

〈一般研究課題〉 屋内環境における遠隔対話コミュニケーションの
品質向上に関する研究

助成研究者 中京大学 舟橋 琢磨



屋内環境における遠隔対話コミュニケーションの 品質向上に関する研究

舟橋 琢磨
(中京大学)

A Study on Quality Improvement of Remote Network Communication in Indoor Environment

Takuma Funahashi
(Chukyo University)

Abstract :

In this research, we pay attention to the disagreement of eye-contact in teleconference caused by the separation between the input camera and output monitor devices. Then we are developing the Eye-Contact Camera System for generate by redraw pupil and sclera eye-contacted motion images to the receiver. In this paper, we proposed eyelid motion modeling caused by eyeball movement for naturality of eye-contacted motion. 8 direction eyelid motion models were generated average eyelid feature points by measured some of person, then arbitrary direction eyelid motion model were calculated by linear interpolation. Proposed method was demonstrated more than naturally conventional redrawing method by subjective.

1. はじめに

近年、顔映像などの非言語情報を含んだ遠隔対面コミュニケーション環境が普及しつつある。しかし、対面時と同様の良好なコミュニケーションを実現するためには様々な課題がある。J. L.Zitnickら[1]は、遠隔会議システムの問題点を次のように指摘した。

- (a):音声に遅延があり、音質も良くない
- (b):通信開始に手間が掛かる(電話の利便性に負ける)
- (c):設備の互換性がない

(d):視線の認識や対面会話同様の感覚を得られない

(a)~(c)については、近年の通信の高速・大容量化とPC・カメラデバイスの低価格化、さらにカメラ付き携帯端末でのTV電話機能の登場により、問題解決への可能性が現実的なものとなってきた。しかし、視線の認識や対面会話同様の感覚を得られないことは、未だに解決されていない問題である。中でも顔インタフェースにおける対話者間の「視線の不一致」の解消という課題において、相手の視線は相手の意思表示を確認する上で重要となる。

この問題に対して先行研究[2,3,4]では、ハードウェアベースのシステムにおいて特殊な環境や機器を使用するという制約条件があり、一般家庭への新規導入は困難である。従ってカメラのみを利用したソフトウェアベースのシステム開発が主流であるが、当該分野における研究件数は少数である。モニタの上下部に設置された2台のカメラからステレオ計測を利用して姿勢位置を描画する手法[5]や、予めモニタの前にカメラを置いて100フレーム分の正面顔画像学習データを用いて、入力フレームを正面顔時の角度になるよう修正する手法[6]が提案されている。しかしながら、実時間処理と修正された顔画像の自然性において、いずれかに問題があり実利用化されたという報告はない。

本研究では、特殊な機器や環境を使用せず、安価なカメラのみを入力機器として、画像処理によって視線のずれた不自然な顔画像を視線の一致した顔画像に修正する Eye-Contact Camera システムの開発と実利用化を目的としている。本システムは、画像処理による瞳検出結果をもとに、カメラ・モニタ・瞳の位置関係から視線が一致する瞳中心座標を再計算し、瞳を再描画することによって視線の一致した顔映像を、実時間で生成することが可能である。心理実験によるシステムの品質評価では、25/27名が視線の一致を感じるという評価を得た。ただし、生成された顔映像の自然さについては、18/27名が自然に感じるという評価を得ており、十分な自然性を得たとは言い難い[7]。よって、映像の自然性を向上させることが目的である。

2. Eye-Contact Camera

2.1 アイコンタクト顔画像生成フロー

図1に示す手順でアイコンタクト顔画像を生成する。本手法は、前処理部・アイコンタクト座標変換部・瞳再描画部に大別される。

前処理部では、入力された画像に対して瞳検出を行ない、瞳中心座標(x_{in}, y_{in})を取得する。次に(x_{in}, y_{in})を基に目領域を抽出し、目を強膜部分、瞳部分、肌、目輪郭に領域分割をする。

アイコンタクト座標変換部では、モニタ、カメラ、ユーザとの位置関係から取得した3種のパラメータから視野角を算出し、検出された瞳中心座標(x_{in}, y_{in})をアイコンタクト座標(x_{out}, y_{out})に変換する。

瞳再描画部では、目領域内の全画素(s_n, t_n)に対して、(x_{out}, y_{out})までの距離 d_n を算出し、 d_n 値を閾値判別することで再描画領域を決定する。判別された領域内では、 d_n 値を用いて描画色を変化させ、強膜、瞳を再描画する。

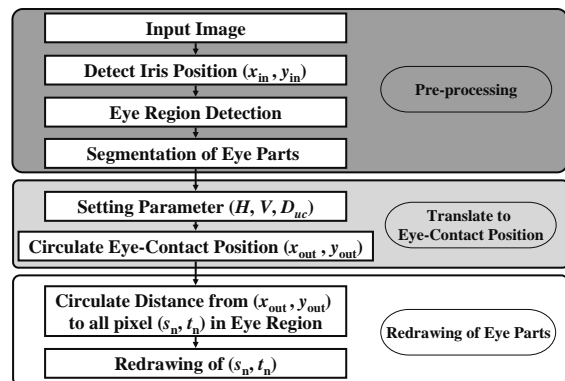


図1 アイコンタクト顔画像生成フロー

2.2 アイコンタクト座標変換

図2に示すようなモニタ,カメラ,ユーザとの重心位置関係を D_{uc} (ユーザ~カメラ間距離), V (モニタ~カメラ間垂直距離), H (モニタ重心~カメラ間水平距離)の3種のパラメータとして指定し, 手動で各距離を計測した後, 垂直視野角度 θ_{ver} と水平視野角度 θ_{hor} を式(4)にて算出する。次に, 元の瞳中心座標 (x_{in}, y_{in}) に対して θ_{ver} , θ_{hor} を顔画像上でのpixel値に変換する関数 Δx , Δy に代入し, 移動後の瞳中心座標 (x_{out}, y_{out}) を式(1)にて算出する。

また, 式(2)の (x_{max}, y_{max}) および (x_{min}, y_{min}) はモニタを目視した場合における, 検出された瞳中心座標の最大値と最小値である。

$$\theta_{ver} = \tan^{-1} \frac{V}{D_{uc}} \quad \theta_{hor} = \tan^{-1} \frac{H}{D_{uc}} \quad (1)$$

$$\begin{aligned} x_{out} &= x_{in} + \Delta x(\theta_{hor}) : \Delta x(\theta_{hor}) = \frac{\theta_{hor}}{x_{max} - x_{min}} \\ y_{out} &= y_{in} + \Delta y(\theta_{ver}) : \Delta y(\theta_{ver}) = \frac{\theta_{ver}}{y_{max} - y_{min}} \end{aligned} \quad (2)$$

$((x, y)_{max} = \text{Detect maximum position of iris})$
 $((x, y)_{min} = \text{Detect minimum position of iris})$

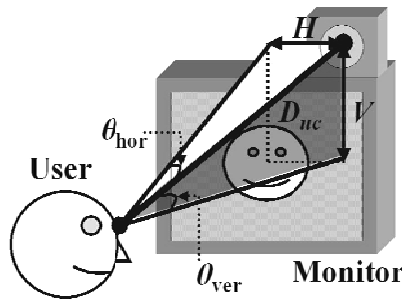


図2 視野角度計算におけるパラメータの位置関係

2.3 瞳再描画

瞳再描画の前処理として, 目を強膜部分,瞳部分,肌,目輪郭に領域分割をする。入力画像から検出された瞳中心座標 (x_{in}, y_{in}) を基に, 目領域を抽出する。その領域に対して, 強膜(白目), 瞳(黒目), 肌の各々の個別領域を濃度差の閾値判定によって図3に示すように領域分割する。個別領域を判定する方法は, 以下に示すような手順で行う。

- (a): 閾値以下の領域に $f_{ij} = 1$ としてラベリングする。
- (b): 目領域の始点から垂直走査して, 上下の濃度差が閾値未満となる領域では $f_{ij} = 3$, 閾値以上となる領域では $f_{ij} = 2$ としてラベリングする。
- (c): 同様に終点から垂直走査して, (b)と同様に行う。
- (d): 残存する部分を $f_{ij} = 0$ としてラベリングする。

この処理を左目領域, 右目領域に対して行い, $f_{ij} = 0$ の領域が強膜, $f_{ij} = 1$ の領域が瞳, $f_{ij} = 2$ の領域が目輪郭, $f_{ij} = 3$ の領域が肌であるとする。

目領域の始点から終点までを水平方向に走査させ, 任意の画素 (s_n, t_n) から瞳中心座標 (x_{out}, y_{out}) までの距離 d_n を求める。 d_n が瞳の半径 r 以内且つ, $f_{ij} = 0$ または $f_{ij} = 1$ ならば, 瞳部分であるとして, 黒色を用いて描画し, 距離 d_n が半径 r 以上且つ $f_{ij} = 1$ であるならば, 強膜部分であるとして, 白色を用いて描画す

るという条件により、式(3)を用いて瞳を再描画する。図4に上記のパラメータ関係を示す。瞳は虹彩と瞳孔によって構成されているため、同色で描画するのは不自然である。また、目は涙で覆われていることから、環境光によって常に反射光が入る。より現実に近い目を再現するために、式(4)にて同心円状にグラデーションをつけ、平滑化を行うことで、瞳再描画を行った際の不自然さを抑制する。

$$\begin{cases} d_n = \sqrt{(s_n - x_{out})^2 + (t_n - y_{out})^2} \\ d_n < r \Rightarrow B_{i,j} \\ d_n \geq r \cap f_{i,j} = 1 \Rightarrow W_{i,j} \\ d_n \geq r \cap f_{i,j} = 0 \Rightarrow F_{i,j} : \text{pixel value} \\ f_{i,j} : \text{label number in Fig.3} \end{cases} \quad (3)$$

$$\begin{aligned} B_{i,j} &= \text{black} + d_n \times \alpha \\ W_{i,j} &= \text{white} + d_n \times \alpha \\ \alpha &: \text{constant} \\ \left(\begin{aligned} \text{black} &= \min \{F_{i,j} | f_{i,j} = 1\} \\ \text{white} &= \max \{F_{i,j} | f_{i,j} = 0\} \end{aligned} \right) \end{aligned} \quad (4)$$

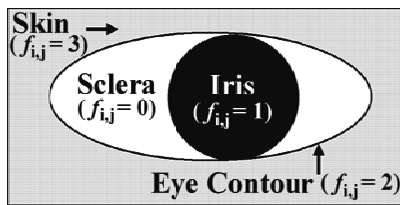


図3 瞳再描画における目領域分割

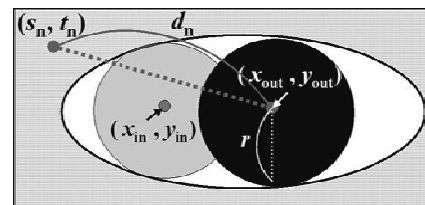


図4 瞳再描画におけるパラメータ関係

3. 瞼輪郭のモデル生成

3.1 眼球運動に伴う瞼動作の計測実験

瞼輪郭モデルを生成するために、被験者6名の瞼動作を記録して、特徴抽出を試みる計測実験を行った(図5)。視野角の関係により、カメラと被験者の距離を50cmとし、カメラを設置し、被験者の頭部を固定する。正面、視野角10°、20°、30°時の各8方向の注目点(計25点)を、眼球のみ動かして注視した時の顔画像を撮像する。注目点25点には効率を考え、注目点番号を付与した(図6)。今回は、後述するモデル生成のデータには使用しなかったが、動画像による撮像を行った。以下、2パターンの眼球運動に伴う瞼動作の撮像を行った。

・[円形状パターン]1→2→3→4→5→6→7→8→9→10→11→12→13→14→15→16→17→18→19→20→21→22→23→24→25(終了)

・[放射線状パターン]1→2→1→3→1→4→1→5→1→6→1→7→1→8→1→9→1→10→1→11→1→12→1→13→1→14→1→15→1→16→1→17→1→18→1→19→1→20→1→21→1→22→1→23→1→24→1→25(終了)

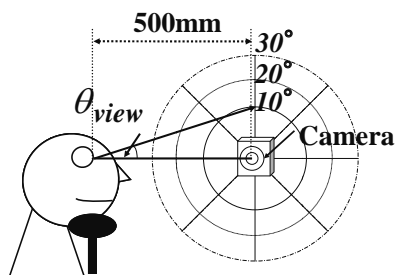


図5 被験者と呈示板の関係

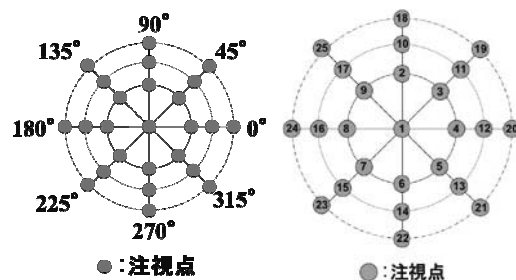


図6 呈示刺激

3.2 瞼輪郭および瞳の特徴点定義

撮像した動画像をフレームレートに従って静止画像に分割し、目視判断によって、注目点番号に対応する静止画像を選出する。選出した各顔画像について、顔特徴点手動入力モジュールを使用し、瞼および瞳特徴点を手動で入力する。このモジュールは、本研究グループで研究・開発を行っている似顔絵生成システムの一部であり、顔部品の特徴点を入力できる機能を持っている。また、入力された特徴点を正規化することによって、データ間の演算処理を容易にすることができる。それぞれ入力される特徴点は、上瞼特徴点(5点)、下瞼特徴点(5点)、瞳特徴点(9点)であり、瞼輪郭および瞳の始点と終点は重複する(図7)。



図7 瞼輪郭および瞳輪郭の特徴点入力

3.3 瞼輪郭および瞳輪郭の平均モデルデータ生成

注視点25点における被験者6名の顔画像150枚(25枚×6名)に対して、特徴点の平均を算出し、瞼および瞳を線で描画する。各モデルは、瞳の大きさを基準として正規化されており、瞳特徴点の重心を算出して位置合わせを行っている。また、描画線については、特徴点入力時に顔画像データから輪郭線を直接トレースしてデータを取得している。瞼モデルを図8に示す。

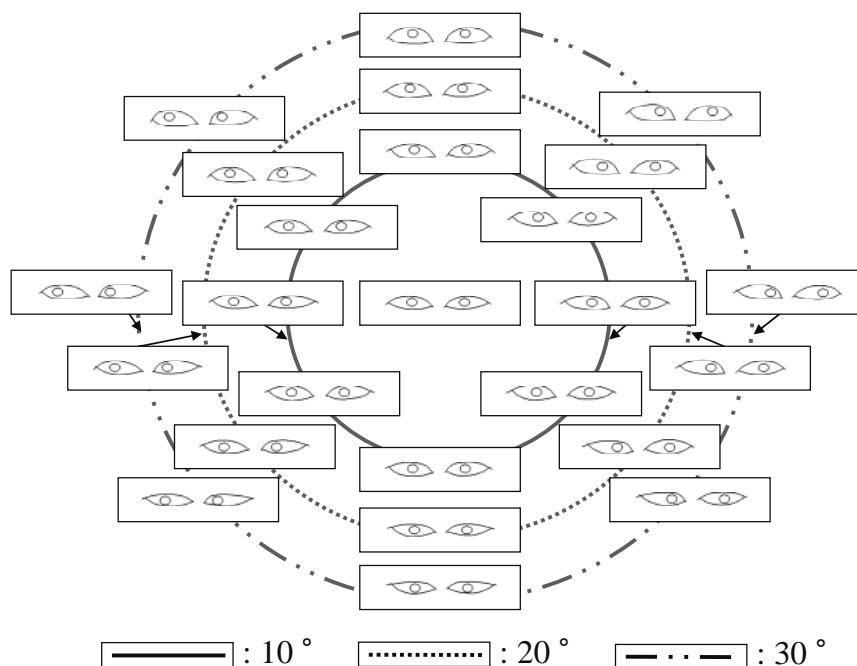


図8 輪郭モデルデータ生成

3.4 任意角度視点の瞼輪郭および眼球輪郭近似モデル生成

前節で取得した注目点25箇所の瞼／瞳輪郭モデルを利用して、方向角 $0^{\circ} \sim 360^{\circ}$ 、視野角 $0^{\circ} \sim 30^{\circ}$ の任意角度の瞼／瞳輪郭近似モデル生成手法を提案する。

瞼／瞳輪郭モデルデータは、図9に示す方向角 $x^{\circ} \sim x+45^{\circ}$ ($0 \leq x \leq 315$)、視野角 10° 、 20° 、 30° で構成されている。作成すべき方向角と視野角を決定すると対応する最寄りのモデルデータを最大4

つ参照し， 験輪郭モデルを適切に推定する。式(5),(6)は， 任意方向角 x と視野角 y を算出するための内分点を求めており， 参照される各既知モデルから線形補間によって任意角度の験輪郭モデルが作成される。作成した任意角度の験輪郭モデル例を図10に示す。

$$\begin{cases} Rx_n = \frac{1}{\theta_a} (\theta_a - |x - \theta_{\alpha_n}|) \\ Ry_n = \frac{1}{\theta_b} (\theta_b - |y - \theta_{\beta_n}|) \end{cases} \quad (5)$$

$(n=1,2,3,4)$

$$D_k = \frac{Rx_k + Ry_k}{\sum_{k=1}^n Rx_k + Ry_k} \quad (6)$$

$(n=1,2,3,4)$

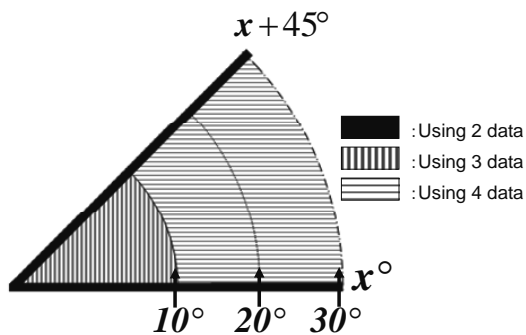


図9 任意視点のモデル推定に使用する参照データ数

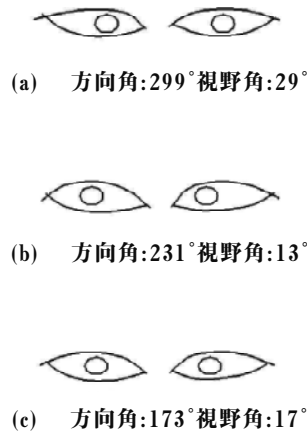


図10 任意視点験／瞳輪郭モデル推定結果例

4. 瞳・験輪郭画像生成実験

4.1 験輪郭および眼球輪郭近似モデル生成結果

推定された座標データを用いて， ペイントツールで手動にて験および瞳を再描画した。撮影された画像と瞳のみの再描画した画像と験・瞳再描画した画像の3種を比較して， 験・瞳が再描画された画像の再現性を試験的に検証した。

本実験では， 最も験輪郭の形状変化が小さいと考えられる視野角10度の注目点(8点)において， 全被験者6名の験および瞳の再描画を行った。再描画結果の一例を図11に示す。各被験者の正面顔左眼画像は左側に示し， 右上段は験・瞳再描画， 右下段は瞳のみの再描画結果である。

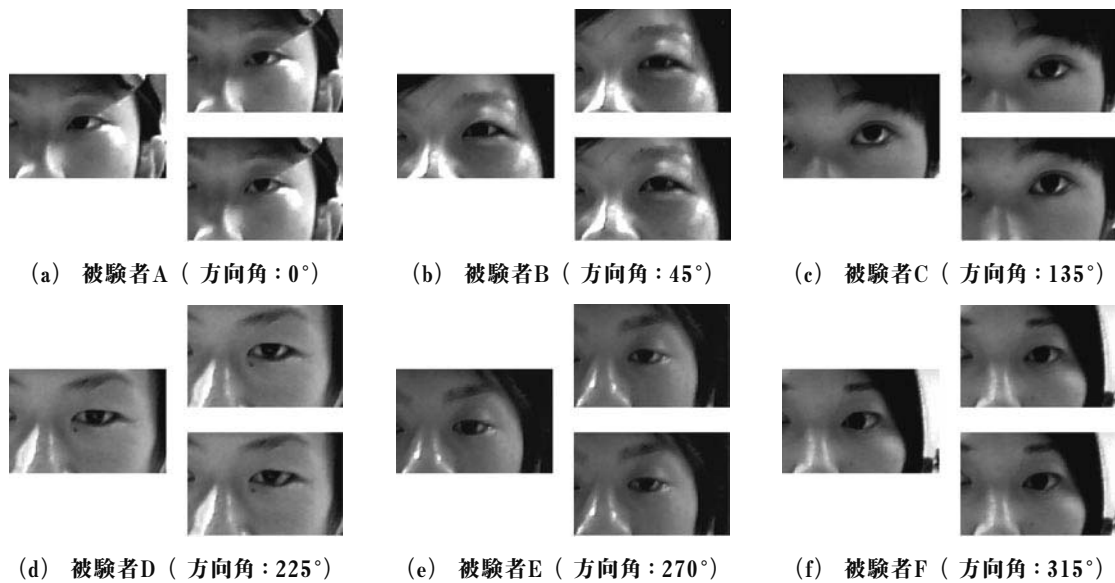


図11 輪郭モデルを用いた瞳・瞼再描画結果例

4.2 考察

方向角 $0^{\circ} \sim 180^{\circ}$ の場合は、瞳のみ再描画処理と瞼・瞳再描画処理を施したアイコンタクト顔画像に不自然さは感じられなかった。要因として上眼瞼下端と下眼瞼上端の変化が共に少ないことが挙げられる。したがって、瞼輪郭が再描画されていない場合でも十分アイコンタクト性が再現できると考えられる。

しかしながら、方向角 $225^{\circ} \sim 315^{\circ}$ の場合は、瞳のみ再描画処理を施すと不自然さを感じる結果を得た。瞼・瞳再描画処理では、二重瞼の場合において不自然さが感じられたが、瞼に対する違和感だけであり、アイコンタクト性は失われていない結果を得た。要因として上眼瞼下端の変化が多いことが挙げられる。より自然なアイコンタクト顔画像を生成するには、二重瞼の再描画処理を行う必要があるが、瞼の判別を行わなければならないと考えられる。

5. 一重／二重瞼判別の検討

人の瞼は一重・二重に大別することができるが、それによって顔の印象は大きく変わると言える。我々は、実世界で人の瞼を目視した場合、見た目ですぐにどちらか判断することは容易にできるが、デジタル画像中において明確な境界情報は失われてしまい、瞼がどちらであるかを判別することは困難である。そこで、平均顔から作成した目領域画像をテンプレートとして使用し、入力した目領域画像の一重・二重を判別する手法を検討した。

5.1 類似度および相違度を用いた判別方法

平均顔は瞼輪郭の形状変化に対応するために、性別、一重・二重瞼を大別して、さらに瞼頂点の位置を図12に示すように目頭寄り、中央、目尻寄りの3種類に大別して作成した。これらの平均顔に対して、瞳を中心に 48×36 ピクセルの矩形領域を切り出し、テンプレート画像として計6枚を作成する。入力画像1枚に対して、一重・二重各3種類のテンプレートを用いてテンプレート・マッチングを行い、類似度 R (式(7))ならびに相違度 D (式(8))を算出する。一重・二重のテンプレート群において、それぞれ極大の指標値を比較して、最大の指標値をもつテンプレートを判別結果とした。

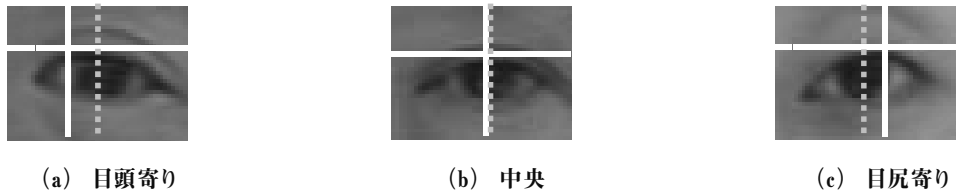


図12 眼頂点位置による顔画像の分類

表1 テンプレート画像一覧

	一重			二重		
	目頭寄り	中央	目尻寄り	目頭寄り	中央	目尻寄り
男性						
女性						

$$R_{NCC} = \frac{\sum_{j=0}^{N-1} \sum_{i=0}^{M-1} (I(i, j) - \bar{I})(T(i, j) - \bar{T})}{\sqrt{\sum_{j=0}^{N-1} \sum_{i=0}^{M-1} (I(i, j) - \bar{I})^2 \times \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} (T(i, j) - \bar{T})^2}} \quad (7)$$

$$D_{SAD} = \frac{1}{255 \times M \times N} \sum_{j=0}^{M-1} \sum_{i=0}^{N-1} |I(i, j) - T(i, j)| \quad (8)$$

5.2 判別実験内容および結果

平均顔作成時と同様に収集した顔画像を男女に分類し、さらに一重・二重を目視にて分類する。また、顔画像から目領域を手動で設定した。実験対象の顔画像は男性の一重で左右の目を合わせて30枚使用し、二重も同じように30枚使用した。女性も同様に枚数を合わせ、合計で120枚の目領域画像を対象とした。男女、一重・二重それぞれのカテゴリにおいて指標別に成功枚数と成功率を表2に示す。最も成功率が高いカテゴリは一重男性、低いカテゴリは二重女性であった。

表2 指標別における成功枚数と成功率

判別指標	男性・一重	男性・二重	女性・一重	女性・二重	合計
相違度D	20枚(67%)	19枚(63%)	25枚(83%)	13枚(43%)	77枚(64%)
類似度R	27枚(90%)	21枚(70%)	24枚(80%)	20枚(67%)	92枚(77%)

5.3 考察

相違度Dと類似度Rによる類似度の結果を比較すると、カテゴリ毎のばらつきはあるが、合計の成功率を比べると類似度の成功率は約77%、相違度の成功率は64%となり、約13%の差があった。正規化相関を用いた類似度は画像の輝度値が正規化されているため、単純に差を求める相違度よりも成功率が高くなるという結果になった。決して良い結果ではないが、その原因として一番大きい

のが瞼位置のずれである。類似度を用いた場合は、ある程度の輝度差を無視できるので、重要なのは瞼の位置合わせであると言える。これらのことから、提案手法である平均目領域画像を用いた一重・二重判別は瞼位置を揃えることで精度が向上すると考えられる。

6. まとめ

本研究では、より自然なアイコンタクト顔映像を作成すべく、瞳再描画技術と連動させて、瞼の動きも再描画させる機能をEye-Contact Cameraシステムに組み込むことを目的とし、その品質向上のための検証実験を行った。

具体的には被験者6名に一定の眼球運動をカメラで撮像し、その映像より25点の注目点の静止画像を作成した後、注視点の瞼動作を詳細に記録した。この記録より平均データを作成し、眼球運動に伴う瞼動作モデルとして導入した。そして瞼の動きモデルを利用し、視野内の任意の方向の眼球運動に対処するため、瞼動作の補間技術を提案した。また、瞳のみおよび瞳と瞼の再描画の画像例をそれぞれ実験的に作成することで、瞼を再描画することの有効性を確認することができた。

また、瞼輪郭描画印象の向上のため、平均目領域画像を用いて入力目領域画像の一重・二重を判別する手法を提案した。

男女、一重・二重それぞれに分類して平均顔を作成し、さらに目の形状による分類を目視にて行い、テンプレート・マッチングによる判別実験を行った。正規化相関による類似度と絶対差の総和による相違度を用いてそれぞれ実験を行い、判別結果を比較すると、類似度を用いた場合の成功率は約72%、相違度を用いた場合は約64%となった。

今後の改善点としては、瞼の動きモデル生成および精度向上である。今回、任意方向の瞼形状の生成を行ったが、Eye-Contact Cameraシステムとの関係を考慮し、頭部の回転方向における判定モジュールも必要である。瞼判別においても最も高い成功率は77%であり、精度が高いとは言えないので改善が必要である。瞼の位置がずれると成功率の低下につながるため、瞼の頂点位置を探索するための特徴量が必要である。

文献

- [1] J.L.Zitnick, J.Gemmell and K.Toyama: Manipulation of Video Eye Gaze and Head Orientation for Video Teleconferencing, Microsoft Technical Report, MSR-TR-99-46 (1999)
- [2] H.Ishii, M.Kobayashi and J.Grudin: "Integration of interpersonal space and shared workspace: ClearBoard design and experiments", ACM Transactions on Information Systems, Volume 11, Issue 4, pp.349-375 (1993)
- [3] 小木哲朗, 山田俊郎, 栗田裕二, 服部陽一, 廣瀬通孝: "仮想空間共有のためのビデオアバタ技術とその利用法", 日本バーチャルリアリティ学会論文誌, Vol.8, No.1, pp.37-46 (2003)
- [4] 石井裕, 渡辺 富夫: "VirtualActorを対面合成した身体的ビデオコミュニケーションシステム", ヒューマンインタフェース学会論文誌, Vol.5, No.2, pp.73-82 (2003)
- [5] R.Yang and Z.Zhang: "Eye Gaze Correction with Stereovision for Video-Teleconferencing ", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.26, No.7, pp.956-960 (2004)
- [6] B.Yip: "Face and Eye Rectification in Video Conference Using Artificial Neural Network ",

Proc. of IEEE International Conference on Multimedia & Expo 2005, CD-ROM (2005)

- [7] 舟橋琢磨, 藤原孝幸, 輿水大和: “遠隔地非対面会話における顔インタフェースメディア生成”, 精密工学会誌vol.75, no.2, pp.321-325 (2009)