

過学習を考慮した AAM パラメータの選択と回帰分析による顔・視線同時推定

高谷 学[†] 滝口 哲也^{††} 有木 康雄^{††}

[†] 神戸大学工学部情報知能工学科 〒 657-8501 兵庫県神戸市灘区六甲台町 1-1

^{††} 神戸大学自然科学系先端融合研究環 〒 657-8501 兵庫県神戸市灘区六甲台町 1-1

E-mail: [†]takatani __ m@me.cs.scitec.kobe-u.ac.jp, ^{††}{takigu,ariki}@kobe-u.ac.jp

あらまし 視線推定は、コンピュータビジョンの分野において重要な技術であり、インターフェースやロボット、人間の監視、支援など、幅広い応用が期待されている。近年では、比較的デバイスのコストも低く、被験者への負担の小さい単眼カメラによる視線推定手法が提案されている。本研究では、Active Appearance Models のパラメータを特徴量として抽出し、赤池情報量規準 (AIC) や最小記述長規準 (MDL) などの規準によって推定に必要な特徴量を選択し、それらの特徴量と顔、視線方向の幾何的な関係をモデル化し、回帰分析を用いて顔、視線方向の同時推定手法を提案する。実験の結果、提案手法は従来の単眼カメラによる視線推定の手法に比べ、平均誤差に約 1.4 ° (約 33%) の改善が見られた。

キーワード 視線推定, Active Appearance Models, 回帰分析, 赤池情報量規準

Simultaneous Estimation of Face and Gaze Orientation by Regression Analysis With Information Criterion

Manabu TAKATANI[†], Tetsuya TAKIGUCHI^{††}, and Yasuo ARIKI^{††}

[†] Department of Computer Sciences and Systems Engineering, Kobe University Rokkodai 1-1, Nada-ku, Hyogo, 657-8501 Japan

^{††} Organizatoin of Advanced Science and Technology, Kobe University Rokkodai 1-1, Nada-ku, Hyogo, 657-8501 Japan

E-mail: [†]takatani __ m@me.cs.scitec.kobe-u.ac.jp, ^{††}{takigu,ariki}@kobe-u.ac.jp

Abstract It is an important technique to estimate gaze orientation automatically in Computer Vision. The application using gaze estimation for interface, robot, and human interaction is expected. In recent years, many approaches for gaze estimation using only a monocular camera with low cost were proposed. In these approaches, the problem was the gaze orientation estimation error mainly caused by the face orientation estimation. To solve this problem, we propose simultaneous estimation of face and gaze orientation by regression using AAM parameters selected by AIC. In the experimental result, the average estimation error was improved by 1.4 degrees (33%) compared with the conventional method. We confirmed the validity of our method through this experimental result.

Key words Gaze estimation, Active Appearance Models, Regression analysis, Akaike Information Criterion

1. はじめに

人間は主に視覚から周囲の環境を知覚、認識する [1]. 人間の視線を推定する技術はロボットビジョン、エージェント開発などの研究において非常に重要であり、ノンバーバルコミュニケーションを行うロボット [2] や、視線情報に基づく動画画像解析 [3], ドライバの眠気感知 [4] など、幅広いアプリケーションへの応用が期待される技術である。

視線推定の研究は近年盛んに行われており、これらの

研究では高精度に視線を推定するために様々なアプローチが提案されてきた。これらのアプローチは大きく 2 種類に分類できる。一つはカメラ以外に、赤外線カメラやヘッドマウントカメラ等の特殊なデバイスを用いて、視線を高精度に推定する手法である [6] [7]。さらに、もう一つのアプローチとして、単眼カメラ画像からの視線推定手法が挙げられる。前者のアプローチに比べ、後者は推定精度が低いが、デバイスのコスト、被験者への負担という点においては優れており、近年様々な手法が提案されている [8] [9] [16]。本研究では、デバイスのコストや、

人間への身体的負担の少ない単眼カメラによる視線の高精度な推定を目的とする。

従来の単眼カメラによる視線推定の手法として、黒眼移動量から視線を推定する手法 [5]、3次元眼球モデルによる視線推定 [10] [11] などが挙げられる。これらの手法では、顔方向推定の後に視線方向推定を行うため、顔方向の推定誤差が視線方向の推定誤差に影響し、視線推定誤差が大きくなり、高精度な視線推定が行えないという問題がある。

本研究では、その問題を解決するため、Cootes らの提案する Active Appearance Models (以下 AAM と略す) [12] の combined パラメータ c を特徴量として抽出し、その特徴量と顔、視線方向の関係を幾何的な位置関係を基にモデル化する。次に、このモデルを基に回帰分析を行うことで、顔、視線方向を同時推定し、より高精度に視線推定を行う手法を提案する。

AAM は顔、視線方向の学習画像列を与えることで、比較的次元の低いパラメータで顔、視線の方向変化に関する顔画像の形状とテクスチャの変化を制御するモデルである。そのため、顔の方向が変化しても特徴量を比較的頑健に抽出可能であるため、本研究では AAM を特徴量抽出の手法として採用した。

さらに、本研究では赤池情報量規準 (AIC) [13] や最小記述長規準 (MDL) などの規準を用いて回帰式の汎化能力を評価し、考えられる特徴量の全ての組み合わせのうち、最も汎化能力の高い回帰式を選択する手法を提案する。これにより、視線や顔方向の推定に不要な次元が、回帰係数の過学習を引き起こすことを緩和し、高精度な視線推定を実現できる。

本稿の構成は次の通りである。2 章で提案手法の流れについて述べ、3 章で Viola らの提案する顔領域探索手法と顔、視線方向同時推定のための特徴量を抽出する手法である AAM について述べる。4 章では、得られた特徴量と、顔、視線の関係をモデル化する方法を提案する。さらに、5 章では回帰分析の過学習を緩和するための赤池情報量規準による特徴量選択の手法と回帰分析の方法について述べる。6 章では提案手法の有効性を示すために行った実験の概要と結果、及び考察を述べる。最後に、7 章では結論と今後の研究について述べる。

2. 提案手法の流れ

図 1 はこれまでに述べた提案手法における処理の流れの概略図である。提案手法では入力画像が与えられると、Viola の提案する手法 [14] により顔領域探索を行う。次に、構築しておいた AAM を用いて、入力画像と最も類似する画像を生成する AAM のパラメータを最急降下法により決定し、このパラメータを特徴量として抽出する。そして AIC によって選択された回帰式を用いて回帰分析を行い、図 2 のように視線と顔方向を同時に推定する。

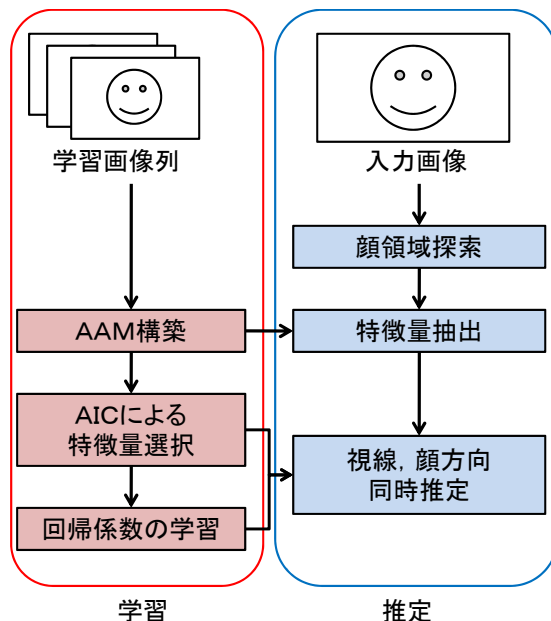


図 1 提案手法における処理の流れ

図 2 において緑色の矢印は視線方向、赤色の矢印は顔方向、矢印の大きさは角度の大きさを表している。



図 2 顔、視線方向同時推定の推定結果

3. 特徴量抽出

入力画像から AAM のパラメータを決定する時、パラメータの値を最急降下法によって変化させながら、入力画像とモデルのマッチングを繰り返し行って探索を行っている。従って、入力画像中における初期探索位置が、顔領域から大きく離れていると安定して AAM のパラメータを決定することができない。

本研究では安定して特徴量を抽出するために、Viola らの提案する顔領域探索手法 [14] を用いて顔領域を探索し、得られた顔領域を初期探索点として AAM のパラメータ探索を行っている。

本章では、Viola らの顔領域探索手法と、AAM によるモデル構築、及びパラメータの決定方法について述べる。

3.1 顔領域探索

AAM を用いて安定に特徴量を抽出するために、Viola らの提案する手法 [14] を用いて顔領域を探索する。この

手法は顔領域探索において一般的に用いられている手法であり、安定して比較的高速に顔領域の探索が可能である [15].



図 3 Haar-like 特徴量抽出に用いられる矩形

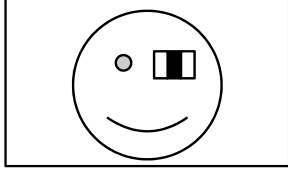


図 4 Haar-like 特徴量抽出のための矩形の当てはめ

この手法では、図 3 に示す矩形領域を、図 4 のように当てはめ、白と黒の各矩形領域間の平均輝度の差を haar-like 特徴量として抽出し、顔を判別するのに有効な矩形の位置、種類、アスペクト比、スケールを弱識別機として AdaBoost によって学習し、それらを線形結合して強識別機を構成し、顔領域を検出する。

3.2 Active Appearance Models

顔画像から特徴量を抽出する際、被験者が必ずしも正面を向いているとは限らない。Cootes らの提案する Active Appearance Models(AAM) [12] は、視線、顔方向の変化のある画像からモデルを構築するため、被験者の顔方向の変化に対して比較的精度良く特徴量を抽出できる。本研究では、AAM を用いて顔画像から特徴量の抽出を行う。以下では具体的に AAM について述べる。

AAM では、モデルの形状 (Shape) $\mathbf{s}$ とテクスチャ (Texture) $\mathbf{g}$ の変化を低次元のパラメータ $\mathbf{c}$ で表現する。 $\mathbf{s}$, $\mathbf{g}$ は式 (1), (2) で表現されるベクトルである。

$$\mathbf{s} = (x_1, y_1, \dots, x_n, y_n)^T \quad (1)$$

$$\mathbf{g} = (g_1, \dots, g_m)^T \quad (2)$$

ここで、 x_i, y_i ($i \leq n$) は各特徴点における座標を表している。学習画像に与えられたベクトル $\mathbf{s}$ を、回転、拡大縮小、位置の正規化を行い、学習画像列から平均形状 $\bar{\mathbf{s}}$ を求める。 g_j ($j \leq m$) は、平均形状 $\bar{\mathbf{s}}$ に画像を正規化したときの $\bar{\mathbf{s}}$ 内部での各画素の輝度値であり、学習画像列から平均テクスチャ $\bar{\mathbf{g}}$ を求める。

平均形状 $\bar{\mathbf{s}}$ 、平均テクスチャ $\bar{\mathbf{g}}$ を用いて学習画像に対して主成分分析を行い、正規直交行列 $\mathbf{P}_s$, $\mathbf{P}_g$ を求め、得られた行列 $\mathbf{P}_s$, $\mathbf{P}_g$ を用いて、式 (3)(4) のように $\mathbf{s}_g$ をモデル化する。

$$\mathbf{s} = \bar{\mathbf{s}} + \mathbf{P}_s \mathbf{b}_s \quad (3)$$

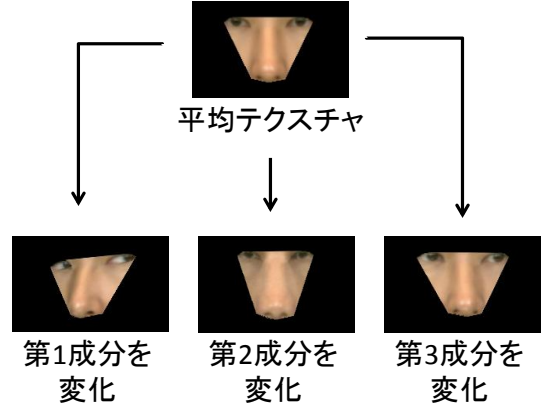


図 5 combined パラメータの各成分を変化させた時のモデル画像の変化 (第 1 成分は顔方向と視線方向の変化、第 2 成分は顔の縦の形状の変化、第 3 成分は視線方向の変化がそれぞれ表出している)

$$\mathbf{g} = \bar{\mathbf{g}} + \mathbf{P}_g \mathbf{b}_g \quad (4)$$

ここで、 $\mathbf{b}_s, \mathbf{b}_g$ はそれぞれ主成分の係数であり、部分空間内で平均からの変化を表すパラメータである。さらに、 $\mathbf{b}_s$, $\mathbf{b}_g$ を式 (5) のように連結し $\mathbf{b}$ とする。

$$\mathbf{b} = \begin{pmatrix} \mathbf{W}_s \mathbf{b}_s \\ \mathbf{b}_g \end{pmatrix} \quad (5)$$

$\mathbf{W}_s$ は形状とテクスチャの単位を正規化するための行列である。さらに、このベクトル $\mathbf{b}$ に対しても主成分分析を行い正規直交行列 $\mathbf{Q}$ を求める。得られた行列 $\mathbf{Q}$ を用いて $\mathbf{b}$ を式 (6) のように表現する。

$$\mathbf{b} = \mathbf{Q} \mathbf{c} = \begin{pmatrix} \mathbf{Q}_s \\ \mathbf{Q}_g \end{pmatrix} \mathbf{c} \quad (6)$$

$\mathbf{c}$ は主成分の係数であり、combined パラメータベクトルと呼ばれる。式 (5) と式 (6) から式 (7) と式 (8) が得られる。

$$\mathbf{W}_s \mathbf{b}_s = \mathbf{Q}_s \mathbf{c} \quad (7)$$

$$\mathbf{b}_g = \mathbf{Q}_g \mathbf{c} \quad (8)$$

ここで、式 (7) と式 (8) より、形状 $\mathbf{s}$ とテクスチャ $\mathbf{g}$ を $\mathbf{c}$ で表現すると式 (9)(10) のようになる。

$$\mathbf{s}(\mathbf{c}) = \bar{\mathbf{s}} + \mathbf{P}_s \mathbf{W}_s^{-1} \mathbf{Q}_s \mathbf{c} \quad (9)$$

$$\mathbf{g}(\mathbf{c}) = \bar{\mathbf{g}} + \mathbf{P}_g \mathbf{Q}_g \mathbf{c} \quad (10)$$

このようにして、低次元のパラメータベクトル $\mathbf{c}$ によって、顔や視線の方向を変化させた顔画像の形状 $\mathbf{s}$ とテクスチャ $\mathbf{g}$ をモデル化できる。

よって、 $\mathbf{c}$ には顔、視線方向に関する情報が含まれていると考えられる。図 5 はパラメータベクトル $\mathbf{c}$ の各成

分を変化させた場合に、顔画像がどのように変化するか、その様子を表したものである。本研究では、入力画像からパラメータベクトル $\mathbf{c}$ を抽出し、顔、視線方向の同時推定を行う。以下では、入力画像から特徴量を抽出する方法について述べる。

入力画像 I に対して、モデル画像 $\mathbf{g}(\mathbf{c})$ との誤差 e を式 (11) のように求め、誤差を最小にするパラメータベクトル $\mathbf{c}$ 及び、 $\mathbf{p}$ を最急降下法によって求める。ここで、 $\mathbf{W}$ は入力画像に対してアフィン変換を行う関数である。

$$e(\mathbf{c}, \mathbf{p}) = \|\mathbf{g}(\mathbf{c}) - I(\mathbf{W}(\mathbf{p}))\| \quad (11)$$

ただし、 $\mathbf{p}$ は、入力画像をアフィン変換するための拡大、回転、平行移動のパラメータベクトルであり、pose パラメータベクトルと呼ばれる。 $I(\mathbf{W}(\mathbf{u}; \mathbf{p}))$ は入力画像をパラメータ $\mathbf{p}$ によってアフィン変換したときの画像である。

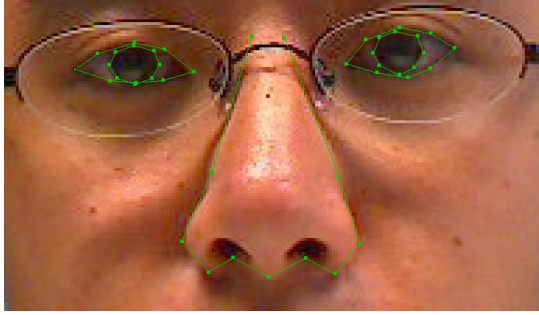


図 6 AAM 構築に用いた 43 点の特徴点 (緑色の点が特徴点)

このようにして、入力画像からパラメータベクトル $\mathbf{c}$ を特徴量として抽出できる。本研究では、図 6 のように眼と鼻を中心とした 43 点の特徴点を用いて AAM を構築した。

4. 回帰式の提案

本章では、Cootes らの提案する顔方向推定手法を拡張して、combined パラメータベクトル $\mathbf{c}$ から顔方向 ϕ 及び、視線方向 θ を回帰分析によって同時推定するための回帰式の提案を行う。

combined パラメータベクトル $\mathbf{c}$ と顔方向 ϕ 及び、視線方向 θ の幾何的な関係をモデル化し、回帰式の導出を行う。Cootes らは顔方向のみの推定のために $\mathbf{c}$ と顔方向 ϕ の関係をモデル化した回帰式を提案している [17]。この回帰式は Peter らによって任意方向の顔画像合成の研究などに採用されている。[18]。本研究ではこの式を視線方向に拡張して、顔、視線方向を同時推定するための回帰式の提案を行う。以下では回帰式の導出の方法について述べる。

図 7 は、顔方向、視線方向が正面を向いている人間の頭部方向から見た時の概略図である。Image plane は画像平面、 R は顔の半径、 r は眼球の半径、 α は顔上での

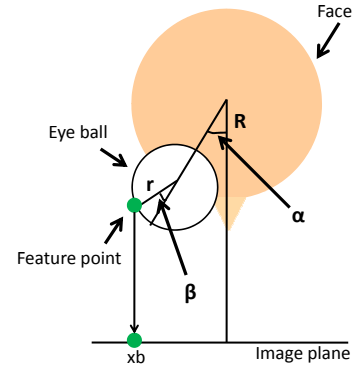


図 7 顔方向=0, 視線方向=0 の時の特徴点の位置

眼球の角度、 β は眼球上での特徴点（例えば黒眼周辺上の点）の角度を表している。ただし、眼球中心が顔の円周上にあると仮定する。このとき、特徴点は画像平面上で座標 x_b の位置に映っている。

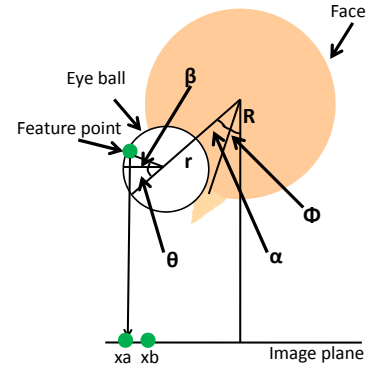


図 8 顔方向= ϕ , 視線方向= θ の時の特徴点の位置

図 8 は図 7 の状態から、顔水平方向が ϕ 、視線水平方向が θ だけ回転したときの頭部の様子を表した概略図である。図 8 において、特徴点は画像平面上の座標 x_a の位置に映っている。このとき、回転による特徴点の画像平面上での変位 Δx は式 (12) によって表現できる。

$$\begin{aligned} \Delta x &= x_a - x_b \\ &= R \sin(\phi + \alpha) + r \sin(\phi + \theta + \alpha + \beta) \\ &\quad - R \sin \alpha - r \sin(\alpha + \beta) \\ &= r \sin(\phi + \theta) \cos(\alpha + \beta) \\ &\quad + r \cos(\phi + \theta) \sin(\alpha + \beta) \\ &\quad + R \sin \phi \cos \alpha + R \cos \phi \sin \alpha \\ &\quad - R \sin \alpha - r \sin(\alpha + \beta) \\ &= r \sin \theta \cos(\alpha + \beta) + r \cos \theta \sin(\alpha + \beta) \\ &\quad + R \sin \phi \cos \alpha + R \cos \phi \sin \alpha \\ &\quad - R \sin \alpha - r \sin(\alpha + \beta) \end{aligned} \quad (12)$$

ただし、 $\hat{\theta} = \phi + \theta$ であり、 $\hat{\theta}$ は画像平面上に対する視線の絶対角度である。

本研究では、この Δx と AAM のパラメータベクトル $\mathbf{c}$ は関係があると仮定して、式 (12) の α, β, R, r を定数とみなして式 (13) のように回帰式を導出する。

$$\mathbf{c} = \mathbf{c}_0 + \mathbf{c}_1 \cos \phi + \mathbf{c}_2 \sin \phi + \mathbf{c}_3 \cos \hat{\theta} + \mathbf{c}_4 \sin \hat{\theta} \quad (13)$$

ここで、 $\mathbf{c}_0, \mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4$ は回帰係数ベクトルである。このようにして回帰式が導出できる。

次に、5 章では回帰係数ベクトルの学習方法と、AIC を用いた顔、視線方向同時推定の方法について述べる。

5. 特徴量選択と視線、顔方向同時推定

本章では、4 章で導出した回帰式 (13) を用いて回帰分析を行い、顔、視線方向を同時に推定する手法を提案する。式 (13) において、求めたい変数は $\cos \phi, \sin \phi, \cos \hat{\theta}, \sin \hat{\theta}$ 、特徴量とする変数はパラメータベクトル $\mathbf{c}$ の各成分である。目的変数と説明変数の関係を明示するために、 $\mathbf{c}$ の各成分を $\mathbf{c} = \begin{pmatrix} e_1 & \dots & e_d \end{pmatrix}$ として、式 (13) を式 (14) のように変形する。

$$\begin{pmatrix} \cos \phi \\ \sin \phi \\ \cos \hat{\theta} \\ \sin \hat{\theta} \end{pmatrix}^T = \begin{pmatrix} 1 \\ e_1 \\ \vdots \\ e_d \end{pmatrix}^T \begin{pmatrix} a_{0,1} & \dots & a_{0,4} \\ \vdots & \ddots & \vdots \\ a_{d,1} & \dots & a_{d,4} \end{pmatrix} \quad (14)$$

ここで、 $a_{i,j}$ ($0 \leq i \leq d, 1 \leq j \leq 4$) は回帰係数であり、 d はパラメータベクトル $\mathbf{c}$ の次元数である。

式 (14) において、AAM のパラメータベクトル $\mathbf{c}$ の全ての成分を用いて顔方向と視線方向をモデル化している。しかし、パラメータベクトル $\mathbf{c}$ には、顔方向や視線方向に関係のある成分だけが含まれるとは限らない。そのため、回帰係数を学習する際、過学習が起こり汎化能力が低下し、未学習データに対する推定の精度が悪くなる。そこで、本研究では赤池情報量規準 (AIC) [13] を用いて、回帰式の汎化能力を評価し、特徴量 $\mathbf{c}$ の成分の選択を行い、より高精度な視線を推定する手法を提案する。

5.1 節では回帰係数の学習方法について述べ、AIC を用いて回帰式の性能を評価する方法について述べる。そして、5.2 節では 5.1 節で求めた回帰式と回帰係数から顔、視線方向を同時推定する方法について述べる。

5.1 AIC による回帰式の評価

式 (14) の回帰係数の学習において、目的変数を推定するのに有効でない説明変数が含まれている場合、過学習が起こり汎化能力が低下し、未知のデータに対して推定の精度が悪くなる。従って、過学習が起きるのを防ぐため回帰式の汎化能力を評価する必要がある。

回帰式の性能を評価する方法として、一般的に赤池情報量規準 (AIC) [13] や最小記述長規準 (MDL) [20] が

用いられる。本研究では、実験で最も良い結果となった AIC を提案手法として採用する。以下では回帰係数の学習方法と、AIC によって回帰式の性能を評価する方法について述べる。

式 (14) において、目的変数 $\cos \hat{\theta}$ について着目すると、回帰式は式 (15) のようになる。

$$\cos \hat{\theta} = a_{0,3} + \sum_{i=1}^d a_{i,3} e_i + \epsilon \quad (15)$$

ここで、 ϵ は推定誤差を表し、平均 0、分散 σ^2 の正規分布に従う確率変数である。

説明変数 $e_1, \dots, e_d$ のうち、回帰分析に用いる説明変数の集合を S_k 、回帰式の番号を k とすると、説明変数の組み合わせは $k = 2^d$ 通り考えられる。そこで、 k の 2 進数に関連づけて以下のように S_k の変数の集合を定義する。

$$\begin{aligned} S_0 &= \{\} \\ S_1 &= \{e_1\} \\ S_2 &= \{e_2\} \\ S_3 &= \{e_1, e_2\} \\ S_4 &= \{e_3\} \\ &\vdots \\ S_{2^d-1} &= \{e_1 \dots e_d\} \end{aligned}$$

よって k 番目の回帰式は式 (15) より式 (16) のように表現できる。

$$\cos \hat{\theta}_k = a_{0,3} + \sum_{i \text{ s.t. } e_i \in S_k} a_{i,3} e_i + \epsilon \quad (16)$$

回帰分析のための学習データの集合を X 、要素数を n とすると、AAM により抽出した特徴量 $\mathbf{c}^t$ の成分 e_i^t ($t \leq n$) と、与えられた顔方向と視線方向の和 $\hat{\theta}^t$ の組が X に含まれている。回帰係数を学習するために、それらの学習データを式 (16) に関して縦に並べ、式 (17) のように表現する。

$$\begin{pmatrix} \cos \hat{\theta}_k^1 \\ \vdots \\ \cos \hat{\theta}_k^n \end{pmatrix} = \begin{pmatrix} 1 & e_1^1 & \dots & e_d^1 \\ & \vdots & & \vdots \\ 1 & e_1^n & \dots & e_d^n \end{pmatrix} \begin{pmatrix} a_{0,3} \\ \vdots \\ a_{\tilde{d},3} \end{pmatrix} \quad (17)$$

ここで、 $\tilde{d}$ は、 S_k に含まれる説明変数の要素数とする。式 (17) において、 $\begin{pmatrix} a_{0,3} & \dots & a_{\tilde{d},3} \end{pmatrix}^T$ について解くと式 (18) で表せる。

$$\begin{pmatrix} a_{0,3} \\ \vdots \\ a_{\tilde{d},3} \end{pmatrix} = \mathbf{A}^+ \begin{pmatrix} \cos \hat{\theta}_k^1 \\ \vdots \\ \cos \hat{\theta}_k^n \end{pmatrix} \quad (18)$$

ただし、 $\mathbf{A}^+$ は行列

$$\begin{pmatrix} 1 & e_1^1 & \dots & e_d^1 \\ & & \ddots & \\ 1 & e_1^n & \dots & e_d^n \end{pmatrix}$$

の一般化逆行列（疑似逆行列）[19] である．これにより、 k 番目の回帰式における回帰係数が得られる．この方法によって得られる回帰係数は最小二乗法によって得られる回帰係数と等価である．

得られた回帰式を用いて、学習データ X で回帰分析を行い、回帰式の最大対数尤度 $l(\hat{\Theta}_k; X)$ を求める． $\hat{\Theta}_k$ は、回帰式の推定パラメータベクトルであり、 S_k に含まれる変数、及び定数係数 $a_{0,3}$ 、誤差の分散 σ^2 から成る．最大対数尤度は学習データ X への当てはまりの良さを表しており、式 (16) において、学習データ X と回帰式の予測推定誤差 ϵ の分散である $\hat{\sigma}_k^2$ を回帰式と学習データから求め、 $\hat{\sigma}_k^2$ を用いて式 (19) のように近似する [21]．

$$l_k(\hat{\Theta}_k; X) = -\frac{n}{2}(1 + \log(2\pi\hat{\sigma}_k^2)) \quad (19)$$

学習データへの当てはまりが良いとき、 $\hat{\sigma}_k^2$ は小さくなるので、結果として式 (19) の最大対数尤度が大きくなる．

一般的に回帰式の自由度（パラメータの数）が増加すると、最大対数尤度は大きくなる．しかし前述の通り、モデルの自由度が増えると過学習の原因となる不要な次元が含まれる可能性が高い．AIC では、その問題を緩和させるため、最大対数尤度に対してモデルの自由度のペナルティを与え回帰式を式 (20) で評価している．

$$AIC_k = n(1 + \log(2\pi\hat{\sigma}_k^2)) + 2(\tilde{d} + 2) \quad (20)$$

式 (20) で計算される AIC の値は小さいほど良いモデル式であることを示しており、最大対数尤度に差のない二つのモデルに対して、説明変数の少ない方のモデルが選ばれる．よって AIC では少ない説明変数で効率よく学習データを表現できるモデルが選ばれる．

本研究では、考えられる特徴量 $\mathbf{c}$ の成分の全組合せ S_k ($k \leq 2^d$) から最も AIC が小さくなる回帰式を選択し、回帰分析に用いる．

5.2 節ではこうして得られた回帰式から視線と顔方向を同時推定する方法について述べる．

5.2 顔、視線方向同時推定

本節では、5.1 節によって求めた最適な回帰式、及び回帰係数から顔、視線方向を同時推定する方法について述べる．

各目的変数について 5.1 節の方法で回帰係数を求めると、式 (14) における回帰係数行列が求めることができる．ただし、 S_k に含まれなかった説明変数の係数は 0 と

する．入力画像から AAM により特徴量 $\mathbf{c}$ が抽出されると、式 (14) に代入して、式 (21) のように値が得られたとする．

$$\begin{pmatrix} \cos \phi \\ \sin \phi \\ \cos \theta \\ \sin \theta \end{pmatrix} = \begin{pmatrix} \hat{z}_1 \\ \hat{z}_2 \\ \hat{z}_3 \\ \hat{z}_4 \end{pmatrix} \quad (21)$$

このとき、顔方向 ϕ 及び、顔と視線方向の和 θ は式 (22)(23) で与えられる．

$$\phi = \tan^{-1} \frac{\hat{z}_2}{\hat{z}_1} \quad (22)$$

$$\theta = \tan^{-1} \frac{\hat{z}_4}{\hat{z}_3} \quad (23)$$

このようにして、顔方向、視線方向を同時に推定できる．

6. 評価実験

表 1 実験に用いたデータベースの概要

被験者数	4 名
学習画像枚数	63 枚 / 人
テスト画像枚数	252 枚 / 人
画像の解像度	640 × 480
被験者の顔角度	-20,-15, ..., 15,20 (deg)
被験者の視線角度	-20,-15, ..., 15,20 (deg)

提案手法が従来手法に比べ、視線水平方向の推定において有効であることを検証するために、評価実験を行った．

6.1 実験条件

評価実験に用いる画像のデータベースの概要を表 1 に示す．

AAM の構築に一人の人間の画像を使用し、学習人物と同一人物の画像を入力画像として与えなければ、AAM による特徴量抽出の精度が低下する [22]．よって実験では各個人のデータに対してそれぞれ AAM を構築し、学習人物と同一の人物の画像を入力画像として与え、実験を行った．

比較する従来手法として、Ishikawa らによって提案されている手法を採用した [16]．この手法では、従来の単眼カメラにおける視線推定において、本実験と同一条件で実験を行った場合、最も精度が高いと考えられる．本研究では、この Ishikawa らの手法 (Conventional) と提案手法である AIC を用いて回帰分析を行う手法 (with AIC) の比較を行う．さらに、AIC を用いて特徴量を選択したことが視線推定の精度向上へどの程度貢献しているかを検証するために、特徴量選択を用いないで式

(14) の特徴量 c 全成分を用いて回帰分析を行った手法 (without AIC), MDL を用いて特徴量選択を行った手法 (with MDL) の計 4 つの手法を比較した。

評価の方法は、視線水平方向の平均推定誤差、及び正解率で評価を行う。正解率とは、推定誤差が 2.5(deg) 以内になった確率である。ここで、誤差の許容範囲を 2.5(deg) と設定したのは、主に以下に述べる理由が考えられる。それは、実験に用いたデータベースの顔、視線方向の角度が 5.0(deg) 単位で与えられているからであり、このデータベースを用いて評価できる精度に限界があると考えられる。

6.2 実験結果と考察

実験の結果を表 2 に示し、平均推定誤差と正解率のグラフを図 9 と図 10 に示す。

平均推定誤差、正解率において提案手法が最も良い結果となった。提案手法 (with AIC) は、従来手法 (Conventional) や AIC を用いない手法 (without AIC) と比べると、有意水準 5% で有意差が見られた。よって、本提案手法は視線水平方向の推定に有効であると考えられる。また、従来手法に比べ、顔、視線方向を同時に推定することにより正解率が 17.8%, AIC による特徴量選択を用いたことで正解率が更に 5.9% 改善していることが確認できた。平均推定誤差に関しても、顔、視線方向同時推定による推定誤差の改善が 0.16(deg), AIC による特徴量選択を用いたことによる推定誤差の改善が更に 1.21(deg) であった。

ただし、AIC により特徴量選択を行う手法 (with AIC) と MDL により特徴量選択を行う手法 (with MDL) の間には、平均推定誤差、正解率において、どちらも有意差が見られなかった。これにより、AIC と MDL、有意水準 5% でどちらのモデル選択の規準を用いても精度に差がないと考えられる。またこれらのことより、過学習を緩和するため、これらの手法を採用することは視線推定において有効であると考えられる。

表 2 実験結果

手法	平均推定誤差 (deg)	正解率 (%)
Conventional	4.17	40.7
without AIC	4.01	58.5
with MDL	2.88	63.7
with AIC (Proposed)	2.80	64.4

次に、前述の実験結果に関して被験者ごとに検証を行った。

図 11、表 3 は、被験者ごとの平均推定誤差の結果を示している。また、図 12、表 4 は、被験者ごとの正解率の結果を示している。

被験者 m02, m04 において全体的に精度が低くなっているのは、両被験者がメガネを着用しているためと考え

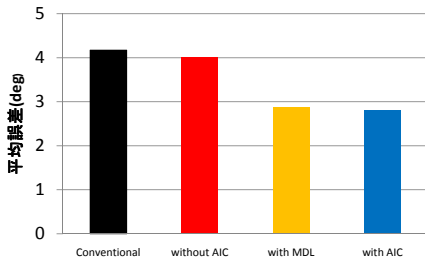


図 9 実験結果 (平均推定誤差)

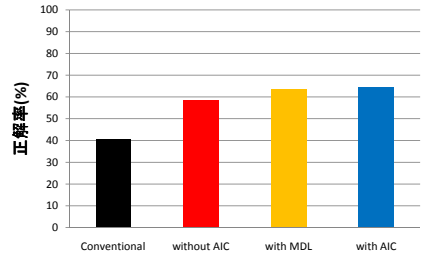


図 10 実験結果 (正解率)

られる。またメガネをかけていない被験者 m01, m03 に関しては従来手法に比べ、提案手法では平均誤差、正解率が共に飛躍的に精度が向上していることが確認できる。

また AIC を用いた手法 (with AIC) と MDL を用いた手法 (with MDL) を比較すると、メガネを着用している被験者に対してわずかに正解率の向上が見られたが、有意差は確認できなかった。よって、AIC, MDL どちらの規準を用いてモデル選択を行って視線推定をしても差がないと考えられる。

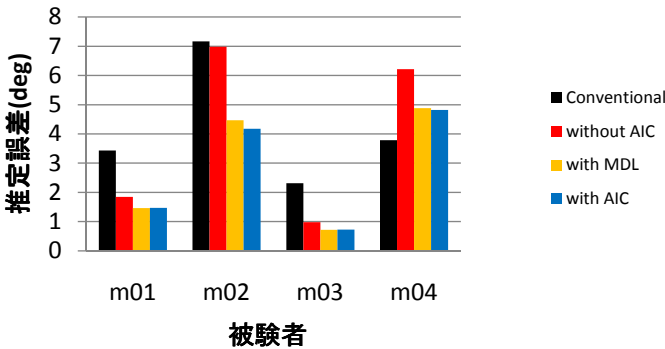


図 11 被験者ごとの実験結果 (平均推定誤差)

7. 結論

本研究では、Active Appearance Models によって、特徴量を抽出し、赤池情報量規準 (AIC) を用いて特徴量を選択し、回帰分析によって顔、視線方向を同時推定することで、従来手法に比べより高精度な視線水平方向の推定が可能となった。また、特徴量を選択する規準として

表 3 被験者ごとの実験結果 (平均推定誤差)
(deg)

手法	m01	m02	m03	m04
Conventional	3.43	7.16	2.31	3.79
without AIC	1.85	6.98	0.98	6.21
with MDL	1.47	4.47	0.72	4.88
with AIC (Proposed)	1.47	4.18	0.73	4.82

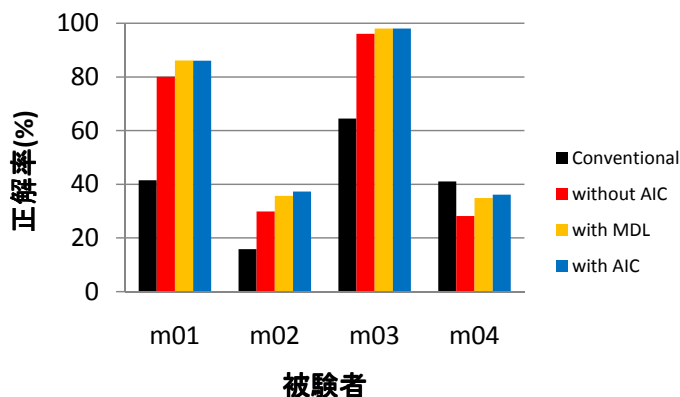


図 12 被験者ごとの実験結果 (正解率)

表 4 被験者ごとの実験結果 (正解率)

手法	m01	m02	m03	m04
Conventional	41.5	15.8	64.5	41.1
without AIC	80.1	29.9	96.0	28.2
with MDL	86.1	35.7	98.0	34.9
with AIC (Proposed)	86.1	37.3	98.0	36.1

用いる AIC と、最小記述長規準 (MDL) の間には大きな差が見られなかった。特徴量の選択にはどちらの規準を用いても問題ないと考えられる。

しかし、眼鏡着用者に関して精度が低い点、AAM が学習人物と同一人物でなければ高精度に特徴量抽出できない点などの問題を抱えているため、今後は特徴量抽出の手法について改善を行う余地があると考えられる。

文 献

- [1] Mehrabian Albert, "Silent Message: Implicit Communication of Emotions and Attitudes," Wadsworth Pub Co, 1980
- [2] Tomoko Yonezawa, Hirotake Yamazoe, "Gaze-communicative behavior of stuffed-toy robot with joint attention and eye contact based on ambient gaze tracking," Proceedings of the 9th international conference on Multimodal interfaces, pp.140-145, 2007
- [3] Derec Pang, Akisato Kimura, Tatsuto Takeuchi, Junji Yamamoto, Kunio Kashino, "A STOCHASTIC MODEL OF SELECTIVE VISUAL ATTENTION WITH A DYNAMIC BAYESIAN NETWORK," Proc, International Conference on Multimedia and Expo, pp.1073-1076, June, 2008
- [4] 白井了, 井東通昌, 乗松有, 足立和正, 中野倫明, 山

本新, "ドライバの運転状態検知のための視線方向の検出", 電子情報通信学会技術研究報告 H I P ヒューマン情報処理, Vol.103, No.455, pp.67-72

- [5] 北川洋介, 加藤丈和, 呉海元, 和田俊和, "アイモデルを用いた視線推定のための黒目追跡," MIRU, 2005
- [6] Takehiko Ohno, Naoki Mukawa, Atsushi Yoshikawa, "FreeGaze: a gaze tracking system for everyday gaze interaction", Proceedings of the 2002 symposium on Eye tracking research & applications, pp.125-132, 2002
- [7] C.H.Morimoto, "Eye gaze tracking techniques for interactive applications," Computer Vision and Image Understanding, Vol.98, Number.1, pp.4-24, 2005
- [8] 三宅哲夫, 今村孝, 章忠, "単眼カメラ画像を用いた視線計測", MSRC, 2007
- [9] 小野泰弘, 岡部孝弘, 佐藤洋一, "低解像度画像からの視線方向推定", MIRU 2006
- [10] Hirotake Yamazoe, Akira Utsumi, Tomoko Yonezawa, Shinji Abe, "Remote gaze estimation with a single camera based on facial-feature tracking without special calibration actions," Proceedings of the 2008 symposium on Eye tracking research and applications, pp.245-250, 2008
- [11] Takahiro Ishikawa, Simon Baker, Iain Matthews, and Takeo Kanade, "Passive Driver Gaze Tracking with Active Appearance Models," CMU, 2004
- [12] T.F.Cootes, G.J.Edwards, C.J.Taylor, "Active Appearance Models", European Conference on Computer Vision 1998
- [13] Akaike H, "A new look at the statistical model identification" Automatic Control, IEEE Transactions Vol.19 p.716-p.723, 1974
- [14] Paul Viola, Michael Jones, "Rapid Object Detection using a Boosted Cascade of Simple Feature", Fourth IEEE International Conference pp.40-45 (2000)
- [15] 林伸治, 長谷川修, "低解像度画像からの顔検出", 画像電子学会誌, 第 34 巻, 2005
- [16] Takahiro Ishikawa, Simon Baker, Iain Matthews, Takeo Kanade, "Passive Driver Gaze Tracking with Active Appearance Models", CMU, 2004
- [17] T.F.Cootes, K Walker, C.J.Taylor, "view-based Active Appearance Models", Automatic Face and Gesture Recognition, 2000
- [18] Peter Huisman, Ruud van Munster, Stephanie Moro-Ellenberger, Raymond Veldhuis, Asker Bazan, "Making 2D face recognition more robust using AAMs for pose compensation", Automatic Face and Gesture Recognition, Vol.10-12, April, 2006
- [19] Hiroshi YAMAGISHI, Kenji AMAYA, "EFFICIENT BOUNDARY ELEMENT METHOD BY MULTI-POLE EXPANSION AND GENERALIZED INVERSE MATRIX FOR ANALYSING TARGET REGION", 計算数理工学論文集, Vol8, No.03-081128
- [20] J. Rissanen, "An Introduction to the MDL Principle", Available online at www.mdl-research.org
- [21] 下平英寿, 伊藤秀一, 久保川達也, 竹内啓, "統計科学のフロンティア 3 モデル選択 予測・検定・推定の交差点", 岩波書店, pp.3-26, 2005
- [22] 小林亮博, 佐竹純次, 平山高嗣, 川島宏彰, 松山隆司, "AAM の動的選択に基づく不特定人物の顔追跡", 情報処理学会研究報告 (CVIM-161), pp.35-40, 2008
- [23] Yusuke SAKASHITA, Hironobu FUJIYOSHI, Yutaka HIRATA "Measurement of 3D Eye Movements Using Image Processing", 実験力学, Vol.6, No.3, September, 2006