

# 3次元ビデオからの人物頭部及び視線の検出

藤本 圭<sup>†</sup> 延原 章平<sup>†</sup> 松山 隆司<sup>†</sup>

概要：本論文では3次元ビデオから人物頭部および視線方向の検出を行う。提案手法のポイントは、(1) 3次元形状と多視点画像を用いた人物頭部の検出アルゴリズムと、(2) 3次元ビデオから仮想的に撮影した正面顔画像を用いた視線方向の推定である。提案手法では対象人物固有の形状モデルを使うことなく、既存の2次元顔検出器を組み合わせることで顔の位置および姿勢の推定を実現した。

## Detection of human face and eye direction from 3D video

KEI FUJIMOTO,<sup>†</sup> SHOHEI NOBUHARA<sup>†</sup> and TAKASHI MATSUYAMA<sup>†</sup>

**Abstract:** This paper aimed to estimate the three dimensional position of a human face and eye direction from 3D video. The key points of the proposed techniques are (1) a detection algorithm for the human head position using three dimensional shape data and images captured from multiple viewpoints and (2) obtaining a frontal face image in order to estimate the eye direction. The proposed methods allow the estimation of the human face position and its posture without any three dimensional models.

### 1. はじめに

#### 1.1 研究の背景と目的

本論文では、自由な状態で運動している対象人物の頭部および視線を非接触・非拘束に検出する新たな手法を提案する。

日本の伝統舞踊では、演者は重要な心情変化を流し目など細かな身体動作により表現することが多く、そのような細やかな身体動作を本来の動作を妨げることなく推定することができれば、演者の意図といった形の無い重要な情報を3次元ビデオと共に保存することができるようになり、無形文化財の保存に繋がると考える。無形文化財の保存のみならず、選手の目線に立った臨場感溢れるスポーツ観戦、視線による計算機の操作など様々な応用が期待できる。

本研究のアイデアは、対象の周囲に配置した多視点カメラ群から対象の3次元形状を推定<sup>1)</sup>した上で、(1) 3次元形状中の人物頭部領域を投票によって推定するとともに、(2) 人物頭部領域を仮想的に正面から撮影することによって虹彩検出に適した仮想正面顔画像を作成する点である。

#### 1.2 関連研究と本研究の位置付け

人物の顔を検出する研究としては、これまで2次元画像から顔を検出する研究が数多くなされ、デジタルカメラのオートフォーカス機能などにこれらの研究

が実用化されている。その方法には、肌色情報を用いた検出方法、顔のモデル画像を用いた方法、学習により顔検出のための識別器を構成する Example-based 顔検出法<sup>2)</sup> などがある。これらはしかし画像における顔の位置、つまり平面中の顔の位置を検出しているに止まっている。3次元空間における顔の位置を推定する方法として、人物頭部の3次元モデルを作成し人物の3次元形状とマッチングを行う手法が考えられるが、人物の頭部形状は固有差がありそのモデルを作るのは煩雑となる。本論文では、人物頭部の3次元モデルを用いずに3次元空間における顔の位置を検出する手法を提案する。

人物の視線を検出する研究としては、顔検出と同じく画像中における虹彩の位置を検出する研究<sup>4)</sup>、3次元のものには赤外線カメラを用いて眼球運動を撮影することで視線を検出する研究<sup>5)</sup>、視線ではないが因子分解法と部分空間法を用いて顔方向を検出する方法<sup>3)</sup> などがある。本論文では特殊なカメラを用いず、非接触・非拘束に3次元ビデオから視線を検出する手法を提案する。

### 2. 提案手法

本論文では入力として、人物の3次元形状、それを撮影した各カメラ画像、及びキャリブレーションデータは与えられているものとする。キャリブレーションデータとは、各カメラについてカメラの内部パラメータを表す  $3 \times 3$  の射影行列  $A$ 、世界座標系からカメラ

<sup>†</sup> 京都大学大学院情報学研究科  
Graduate School of Informatics, Kyoto University

座標系へ変換するための回転行列  $R$  と並進ベクトル  $T$  である．また，各視点のカメラ画像はいずれもレンズ歪みのない理想画像とし，人物の3次元形状は三角形メッシュモデルとして与えられているものとする．

## 2.1 人物頭部の検出

各撮影画像から顔を探し，検出された顔を撮影したカメラと人物の位置関係を評価し，人物の3次元形状に投票することで，顔の領域を決定する．画像から人物の顔を検出する代表的な手法には以下のものがある．

- 学習データを用いる Example-based 顔検出法
- 2次元の顔モデルを用いた検出方法

後者は精度良く検出を行うには顔領域を大きな解像度で撮影した画像が必要である．これに対して，前者は後者に比べて低解像画像でも検出できる．本論文の目的から，舞踊など人物の全身を保存したいので，顔領域を大きく撮影する事は難しい．よって，前者に属する Viola-and-Jones の顔検出アルゴリズム<sup>2)</sup>を採用する．提案手法の詳細を以下に示す．

**step1** 視点  $C_i$  ( $i = 1, \dots, N$ ) における撮影画像を  $I_i$  ( $i = 1, \dots, N$ ) とする． $I_i$  それぞれにおいて，Viola-and-Jones の顔検出器によって顔領域を検出する．顔領域は画像  $I_i$  中の矩形領域の集合  $R_i = \{r_{ij} | r_{ij} \in I_i\}$  であり，空集合もありえる．

**step2** 人物の3次元形状の三角形パッチ頂点集合を  $V$  とする．ここで各頂点  $v \in V$  について， $v$  を顔領域の一部として検出した視点数  $c_v$  を次のようにして求める．

- (1)  $c_v = 0$  とする
- (2) 各視点  $C_i$  において， $v$  を画像  $I_i$  上に投影した点  $v_i$  が  $R_i$  中に含まれていれば  $c_v = c_v + 1$  とする

**step3**  $c_v > 0$  なる  $v$  の集合  $V_c$  を，メッシュの接続関係に基づいて部分集合  $S = \{s_i | s_1 \cup s_2 \cup \dots \cup s_n = V_c \in V\}$  へと分割する．ここで  $n$  は部分集合の数であり， $s_i$  は互いに疎であるとする．

**step4** 集合  $S$  に含まれる各頂点集合  $s_i$  について， $c_v$  の平均値  $\bar{c}_i$  を求める．

$$\bar{c}_i = \frac{\sum_{v \in s_i} c_v}{N(s_i)} \quad (1)$$

ここで  $N(s_i)$  は  $s_i$  に含まれる頂点の数である．

**step5** 全  $s_i$  の中で最大の  $\bar{c}_i$  をもつ部分集合を顔領域  $F$  とする．

$$F = \operatorname{argmax}_{\bar{c}_i} s_i \quad (2)$$

## 2.2 虹彩の検出

撮影画像から直接視線を検出することができればいいが，自由な状態で運動している対象人物の視線を非接触・非拘束に推定したいので，カメラは必ずしも常には正面から対象を撮影できない．従って，正面顔を常に撮影する仮想カメラを設け，正面顔を投影し，虹彩を検出することで実現する．その提案する虹彩検出

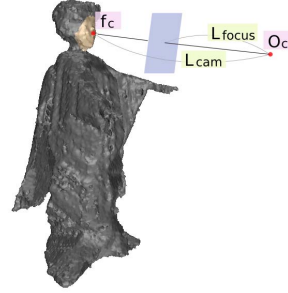


図1 投影するスクリーン

の流れを示す．なお，2.1 で述べた人物頭部の検出により，人物の3次元形状中における顔領域は既知であるものとする．虹彩検出の流れは次の通りである．

- (1) 人物の顔方向ベクトルを求める
  - (2) 仮想カメラとスクリーンを作り，人物の正面顔を投影する
  - (3) 投影した正面顔の回転を補正する
  - (4) 得られた正面顔画像から虹彩を検出する
- 以下，その詳細を示す．

### 2.2.1 顔方向ベクトルの検出

**step1** 人物の3次元形状を表したメッシュのうち，顔領域に含まれる三角形パッチの頂点集合を  $F$  とする．各頂点  $f \in F$  について，その単位法線ベクトル  $\vec{n}_f$  を求める．

**step2** 全ての  $f$  について， $\vec{n}_f$  の平均を取る．

$$\vec{n}_{\text{face}} = L_{\text{cam}} \frac{\sum_{f \in F} \vec{n}_f}{N(F)} \quad (3)$$

ここで， $L_{\text{cam}}$  は顔方向ベクトルの長さ（顔の中心から投影中心までの距離となる）， $N(F)$  は  $F$  に含まれる頂点数とする． $\vec{n}_{\text{face}}$  を顔方向ベクトルとして採用する． $\vec{n}_{\text{face}}$  の始点は，step3 で求める顔の中心位置  $f_c$  とする．

**step3** 集合  $F$  に含まれる各頂点  $f$  について，式(4)のように， $f$  の位置の平均を取ることで， $f_c$  を求める．ここで， $\vec{f}$  を  $f$  の位置ベクトル， $\vec{f}_c$  を  $f_c$  の位置ベクトルとする．

$$\vec{f}_c = \frac{\sum_{f \in F} \vec{f}}{N(F)} \quad (4)$$

### 2.2.2 正面顔の投影

正面顔画像をいかに高解像で目の特徴を捉え易い画像として得るかが，後の虹彩検出において非常に重要となる．正面顔画像を取得する手順を以下に示す．

**step1** 顔方向ベクトル  $\vec{n}_{\text{face}}$  の終点を投影中心  $O_c$  とする仮想カメラを作り，投影中心  $O_c$  から顔の中心  $f_c$  に向かって焦点距離  $L_{\text{focus}}$  の位置に正面顔を投影するスクリーンを作る（図1参照）．このスクリーンサイズや焦点距離  $L_{\text{focus}}$  は，後の虹彩検出に都合の良いものにしておく（実験では，スクリーンサイ

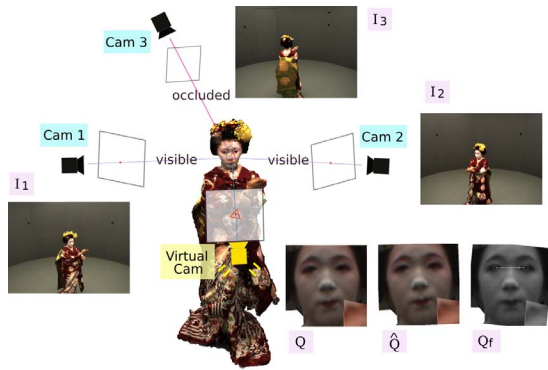


図 2 正面顔投影の詳細

ズと  $L_{focus}$  は手で与えた)。

**step2**  $f$  を観測可能なカメラの撮影画像を用いて、作成したスクリーンに正面顔を投影する (図 2)。こうして得られた正面顔画像を  $Q$  とする。

### 2.2.3 正面顔画像の回転を補正

人物が首を傾げている場合、このままでは回転した正面顔画像が得られてしまう。後の虹彩検出における検出精度を向上させるため、この段階でその回転を補正しておく。2 次元画像からの顔検出の手法である Viola-and-Jones の顔検出器は、ある一定の回転角 (学習用データにより決まる) 内にある顔のみを検出する特性を持っている。この特性を利用し、投影により得られる正面顔について、Viola-and-Jones の顔検出器で顔として検出される回転範囲を求め、得られた範囲から顔の回転を補正する。その詳細手順を示す。

**step1** 投影して得られた正面顔画像を  $1^\circ$  から  $360^\circ$  まで回転させた画像を  $Q_i$  ( $i = 1, \dots, 360$ ) とする。各  $Q_i$  に対して Viola-and-Jones の顔検出器を用いて顔検出を行う。検出された顔の個数が 1 の場合、 $Q_i$  の投票数を  $N(i) = 1$  とし、それ以外を  $N(i) = 0$  とする。

**step2**  $N(i) = 1$  なる  $i$  の集合を  $D$  とする。 $D$  を  $i$  の連続性により部分集合  $d = \{d_j | d_1 \cup d_2 \cup \dots = D\}$  へと分割する。隣り合う部分集合間の ( $i$  についての) 距離が 1 であれば、その部分集合同士を連結する。つまり、式 (5) を同時に満たす  $i$  について、 $N(i) = 1$  とする。

$$N(i-1) = 1, N(i) = 0, N(i+1) = 1 \quad (5)$$

**step3**  $D$  及び  $d$  を更新する。 $d$  のうち最も大きな (最も  $i$  が連続した) 部分集合  $d_{max}$  について、 $\theta_b$  を  $d_{max}$  の  $i$  の最小値、 $\theta_e$  を  $d_{max}$  の  $i$  の最大値として検出された顔の回転範囲 ( $\theta_b \leq \theta \leq \theta_e$ ) とする。また、その回転範囲の中心を  $\theta_m$  とする。投影して得られた正面顔画像を  $-\theta_m$  だけアフィン変換により回転させる。こうして得られた回転補正した正面顔画像を  $\hat{Q}$  とする (図 2)。

### 2.2.4 虹彩の検出

得られた正面顔画像から虹彩を検出する。投影して得られる正面顔画像は撮影画像に比べ幾分ぼやけや歪みが生じる。従って、より頑健な虹彩検出の手法を選択する必要がある。分離度フィルタとテンプレートを用いた手法<sup>4)</sup>を用いて虹彩の位置の検出を行う。

### 2.3 視線の検出

#### 2.3.1 正面顔画像の水平化

正面顔を投影する際に顔の傾きを一度補正した。しかし、それだけでは完全に傾きを正すことは難しい。後の視線方向の偏り具合を検出する際、画像に対して水平に左右の虹彩中心があると効率が良い、よって、虹彩の検出により得られた左右の虹彩の位置を利用して顔の傾きをさらに補正する。以下その詳細を示す。なお、以降の記述で、左目 (左の虹彩) とは検出対象人物から見て左側にある目 (虹彩) を指すものとする、右についても同様とする。

**step1** 正面顔画像  $\hat{Q}$  において検出された左右の虹彩中心をそれぞれ点  $E_l$  及び点  $E_r$  とし、それらの中点を点  $M$  とする。点  $E_l$  と点  $E_r$  を結ぶ線分と画像の水平方向とが成す角  $\phi$  を求める。

**step2** 回転中心を  $M$ 、回転角度を  $\phi$  として、アフィン変換により傾きが正される方向に  $\hat{Q}$  を回転させる。得られた画像を  $Q_f$  とする (図 2)。

#### 2.3.2 虹彩の偏り具合の検出

左右の虹彩がどの程度、左・右に偏っているかを検出する。そのために左右の目頭の中心点を検出し、その中心点から左右の虹彩までの距離の差分を計算することにより行う。その詳細を以下に示す。

**step1** 虹彩の位置を元に水平化した正面顔画像  $Q_f$  について、検出された左右の虹彩中心をそれぞれ点  $E_l (X_l, Y_l)$ 、点  $E_r (X_r, Y_r)$  とし、それらの中点を点  $M (X_m, Y_m)$  とする。次の式 (6) の範囲にある点  $(x, y)$  の平均輝度値  $\overline{lum}$  を求める。

$$\begin{aligned} X_m - w \leq x \leq X_m + w \\ Y_m - h \leq y \leq Y_m + h \end{aligned} \quad (6)$$

**step2** 点  $M$  から  $x$  の負の方向 (右目方向) へ向かって、点  $(X_a, Y_m)$  を移動させ、次の式 (7) の範囲にある点  $(x, y)$  の平均輝度値が初めて  $\overline{lum}$  の  $\beta$  倍を下回る点  $h_r (X_a, Y_m)$  を右の目頭として採用する (図 3 参照、実験では  $\beta$  の値は手で与えた)。

$$\begin{aligned} X_a - W \leq x \leq X_a \\ Y_m - H \leq y \leq Y_m + H \end{aligned} \quad (7)$$

左の目頭についても同様に、点  $M$  から  $x$  の正の方向 (左目方向) へ向かって、点  $(X_a, Y_m)$  を移動させ、次の式 (8) の範囲にある点  $(x, y)$  の平均輝度値が初めて  $\overline{lum}$  の  $\beta$  倍を下回る点  $h_l (X_a, Y_m)$  を左の目頭として採用する。

$$\begin{aligned} X_a \leq x \leq X_a + W \\ Y_m - H \leq y \leq Y_m + H \end{aligned} \quad (8)$$

**step3** 得られた左右の目頭  $h_r, h_l$  の中点を  $C$  とする。

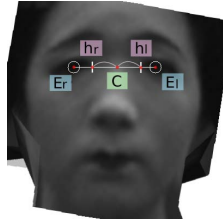


図3 眉間の中心点の検出

左右の虹彩中心から点  $C$  までの距離を  $L_l, L_r$  とする． $\alpha$  を式 (9) のように定める． $\alpha$  が正のとき人物は左を見ており，負のときは右を見ていると言える，また， $\alpha$  の大きさによりその度合いを知ることができる．

$$\alpha = L_l - L_r \quad (9)$$

### 2.3.3 視線方向の検出

虹彩の偏り具合を  $\alpha$  として得た．この  $\alpha$  と角度を対応付けることで視線方向を特定する．

**step1** 視線方向を表す角度を  $\theta_e (-90^\circ < \theta_e < 90^\circ)$  とする． $\theta_e$  が負の値であれば右を，正の値であれば左を見ているものとする． $\theta_e$  と  $\alpha$  を次のように対応付ける．

$$\theta_e = \arcsin\left(\frac{\alpha}{A}\right) \quad (10)$$

$A$ :  $\alpha$  の最大値 (目頭と目尻の距離)

**step2** 顔方向ベクトル  $\vec{n}_{\text{face}}$  の始点を左右の目頭の中心  $C$  としたベクトルを， $\vec{n}_{\text{eye}}$  とする． $C$  を中心に  $\vec{n}_{\text{eye}}$  を  $z$  軸の方向に  $\psi$  だけ回転する (顔方向ベクトル  $\vec{n}_{\text{face}}$  が下向き加減になるため，それを補正する)． $C$  を中心に  $\vec{n}_{\text{eye}}$  を  $\theta_e$  だけ  $x y$  平面に対して平行に回転させ，2.2.3 及び 2.3.1 において正面顔を回転した角度分，そのベクトルを  $C$  を中心に  $\vec{n}_{\text{eye}}$  の回りに回転させたベクトルを視線ベクトル  $\vec{E}_{\text{eye}}$  とする．

## 3. 実験結果

### 3.1 実験条件

本論文で提案した人物顔の検出および視線検出手法の性能を評価するため実験を行った．使用したデータは，解像度 XGA，フレームレート 25fps，シャッタースピード 1ms のカメラ 15 台を用いて人物を撮影した撮影画像 (図 5)，それらから 3 次元形状復元を行った人物の 3 次元形状，各カメラのキャリブレーションデータである．正面顔の投影では，スクリーンサイズを  $400 \times 400$  pixel， $L_{\text{focus}} = 43.0\text{cm}$ ， $L_{\text{cam}} = 50.0\text{cm}$  とした，なお，正面顔を大きく写した撮影画像 (cam13) は使っていない．また，虹彩の検出ではテンプレート画像に  $300 \times 170$  pixel の図 4 を用いた．虹彩の偏り具合の検出においては， $w = 15, h = 1, W = 10, H = 1, \beta = 0.6$  とし，視線方向の検出では  $\psi = 18^\circ$  とした．

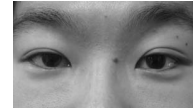


図4 虹彩検出に用いたテンプレート



図5 cam01 - cam15 撮影画像

## 3.2 人物頭部の検出結果

顔の検出には約 5 分間の舞踊から，2 秒毎に抽出した総フレーム数 156 のデータを用いた．提案手法により，図 6 のように正しく顔の位置を検出したフレーム数が 149，その検出精度は 95.5% (表 1) であった．検出に失敗した 7 フレームの内訳は，図 7 のようにお辞儀をして完全に顔が隠れてしまっている (重度に体勢が悪い) ため検出に失敗したものが 2，図 8 のように片手 (着物) で顔を遮ってしまったために失敗したものが 1，人物が下を向いているために失敗したものが 3，体勢は良いのに失敗したものが 1 という結果であった．重度に体勢の悪いフレームを除く検出率は 96.8% であった．また，図 9 のように，顔の検出には成功したが，体勢が悪く 3 次元形状復元が不十分であり，その検出された顔領域が不確かなものが 8 フレームあった．

### 3.2.1 人物頭部検出の考察

提案手法は高い精度で顔を検出した．体勢が良いものに関しては，検出に失敗したものはわずか 1 フレームであった．検出に失敗したフレームは，下を向いているものや，手や着物などの障害物により顔が隠されて

表 1 人物頭部検出の結果

条件	検出率 [%]
正検出 / 総数	95.5 (149/156)
重度に体勢が悪いもの省く	96.8 (149/154)

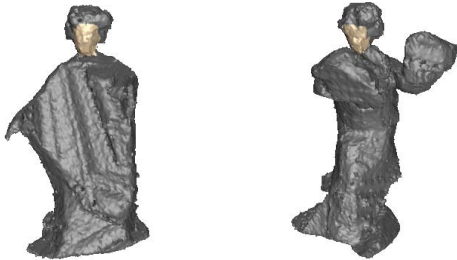


図 6 顔検出に成功した画像



図 7 重度に体勢が悪く顔検出に失敗した例



図 8 片手で顔を隠している、下を向いているため顔検出に失敗した例

いるものであった。人物が下を向いていたフレームは、撮影画像に顔が写っていても、顔に照明が十分に当たらず暗く写っていたために、Viola-and-Jones の顔検出器により撮影画像から顔が検出されていなかった。また、障害物が軽度に顔を隠していたために検出に失敗してしまったもの原因として、実験に用いた 3 次元ビデオを撮影したカメラ 14 台中、顔を撮影したカメラが概ね 3 台程度しかなく、顔候補となる 3 次元形状への投票数が少ない点が上げられる。その顔を撮影した少ないカメラのうち 1 台でも障害物に遮られてしまうと、残るカメラのみで投票を行わざるを得ず、その精度は落ちる。また、投票数が少ないと、着物の紋様など顔以外の部分が顔として検出された場合に、その誤検出による影響が大きくなってしまふ。

### 3.3 正面顔の投影結果

提案手法による正面顔の投影結果として典型的な 6



図 9 顔検出に成功したが、形状復元が不十分で顔領域が不確かな例

表 2 正面顔投影の結果

正面顔の状態	フレーム数 (総数 141)
白目を確認できる	56 (主観値)
白目が確認できない	36 (主観値)
目を瞑っている	31
目を投影できなかった	18

枚を図 10 に示す。人物頭部の検出に成功した 141 フレーム中、図 10 の No1, No2 のように白目を確認できるものが 56 フレーム、No3, No4 のように白目が確認できないものが 36 フレーム、No5, No6 のように人物が目を瞑っているものが 31 フレーム、仮想カメラが障害物に阻まれ目を投影できなかったものが 18 フレームあった (表 2)。なお、白目が確認できるか否かの判定はその線引きが難しいため、主観的数値である。目を瞑っているかの判定は撮影画像を見て判断した。

#### 3.3.1 正面顔投影の考察

No3, No4 のように白目を確認することのできない正面顔が得られてしまった原因として、人物の形状復元誤差やキャリブレーション誤差の影響を受けて、ぼやけや歪みが生じたと考えられる。また、そもそも、正面顔の投影は撮影画像から色情報を獲得するため、必然的に撮影画像を上回る画像は得られない。撮影画像に顔が明瞭に写っていないければ必然的に明瞭な正面顔画像を得ることはできない。実際、図 5 のように撮影画像に写った顔が小さかった。これも、原因として上げられる。

仮想カメラが手や着物などの障害物に阻まれたために目を投影できなかったものが 18 フレームあった。このようなフレームを抑制するためには、障害物が顔と仮想カメラを遮らないよう、動的に仮想カメラの位置を決定するアルゴリズムを導入する必要がある。

#### 3.4 正面顔画像の回転補正の結果

提案した手法により、正面顔の回転補正を正しく行ったものは人物頭部の検出に成功した 141 フレーム中、133 フレーム (94.3%) であった。しかし、左右の虹彩が水平になるまで厳密に補正するまでには至っていない。また、残る 8 フレームは、上下反転したり、横を向いた状態に補正してしまった (上下反転したものが 6、横を向いたものが 2)。

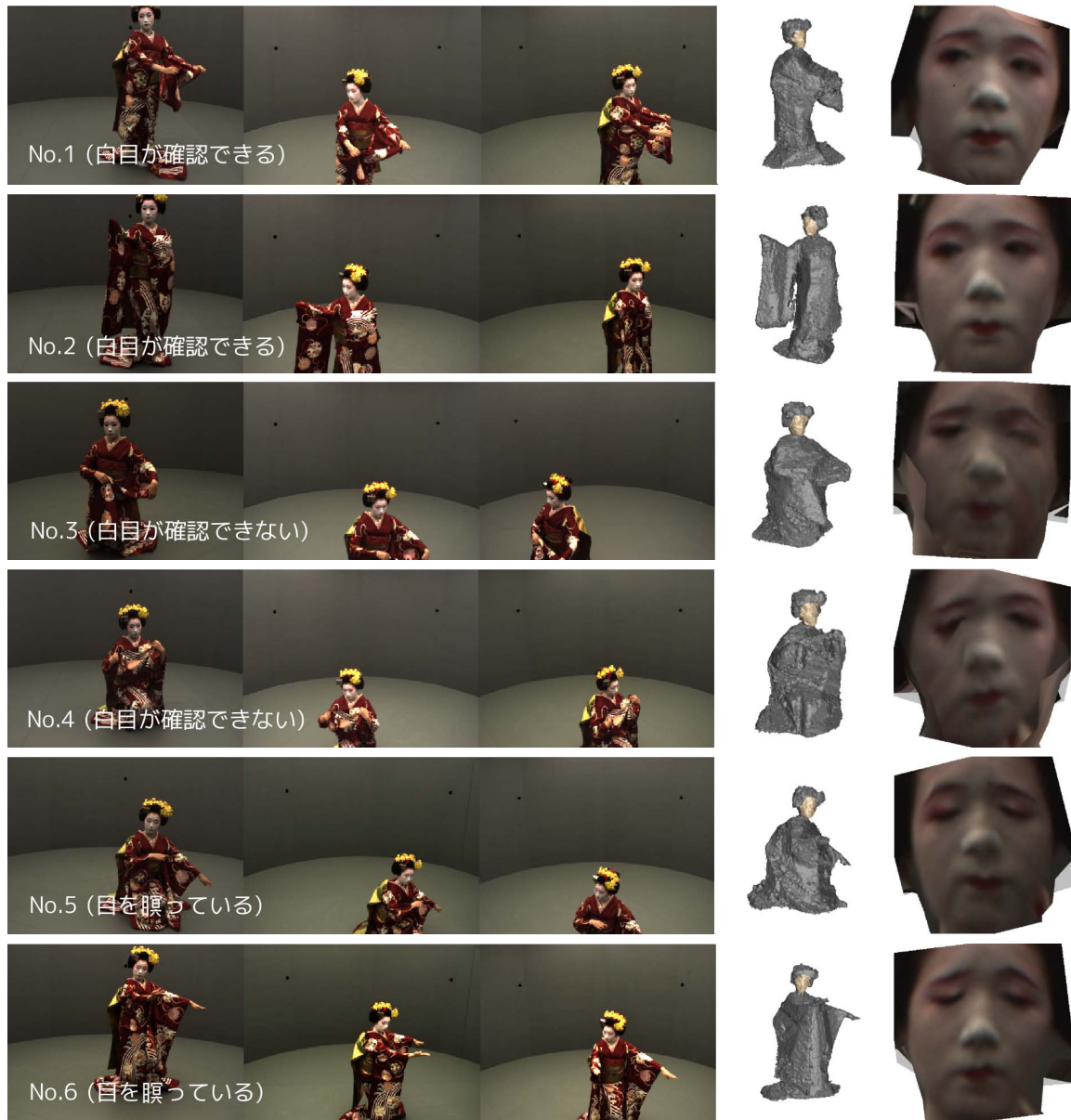


図 10 正面顔の投影結果（顔の回転補正後）撮影画像，3次元形状，投影して得られた正面顔

### 3.4.1 正面顔画像の回転補正の考察

回転補正に失敗したものが8フレームあった。これらはいずれも，2.2.3において，正しく補正する  $N(i) = 1$  なる  $i$  の部分集合  $d$  が，距離 2 以上で分断されて歯抜け状態になったため，誤った部分集合が最大連結集合となっていたために失敗していた。集合  $D$  について，連結しているとみなす距離を大きく取ることでの失敗を回避することができるが，大きくした分だけ補正精度が落ちる。

この正面顔画像の回転を補正する別の方法として，顔の左右対称性を利用する手法を試みた。つまり，顔

方向ベクトル  $\vec{n}_{\text{face}}$  を回転の軸とする対称面候補を回転させ，その面に対して最も3次元形状が対称である面を左右対称面として採用し，その面の傾きから正面顔画像を回転するというものである。しかし，この左右対称性を利用する手法は，90° 顔を横に倒したように補正してしまうものが目立った。補正精度も提案手法に比べると劣っていた。

### 3.5 虹彩の検出結果

提案手法により，虹彩の検出に成功した典型的例を図 11 に，失敗した例を図 12 に示す。白目を確認することのできた正面顔 52 フレーム中，虹彩の検出に成

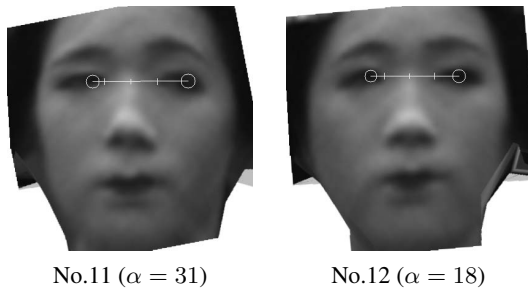
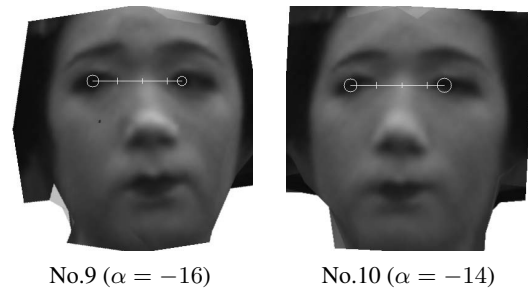
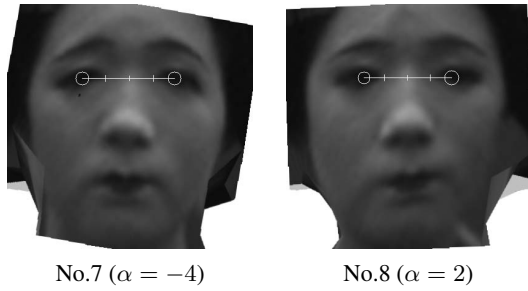


図 11 虹彩検出，虹彩の偏り具合の結果（成功例）

功したものが 47 フレーム（90.4%），失敗したものが 5 フレームあった．これについても，主観的観測値である．失敗した 5 フレームは，いずれも図 12 のように目を検出できてはいるが，虹彩の位置が厳密には誤っているものであった．

### 3.5.1 虹彩検出の考察

主観的観測ではあるが，提案手法は白目が確認できる正面顔に対しては高精度に虹彩を検出した．虹彩の検出に失敗した 5 フレームは，いずれも虹彩ではなく目尻や目頭を虹彩として検出してしまっていた．眉など，目ではない部位を誤検出することはなかった．画像中心から blob までの距離のコスト付けが，この失敗に起因しているかを判断すべく，このコストを省く残りのコストで虹彩の検出を試みたが，白目が確認できる正面顔に対しては完全に等しい結果となった（白目が確認できない正面顔に対しては，このコストを用いた方が眉が検出されてしまうのを抑制していた）．

### 3.6 虹彩の偏り具合の検出結果

虹彩の検出に成功した 47 フレーム中，2.3.2 の目頭

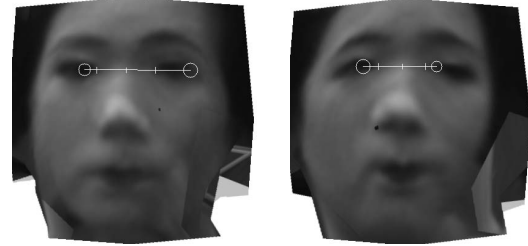


図 12 虹彩検出の結果（失敗例）



図 13 視線方向検出の結果

の検出に失敗したものはなかった．つまり，47 フレーム全てについて  $\alpha$  の値を得ることができた．虹彩の検出に成功した典型的例の  $\alpha$  の値を図 11 に示した．

### 3.6.1 虹彩の偏り具合検出の考察

虹彩の検出に成功した 47 フレームでは，全てのフレームについて  $\alpha$  の値を取得することができた．その  $\alpha$  の値は，直感的に捉えた視線の方向と大きはずれないことから，本論文で導入した虹彩の偏り具合  $\alpha$  は，視線の定量的な尺度と言えるだろう．

評価実験に用いたデータには，左右の目の間に大きなほくろやあざなどの黒い領域は無かったが，それらがあるとその黒い領域の端点を目頭として検出してしまう恐れ，また，式 (6) の範囲内の平均輝度値  $\bar{l}_{um}$  が大きくなってしまいうため正しく目頭を検出できない恐れがある．

### 3.7 視線方向検出の結果及び考察

視線方向ベクトル検出結果の例（No.9, No.11）を図 13 に示した．主観的観測ではあるが，虹彩の偏り具合  $\alpha$  を基に計算した 3 次元視線方向ベクトルは，撮影画像の虹彩位置から推測される視線方向と比べて，大きくずれていないと言える．

## 4. 結 論

### 4.1 人物頭部の検出

人物頭部を検出するために，これまで平面における顔の位置を検出する研究が数多くなされてきたが，それらはいずれも 2 次元にとどまっている．本論文では人物頭部の 3 次元空間における位置の検出を試みた点

が新しい。提案手法を用いた評価実験では、非常に高い精度で顔を検出することができた。提案手法は、顔を撮影したカメラが少ない3次元ビデオの場合、撮影画像からの顔検出において顔ではない部位が顔として検出されたり、手などの障害物がある顔を撮影したカメラを遮ってしまうと、誤検出をする可能性が高まる特性を持っていることが明らかになった。これは、顔候補となる3次元形状への投票数が不十分となるからである。このことから、顔候補となる3次元形状への投票機会を増やすことが、検出精度向上につながると言える。その方法として次のものが考えられる。

- 2次元画像からの顔検出に、正面顔のみではなく、左顔や右顔を検出するようにし、検出された左・正面・右顔の組み合わせにより、顔の位置を評価する(3次元形状への投票機会が左・右顔の分増える)
  - 時間方向の顔位置の連続性も評価に加える
- 横顔も含めた検出を行うためには、2次元画像から精度良く横顔を検出することのできる手法を考案しなければならない。

また、提案手法の性質から、より多くのカメラが顔を撮影した3次元ビデオが入力として与えられれば、提案手法による検出精度はより高いものとなるだろう。

#### 4.2 視線の検出

視線を推定するために、これまで顔を写した画像を解析することで平面上での虹彩の位置を検出する研究、赤外線カメラを用いて眼球を直接撮影することで視線を得る研究などがなされてきた。自由な状態で運動している対象人物の頭部および視線を非接触・非拘束に視線方向を推定にあたり問題となるのは、必ずしも常に対象人物の正面から撮影し続けることができないことである。本論文では、仮想的に新たにカメラを設け、正面顔を常に投影するというアイデアの下、特殊なカメラを用いずに3次元ビデオから視線を推定する手法を提案した。

評価実験の結果、顔の検出に成功し、人物が目を瞑っていない110フレーム中、正しく視線方向を推定することのできたもの(主観的観測)は47フレーム(42.7%)であった。本論文で導入した、虹彩の偏り具合 $\alpha$ を基に推定した視線方向は、その成功した47フレームについては、撮影画像を見ることで推定した視線方向と大きな相異はなかった。

視線方向検出の失敗詳細は、(1)仮想カメラが手や着物などの障害物により遮られたために正面顔の投影に失敗した、(2)白目を確認できない正面顔が投影されてしまった、(3)正面顔画像の回転補正に失敗した、(4)虹彩の検出に失敗した、というものであった。(1)については、仮想カメラが障害物により遮られないように、仮想カメラの位置を動的に計算するアルゴリズムを導入することで回避できるだろう。(2)については、人物の形状復元誤差やキャリブレーション誤差の

影響を受けて、白目を確認できない正面顔が得られてしまったと考えられる。(3)についての新たな手法として、ラグビーボールのような楕円球モデルを人物顔の3次元形状とマッチングし、楕円球の長軸と短軸の傾きから、顔の左右対称面を検出することにより、回転を補正する方法が考えられる。また、この段階で高精度に回転を補正することができれば、後のblob対の検出において厳しい制約を設けることができ、眉が検出されてしまうのを抑制することができる。(4)については、正面顔の投影が改善され、より高解像度ではっきりした正面顔画像を獲得することができるになれば、現状のアルゴリズムでの精度も向上すると考えられる。また、虹彩検出のコスト付けとして、時系列運動情報を利用して、その虹彩の位置のコスト付けを行う方法も考えられるが、他の部位に比べ眼球の動きは非常に早く、フレーム間で対応を得るのは難しい。従って、このコスト付けを行うには、ハイスピードカメラなど、新たなカメラを導入しなければならないだろう。

これらの改善を行い、より高精度に視線を検出することが今後の課題である。また、さらに視線の定量的な検出精度評価実験を行いたい。

#### 謝 辞

本研究の一部は、文部科学省グローバルCOEプログラム「知識循環社会のための情報学教育研究拠点」および「大型有形・無形文化財の高精度デジタル化ソフトウェアの開発」プロジェクトの補助を受けて行った。本論文の執筆にあたりご協力を頂いた高井勇志特任助教、Lyndon Hill 特任助教に感謝致します。

#### 参 考 文 献

- 1) T. Matsuyama, X. Wu, T. Takai, and S. Nobuhara: Real-Time 3D Shape Reconstruction, Dynamic 3D Mesh Deformation and High Fidelity Visualization for 3D Video. CVIU(2004): Vol.96 No.3, pp.393-434.
- 2) Paul A Viola, Michael J. Jones: Robust Real-Time Face Detection. ICCV(2004): Vol.57 No.2, pp.137-154.
- 3) 山田真己, 中島朗子, 福井和広: 因子分解法と部分空間法による顔向き推定. 信学技報 (2002): PRMU2001-194, pp.1-8.
- 4) 川口 剛, リゾン モハメッド, 日高 大輔: ハフ変換と分離度フィルタによる人物顔からの両目の検出. 電子情報通信学会論文誌 (2001): Vol.J84-D2 No.10, pp.2190-2200.
- 5) 竹上 健, 後藤 敏行, 大山 玄: 視線方向計測のための高精度瞳孔検出アルゴリズム. 電子情報通信学会 (2003): Vol.J86-D2 No.2, pp.252-261.