対してFriedman検定を行なった結果、有意水準1%で有意性があることが確認できた。

また、比較手法では視線一致の度合いがあまり改善していない。これは、この手法が3次元構造を考慮せずに画像全体を回転させているためである。これに対して、提案手法では元画像

表1 視線一致の評価値

	元映像	提案手法	比較手法
平均	1.60	3.65	1.70
分散	1.17	1.00	1.09

の目領域のみを変換することにより視線一致の度合いを改善することができた。

4.2 視線一致のタイミング評価

4種類の映像を1枚ずつ注意深く見てもらうことで、被験者が視線一致を感じるタイミングを評価した。この実験では、カメラをディスプレイの上部と正面に2台設置し、同時に撮影した。撮影した映像の長さは200フレームである。正面のカメラは現実の対面対話を想定した理想的な状況で顔画像を撮影するために設置した。被撮影者が正面のカメラや、それ以外の場所を注視する映像を撮影した。提案手法と比較手法をディスプレイ上部に設置されたカメラで撮影した映像に適用した。12人の被験者が図6に示すような映像を含む4種類の映像を評価した。4種類はそれぞれ、(a)ディスプレイの上部に設置されたカメラから撮影された一般的な状況の映像、その映像に(b)提案手法と(c)比較手法を適用した映像、(d)ディスプレイの正面に設置されたカメラから撮影された理想的な状況の映像である。テレビ会議をしていると想定して被撮影者と視線が一致していると感ずるか否かについて被験者から回答を得た。

図7に各フレームにおける被験者が視線一致を感じる割合を示す。正面のカメラから得られた映像における視線一致を感じる割合を真値とした場合、上部のカメラから得られた変換前の映像に対する評価と真値との誤差の平均は43.9%であったが、比較手法により変換した映像では39.3%に、提案手法により変換した映像では14.6%に減少した。

また、上部のカメラから得られた映像では、正面のカメラから得られた映像において視線一致を感じているタイミングでも視線一致を感じていない。また、比較手法により変換された映像では、元映像に比べて視線一致を感じる割合は改善しているものの、まだ多くの人が視線一致を感じていない。それに対して、提案手法により変換された映像では、正面のカメラから得られた映像において視線一致を感じているタイミングとほぼ同じタイミングに視線一致を感じていることがわかる。これにより、提案手法を適用したテレビ会議では、実際に対面したコミュニケーションと同じように自然に視線一致を感じる可能性があることを確認した。

提案手法において、5フレーム目や50フレーム目付近で視線一致を感じていない箇所が存在する。これは、実際にユーザが視線一致知覚範囲内を見ているのにもかかわらず、提案手法における視線一致判定で視線一致知覚範囲外であると判定され、目領域画像の変換が行なわれなかったためである。より自然なコミュニケーションのためには、視線一致判定の精度を改善することが必要である。そのためには、視線一致判定の性能を詳細に評価する必要がある。



図 6 4種類の映像に含まれる画像の例

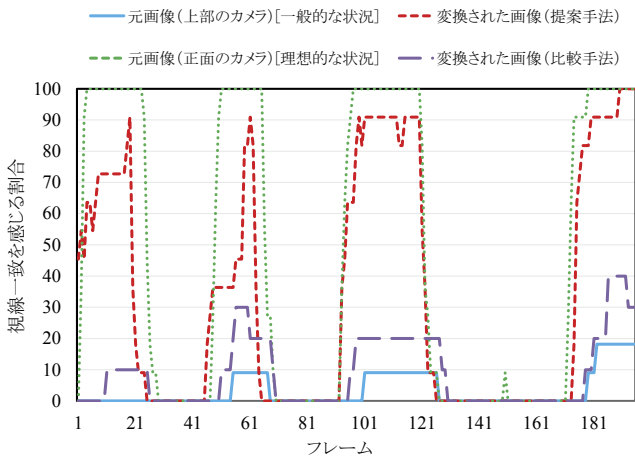


図 7 各フレームにおいて視線一致を感じた被験者の割合

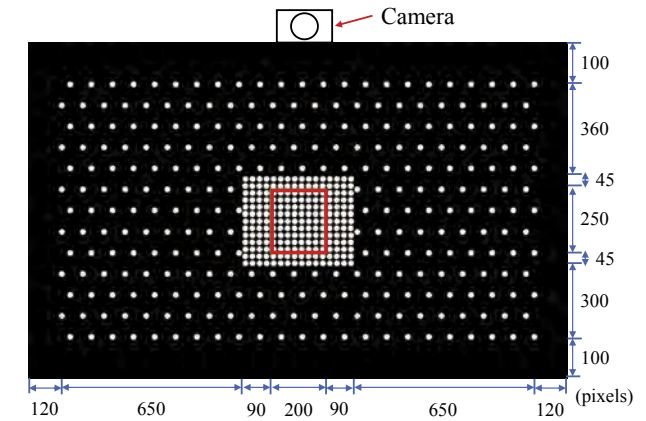


図 8 データセット収集のために使用した注視点

5. 視線一致判定の性能評価

視線一致判定の精度を評価するために、提案手法である HOG 特徴を使用する手法、輝度値を特徴として使用する手法、黒目の中心位置を特徴として使用する手法を比較した。

本実験は、4.1 節で行なった実験と同じ条件で行なった。1,920×1,200 ピクセルの解像度をもつ 24 インチのディスプレイ上部にカメラを設置し、1,280×980 ピクセルの解像度で画像を撮影した。5 人の被験者の画像はディスプレイから 50 cm 離れた位置で撮影された。図 8 に示すような 486 個の白点を 1 つずつ被験者に提示し、被験者が提示された点を注視した顔画像を各点で撮影した。撮影したデータセットの例を図 9 に示す。図 8 において、視線一致知覚範囲は赤の矩形で囲まれた領域として表され、魚野らの知見 [2] を参考に、中心からおおよそ 4 度の領域としている。視線一致知覚範囲を注視している画像 (positive) とそれ以外を注視している画像 (negative) として識別器を学習し、視線一致知覚範囲を見ているか否かを判定する識別器を構築した。

HOG は勾配ヒストグラムからなる特徴ベクトルとして表現され、輝度値特徴については画像の輝度値をそのまま特徴ベクトルとして表現した。黒目の中心位置の特徴は、データセットの黒目の中心の位置を人手により与え、左右それぞれの目頭を基準とした位置を表現した特徴とし、左右の目それぞれの x, y 座標で成される 4 次元特徴である。評価には、各被験者ごとに 3, 5, 10 分割交差検定で実験を行なった。

表 2 に視線一致判定の精度を示す。提案手法は比較手法に比べて高い精度を示した。輝度値を特徴とする手法では、目領域画像の切り出しの際のわずかな位置ずれにより精度が低下した。また、黒目の中心位置を特徴とする手法では、目頭の位置を基準にしたため、目頭の位置が精度良く検出できなかった場合、その影響を大きく受け、精度が低下した。それに対して、提案手法では HOG がわずかな位置ずれや回転に対して頑健な特徴として抽出されるため、高い精度を示したと考えられる。また、交差検定に用いる分割数が減ることで学習用画像の数が減ったとしても、ほとんど精度に変化がないことが確認できた。これにより、学習用画像収集に対するユーザの負担を低減することが可能である。



図 9 データセット中の画像の例

表 2 視線一致判定の精度

		適合率	再現率	F 値
10 分割	HOG (提案)	0.612	0.903	0.729
	輝度値	0.429	0.979	0.597
	黒目の中心位置	0.362	0.938	0.522
5 分割	HOG (提案)	0.606	0.899	0.724
	輝度値	0.432	0.986	0.600
	黒目の中心位置	0.364	0.936	0.525
3 分割	HOG (提案)	0.606	0.896	0.723
	輝度値	0.421	0.986	0.590
	黒目の中心位置	0.369	0.961	0.534

表 3 視線一致判定の精度 (境界付近のデータ除外後)

		適合率	再現率	F 値
10 分割	HOG (提案)	0.662	0.964	0.785
	輝度値	0.400	1.000	0.571
	黒目の中心位置	0.310	0.958	0.468
5 分割	HOG (提案)	0.662	0.958	0.783
	輝度値	0.400	1.000	0.571
	黒目の中心位置	0.313	0.959	0.473
3 分割	HOG (提案)	0.651	0.952	0.773
	輝度値	0.388	1.000	0.560
	黒目の中心位置	0.316	0.976	0.478

視線一致知覚範囲の境界付近では、中心位置から離れているため、視線一致を知覚する感度が低下し、視線一致を知覚しているか否か不明瞭であると考えられる。そこで、視線一致知覚範囲の矩形の最近傍点はどちらに判定されても良いとし、評価から除外する。表 3 に境界付近のデータ除外後の結果を示す。表 2 の結果と比較して、提案手法では、比較手法に対してさらに精度が向上していることがわかる。

6. まとめ

一般的にウェブカメラはディスプレイの外部に設置されており、ユーザはディスプレイ上の相手のユーザを見るため、互いに視線が一致しない問題が発生している。

本研究では、単一のカメラのみを用いて視線を一致させる手法を提案した。提案手法では、相手が視線一致を感じる範囲を

視線一致知覚範囲と定義し、ユーザの視線が視線一致知覚範囲内であると判定された場合にのみ、ユーザの目領域を変換し、視線の一致した顔画像を生成し、対話相手に送信する。

視線一致判定に用いる学習データを収集することはユーザにとって負担であると考えられる。しかし、菅野らが提案したようなマウスのクリックイベントを利用したオンライン視線キャリブレーション手法 [12] などを利用することにより、学習データ収集に関する負担を低減することが可能である。

提案手法は、Intel Core i7 3.59GHz CPU, 8GB RAM の利用下で、入力映像が 1,280×960 ピクセルの解像度の場合、処理速度は 5 fps であった。しかし、入力映像の解像度を縮小することや、処理を並列化することで高速化が可能である。

今後の課題としては、新たな特徴量の検討による視線一致判定のさらなる高精度化や、提案手法を実際のテレビ会議システムに実装すること、メガネをかけたユーザに対する対応などが挙げられる。

謝辞 実験に協力して下さった方々に感謝いたします。本研究の一部は、科学研究費補助金による。

文 献

- [1] L. Muhlbach, B. Kellner, A. Prussog, and G. Romahn, "The importance of eye contact in videotelephone service," Proc. 11th Int. Symposium on Human Factors in Telecommunications, no. O-4, pp. 1–8, Sept. 1985.
- [2] S. Uono and J.K. Hietanen, "Eye contact perception in the West and East: A cross-cultural study," Plos one, vol. 10, no. 2, e0118094, Feb. 2015.
- [3] S.M. Anstis, J.W. Mayhew, and T. Morley, "The perception of where a face or television 'portrait' is looking," American J. of Psychology, vol. 82, no. 4, pp. 474–489, Dec. 1969.
- [4] R. Kollarits, C. Woodworth, and J. Ribera, "An eye-contact cameras/display system for videophone applications using a conventional direct-view LCD," Digest of 1995 SID Int. Symposium, pp. 765–768, May 1995.
- [5] R. Yang and Z. Zhang, "Eye gaze correction with stereovision for video-teleconferencing," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 26, no. 6, pp. 956–960, July 2004.
- [6] C. Kuster, T. Popa, J.C. Bazin, C. Gotsman, and M. Gross, "Gaze correction for home video conferencing," Proc. ACM SIGGRAPH Asia, ACM Trans. Graphics, vol. 31, no. 6, pp. 174:1–174:6, Nov. 2012.
- [7] D. Giger, J.C. Bazin, C. Kuster, T. Popa, and M. Gross, "Gaze correction with a single webcam," Proc. 2014 IEEE Int. Conf. on Multimedia and Expo, pp. 68–72, July 2014.
- [8] B. Yip, "Face and eye rectification in video conference using affine transform," Proc. 2005 IEEE Int. Conf. on Image Processing, vol. 3, pp. 513–516, Sept. 2005.
- [9] F. Solina and R. Ravnik, "Fixing missing eye-contact in video conferencing systems," Proc. 33rd Int. Conf. on Information Technology Interfaces, pp. 233–236, June 2011.
- [10] J. Saragih, S. Lucey, and J. Cohn, "Deformable model fitting by regularized landmark mean-shift," Int. J. of Computer Vision, vol. 91, no. 3, pp. 200–215, Jan. 2011.
- [11] N. Dalal and W. Triggs, "Histograms of oriented gradients for human detection," Proc. 2005 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, vol.1, pp. 886–893, June 2005.
- [12] Y. Sugano, Y. Matsushita, Y. Sato, and H. Koike, "Appearance-based gaze estimation with online calibration from mouse operations," IEEE Trans. on Human-Machine Systems, vol. 45, no. 6, pp. 750–760, Feb. 2015.