

## AAMの動的選択に基づく不特定人物の顔追跡

小林 亮博<sup>†</sup> 佐竹 純二<sup>†</sup> 平山 高嗣<sup>††</sup> 川嶋 宏彰<sup>††</sup> 松山 隆司<sup>††</sup>

<sup>†</sup> 情報通信研究機構 〒619-0289 京都府相楽郡精華町光台 3-5

<sup>††</sup> 京都大学 〒606-8501 京都市左京区吉田本町

E-mail: <sup>†</sup>{akihiro-k,j-satake}@nict.go.jp, <sup>††</sup>{hirayama,kawashima,tm}@i.kyoto-u.ac.jp

**あらまし** 本研究では、情報提示システムの前に訪れた不特定のユーザに対して最適なモデルを選択することで、カメラ画像からリアルタイムで高精度に顔を追跡する方法を提案する。Active Appearance Model (AAM) は形状変化と見えの変化の相関を学習する手法で、顔などの非剛体の追跡に適している。しかし不特定人物の顔を追跡する上で学習データを過剰に用いると、個人の特徴を表す軸が失われるだけでなく、多くの局所解が生まれるため追跡精度が低下する。そこで本研究では学習データを、類似した人物が併合されるようにクラスタリングを行い、未知のユーザに類似した顔のクラスタを選択することで追跡精度を向上させる。本稿では、相互部分空間法を用いてクラスタリングを行う手法を提案し、leave-one-out 法で評価した結果について報告する。

**キーワード** Active Appearance Model, 相互部分空間法, 顔追跡

## Person-Independent Face Tracking Based on Dynamic AAM Selection

Akihiro KOBAYASHI<sup>†</sup>, Junji SATAKE<sup>†</sup>, Takatsugu HIRAYAMA<sup>††</sup>, Hiroaki KAWASHIMA<sup>††</sup>, and Takashi MATSUYAMA<sup>††</sup>

<sup>†</sup> National Institute of Information and Communication Technology 3-5 Hikaridai, Seika-cho, Soraku-gun, Kyoto, 619-0289 Japan

<sup>††</sup> Kyoto University Yoshida-Honmachi, Sakyo-ku Kyoto 606-8501, Japan

E-mail: <sup>†</sup>{akihiro-k,j-satake}@nict.go.jp, <sup>††</sup>{hirayama,kawashima,tm}@i.kyoto-u.ac.jp

**Abstract** This paper proposes a real-time high-precision method to track an unknown face in front of an information system by selecting appropriate model to a video image. Active Appearance Model (AAM) can track a non-rigid object such as a face, because AAM learns the correlation between shape and texture. However, when AAM tracks an unknown face, excessive training data increases tracking errors because of truncating axes which shows characteristics of individuals and of additional local minima generated by the data. In order to increase accuracy of tracking unknown face, we build clusters from training data sets and select a model based on the cluster includes similar face with the unknown face. This paper shows a method of clustering and selecting the cluster based on Mutual Subspace Method, and the result of estimation by leave-one-out.

**Key words** Active Appearance Model, Mutual Subspace Method, Face Tracking

### 1. はじめに

本研究では、ユーザの非言語情報（視線や顔きなど）からユーザの興味や反応を認識し、インタラクティブなやり取りを行う対話システムの開発を目指している [1]。図 1 に示される大型ディスプレイとカメラを組み合わせた情報提示システムを構築し、ユーザの注意（顔向き、視線）や動作（顔き、首傾げ、指差しなど）からユーザの興味や反応を自動的に認識することで、ユーザの好みに合わせた旅行プランニングシステムやユーザの興味を先取りして説明するガイダンスシステムなどへの応用を考えている [2]。

以上のようなユーザの非言語情報を利用するシステムを構築



図 1 情報提示装置

Fig. 1 Information Display

する上で、顔追跡の技術はユーザの顔向き検出や視線検出の精度を左右するため非常に重要であり、様々な研究が行なわれている。中でも Active Appearance Model (AAM) [3] は、shape (特徴点の座標値) と grey-level (輝度値) の相関をパラメタとして持つ統計学的なモデルを構築し、変形を伴う物体を高速で安定的に追跡することが可能であるため、表情解析等の分野でも用いられている [4]。しかし不特定人物を追跡する上で、学習データを過剰に用いると、個人の特徴を表す軸が失われるだけでなく、多くの局所解が生まれるため追跡精度が低下する。そこで本研究では学習データを、類似した人物が併合されるようにクラスタリングを行い、未知のユーザに類似した顔のクラスタを選択することで追跡精度を向上させる手法を提案する。

## 2. 関連研究

AAM は他の多くの顔認識の手法と同様に、主成分分析により顔画像を次元圧縮することで得られる部分空間を利用している。顔の部分空間を追跡に利用する手法としては、Condensation と組み合わせ安定性を向上させた疎固有テンプレート追跡 [5] が上げられるが、これは学習済みの人物のみを扱っている。

従来、画像から不特定人物の顔領域を切り出す研究においては、多数の人物から部分空間を学習し空間の表現能力を向上させるアプローチが取られてきた [6], [7]。また、本研究がめざす顔の特徴点の追跡においても、多人数の顔画像から目、鼻、口といった部品ごとに固有空間を構築することで、画像から顔部品を検出する試みが行なわれている [8], [9]。しかし、過剰なデータから部分空間を学習すると、個人に特徴的な軸が切り捨てられ表現できなくなり追跡精度が落ちてしまうといった問題がある [10]。

本研究ではこの問題を解決するため、未知のユーザに類似した顔のクラスタを選択することで追跡精度を向上させる。固有空間を用いない手法にも目を広げると、個人差を吸収するため多数のモデルを切り替えるというアイデアは、菅野らの手法 [11] や Elastic Bunch Graph Matching (EBGM) [12] で用いられており、高精度な顔特徴点の検出を実現している。菅野らはバンドル調整により個人差パラメタの最適化を行い顔形状の個人差を吸収している。また、EBGM では個人性、照明、肌の色、向きなど様々な見え毎にガボール特徴量モデルを持ち入力画像に適したものを選択しているが、本研究では、部分空間に適切な表現能力を持たせるため複数の人物の学習データからクラスタを生成している。

## 3. AAM [3] を用いた追跡

### 3.1 モデルの学習

AAM では shape と grey-level の間の関係を学習するため、先ず画像から shape 依存の変化を取り除く。学習データとしては特徴点を配置した画像を用意する。図 2 は図 1 の装置で撮影した画像に 45 の特徴点を配した学習データの例である。以下、画像上の特徴点座標を並べたベクトルを shape ベクトル  $s$  と呼ぶ。AAM では、学習セットの shape ベクトル群に対し、スケール・回転・平行移動に対して正規化を行った後、平



図 2 AAM の学習データ  
Fig. 2 Training Data for AAM

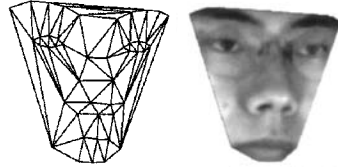


図 3 Warp に用いるメッシュと Warp 後の画像  
Fig. 3 A Mesh for Warp and Warped Image

均 shape  $\bar{s}$  を求める。また、学習画像から特徴点座標に沿って顔領域を切り出し、形状の正規化を行い、切り出した画像を平均形状形に変形する。AAM ではこの操作を Warp と呼び、画像の変形には piecewise affine や thin plate spline が用いられている [13], [14]。本研究では特徴点を結ぶメッシュを構成し piecewise affine によって Warp する。Warp に用いたメッシュと Warp 後の画像を図 3 に示す。以下、Warp した画像の輝度値をベクトルとして表したものを grey-level ベクトル  $g$  と呼ぶ。

次に、shape と grey-level の分布と相関を求める。shape ベクトル  $s$  と grey-level ベクトル  $g$  の系列をそれぞれ主成分分析すると、各学習データの  $s$ ,  $g$  は次式で近似される。

$$s = \bar{s} + U_s c_s, \quad g = \bar{g} + U_g c_g \quad (1)$$

ここで、 $\bar{s}$ ,  $\bar{g}$  はそれぞれ shape と grey-level の平均ベクトル、 $U_s$ ,  $U_g$  は主成分を列ベクトルとして並べた列直交行列、 $c_s$ ,  $c_g$  は主成分の係数である。shape の変化と grey-level の変化には相関があるので、 $c_s$ ,  $c_g$  を連結させた  $c$  を定義し、この  $c$  で表されるベクトル系列に対して主成分分析をさらに行うことで次式のモデルが得られる。

$$\begin{bmatrix} W_s c_s \\ c_g \end{bmatrix} = c = \begin{bmatrix} V_s \\ V_g \end{bmatrix} d = V d \quad (2)$$

ここで、 $W_s$  は shape と grey-level 間の単位の違いを吸収する対角行列である。また、 $V$  は固有ベクトル系列、 $d$  は shape と grey-level の両方を制御するパラメタベクトルである。モデルの線形性に注目すれば、shape ベクトル  $s$  と grey-level ベクトル  $g$  は、 $d$  の関数として次式で表せる。

$$s = \bar{s} + U_s W_s^{-1} V_s d, \quad g = \bar{g} + U_g V_g d \quad (3)$$

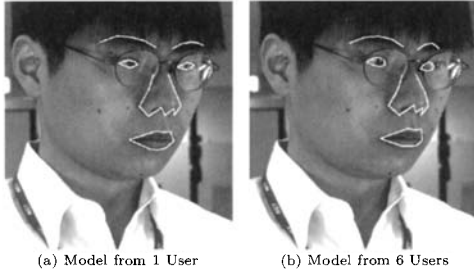


図 4 本人/多人数の AAM による追跡結果

Fig. 4 Tracking Result Based on Personal / Multi-Users' AAM

ここで、 $V = (V_s^T, V_g^T)^T$  である。つまり、 $d$  を与えると式 (3) より画像の grey-level ベクトル  $g$  と shape ベクトル  $s$  が求まる。そして、この  $g$  と  $s$  を用いて画像を合成することができる。Cootes らは、モデルと学習データ画像の間の grey-level の残差ベクトルとモデルパラメタの修正ベクトルの関係は線形であると仮定して、二者の関係を学習している [3]。

### 3.2 AAM における固有空間の探索

追跡対象画像と上述のモデルが与えられた時、固有空間の探索は追跡対象画像とモデルから合成された画像の間の grey-level の残差を最小化する最適化問題として考えることができる。つまり、マッチングによって得られるモデルパラメタ  $d^*$  は

$$d^* = \arg \min_d |g_u - g_v| \quad (4)$$

により表される。 $g_u$  は探索対象画像の grey-level ベクトルであり、 $g_v$  はモデルパラメタから合成された画像の grey-level ベクトルである。得られた最適解  $d^*$  と式 (3) より shape ベクトル  $s^*$  が求められる。

### 3.3 AAM の問題点

AAM では 3.1 で述べたように、grey-level の残差ベクトルとモデルパラメタの修正ベクトルは線形であるといった単純な仮定を用いて式 (4) を最適化する。したがって、学習データが増加すると、個人に特徴的な shape や appearance が切り捨てられるとともに、局所最適解が増えて追跡精度が低下する。図 4(a) は本人のデータから学習した AAM を用いて追跡した結果であり、図 4(b) は本人の他に 5 人分のデータを加えて学習した AAM を用いた結果である。図 4(a) に比べ図 4(b) では、局所解に落ち込んだため目もとに大きなずれが発生しており追跡精度が低下している。この問題に対し本研究では、類似する顔データ毎にクラスタを生成することで、追跡に適した学習データセットを構築し、AAM の追跡精度向上を目指す。

## 4. クラスタの生成と判別

### 4.1 クラスタ間距離の定義

顔画像のクラスタリングや入力画像系列に対するクラスタの選択を行う上で、クラスタ間の類似度を定義する必要がある。本研究では相互部分空間法 [15] で定義された類似度を用いる。相互部分空間法は、従来の静止画ベースの認識法に比べ、入力

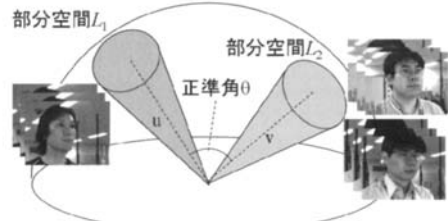


図 5 相互部分空間法

Fig. 5 Mutual Subspace Method

画像全体の分布を利用して判断することで、顔の向きや表情などの変動を吸収することが可能であり、動画像を用いた顔認識に応用されている [16]。図 5 は相互部分空間法の概念図である。相互部分空間法は部分空間  $\mathcal{L}_1$  と部分空間  $\mathcal{L}_2$  のなす最小角度  $\theta_1$  に基づいて空間同士の距離を定義している。最小角  $\theta_1$  は次式で定義される。

$$\cos^2 \theta_1 = \max_{u \in \mathcal{L}_1, v \in \mathcal{L}_2, \|u\| \neq 0, \|v\| \neq 0} \frac{|(u, v)|^2}{\|u\|^2 \|v\|^2} \quad (5)$$

二つの部分空間のなす角度は正準角と呼ばれ、 $M$  次元部分空間  $\mathcal{L}_1$  と  $N$  次元部分空間  $\mathcal{L}_2$  の間には  $N (\leq M)$  個の正準角が定義できる。 $F$  次元特徴空間における部分空間  $\mathcal{L}_1$  と  $\mathcal{L}_2$  の基底ベクトルを  $\Phi_i, \Psi_i$ 、これらから計算される  $F \times F$  の射影行列を  $\mathbf{P}_1 = \sum_{i=1}^M \Phi_i \Phi_i^T$  と  $\mathbf{P}_2 = \sum_{i=1}^N \Psi_i \Psi_i^T$  とすると、 $\mathbf{P}_1 \mathbf{P}_2$  あるいは  $\mathbf{P}_2 \mathbf{P}_1$  の  $i$  番目に大きい固有値  $\lambda_i$  が  $\cos^2 \theta_i$  となる。前田らは、最小正準角  $\theta_1$  に関する  $\cos^2 \theta_1$  が射影行列の積  $\mathbf{P}_1 \mathbf{P}_2 \mathbf{P}_1$  あるいは  $\mathbf{P}_2 \mathbf{P}_1 \mathbf{P}_2$  の最大固有値となることを利用し、最小正準角の高速な導出方法を提案している [15]。

### 4.2 処理の概要

本研究では、特徴点に沿って顔領域を切り出した gray-level ベクトル  $g$  から部分空間を構成しクラスタリングに用いる。3.1 で述べたように、ベクトル  $g$  は平均 Shape に正規化されているため、表情変化といった個人内の変動が抑えられ顔の個人差を比較しやすい。図 6 は、本手法の概要を示す。図中の A-C は顔画像 DB 上のクラスタを示す。(a)-(d) は以下通りである。

- (a) 全クラスタから学習した AAM (初期追跡用)
- (b) 各クラスタから学習した AAM (追跡用)
- (c) 各クラスタの部分空間 (クラスタ選択用)
- (d) 入力画像の部分空間 (クラスタ選択用)

本研究では、(a)-(c) はあらかじめ学習し、(d) はカメラ入力に応じてオンラインで生成するシステムを想定している。(a) と (b) は特徴点追跡に、(c) と (d) はクラスタ選択に用いる。先ずシステムはカメラ入力に対して (a) を用いて初期追跡を行う。得られた特徴点に沿って顔領域を切り出し、全学習データの平均 shape に Warp したベクトル  $g$  を求め、それらを KL 展開して得られた空間を (d) とする。次に各クラスタの (c) と (d) の類似度を求め、最も類似度の高い (c) を選択し、選ばれたクラスタから生成した (b) を用いてより精度の高い追跡を行う。

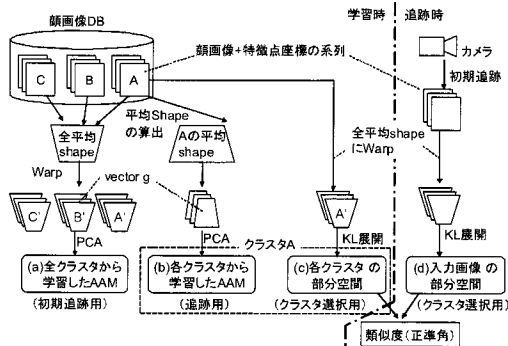


図6 部分空間算出の概要  
Fig. 6 Overview of Building Subspace

AAM どうしを直接比較した方がシンプルなシステムを構築できるが、以下の理由で AAM 間の類似度を求めることは困難である。図 3 に示すように AAM では、モデル毎の平均 Shape  $\bar{s}$  に沿って Warp するため、(b) の画像空間はモデル毎に独立である。しかし、図 5 に示すように相互部分空間法で比較する  $L_1$  と  $L_2$  は一つの画像空間に含まれることを前提としているため、AAM で学習した部分空間をそのまま比較することができない。(b) を全平均 Shape にそそえて構築することも可能であるが、追跡の際は個々のクラスタ毎に特化した shape に Warp した方が精度が向上すると考えられるため、今回は (b) と (c) を別個に生成した。両者を統合して用いる方法については、6.3 で議論する。

### 4.3 クラスタの生成

本研究では、先ず個人毎に部分空間を生成する。得られた空間に対して以下のアルゴリズムを適用し、階層的にクラスタリングを行う。以下、個人部分空間とは最初に計算する個人ごとの部分空間を指す。

**step 0.** 個人部分空間どうしの距離は  $1 - \cos^2 \theta_1$  とする ( $\theta_1$  はそれらの空間の最小正準角)

**step 1.** クラスタ間の距離を計算する (1 回目は全組合せの距離を計算。2 回目以降は、step 2 で併合されたクラスタに関与する距離だけ再計算)

**step 2.** 最も距離の近いクラスタを併合し、クラスタ間距離を再計算する。

- (1) クラスタの部分空間を再計算し、新たな部分空間と他の部分空間との正準角から距離を計算する。
- (2) 個人部分空間は残しておき、クラスタ間距離は、各クラスタに含まれる個人部分空間に基づいて定義する。

**step 3.** クラスタ数がひとつになるまで、step1,2 を繰り返す。

step 2. におけるクラスタ間距離の再計算の方法は、大きく分けて (1) 部分空間の再構成、(2) 個人部分空間による距離の定義の 2 種類が考えられる。しかし (1) を用いた場合、併合が

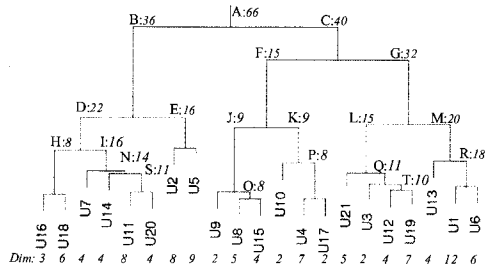


図7 個人部分空間 (被験者 21 名) のクラスタリング結果  
Fig. 7 Result of Clustering 21 Individual-Subspaces

起きたクラスタが平均化されることで他のクラスタとの距離が減少し、また次のターンでも併合されるといった現象が発生し、鎖状のクラスタが形成されてしまうため、今回は各クラスター内の分散が小さくなるよう階層的に併合する Ward 法 [17] を用いてクラスタ間の距離を定義した。

## 5. 実験

### 5.1 実験条件

実験には図 1 に示す情報提示ディスプレイを使用した。50 インチプラズマディスプレイの右、左、下に設置した計 3 台のカメラの内、今回は右のカメラ 1 台のデータのみを用いた。使用したカメラは Point Grey Research 社製の Dragonfly2 (1024 × 768, 30fps, 8bit グレー画像, 1/3 インチ CCD) であり、レンズは FUJINON 社製の HF12.5HA-1B ( $f=12.5\text{mm}$ ) を用いた。また、今回は照明条件を一定にするため、ディスプレイの両脇に写真電気工業社製のライト RIFA-F (50 × 50cm) を配置した。

被験者として 20 代から 40 代の男性 14 名、女性 7 名、計 21 名のユーザを用意し、それぞれ、ディスプレイから 1m はなれた場所からディスプレイ上に表示した 20cm 間隔のマーカ 15 点 (横 3 × 縦 5) を注視する様子を撮影した。また、ユーザには意識して顔をマーカの方に向けるよう指示し、眼鏡を掛けているユーザには眼鏡を外すよう指示した。

撮影した動画から以下の手順で顔データのクラスタを生成した。各マーカを注視している画像をそれぞれ 1frame 合計 15frame 取り出し、図 2 に示す 45 の特徴点を手動で記録した。以上のデータに対して 4.2 で述べた手順で、個人毎に部分空間を構成した。このとき累積寄与率 95% 以上となるよう次元数を決定した。4.3 に述べた手順で統計解析ソフト R [18] を用いてクラスタリングを行なった。その結果を図 7 に示す。Uxx はユーザ番号、A-T は便宜上定義したクラスタ ID を示し、斜体の数字は各クラスタの次元数を示す。

### 5.2 追跡精度の評価

得られたクラスタに対し、leave-one-out 交差検定法により追跡精度を評価した。1 人分のデータをテストセットとしてクラスタから抜いた上で、図 6 の (b) 各クラスタから学習した



図8 クラスタ $\alpha$ / $\beta$ から生成したAAMによる追跡結果  
Fig.8 Tracking Result Based on AAM from Cluster  $\alpha$  /  $\beta$

AAMを構築した。そのうち第1層(B,C)および次の第2層(D,E,F,G)の(b)を用いて、テストセットの動画に対し追跡を行った。例えば図8(a)は、クラスタCからユーザ12のデータを抜いて学習したAAMを用いて、ユーザ12の顔を追跡した結果である。以降、leave-one-outを行なう前にテストセットを含有していたクラスタをクラスタ $\alpha$ と呼ぶ<sup>(注1)</sup>。また、階層的分類を行う際にクラスタ $\alpha$ の比較対象となるクラスタをクラスタ $\beta$ と呼ぶ<sup>(注2)</sup>。図8(b)は、クラスタCの比較対象にあたるクラスタBのAAMを用いてユーザ12の追跡した結果である。

次に、学習データを作成する際に手動で与えた特徴点座標を正解として、各クラスタの追跡性能を評価した。表1は、各クラスタのAAMを用いてテストセットを追跡した際に得られた特徴点座標と正解値との平均誤差を示す。平均誤差は15frameの学習データにおける45点の全特徴点それぞれの誤差を平均して求めた。表1中のクラスタ $\alpha$ の誤差がクラスタ $\beta$ の誤差より大きくなるケースに色をつけた。色をつけたケースを除いて教え上げると、第1層においては21ケース中16ケース(76.2%)、第2層においては21ケース中17ケース(81.0%)において、クラスタ $\alpha$ の方が誤差が小さくなっており、もともと含まれていたクラスタの方が誤差が小さくなっている。すなわち、類似した人物から生成したクラスタを選択することで追跡時の誤差が小さくなることを示し、4.3で生成したクラスタの妥当性を示す。

### 5.3 クラスタ選択性能の評価

未知の人物に対し前節で述べた最適なクラスタを選択できるかどうかテストした。前節と同様にleave-one-out交差検定法を用いて上記の性能を評価した。1人分のデータをテストセットとしクラスタから抜いた上で、図6の(c)クラスタ部分空間を生成する。テストセットから(d)入力画像の部分空間を生成し、各クラスタとの類似度を評価した。実験は、第1層(B-C)および第2層(D-G)に対して、それぞれのユーザについて合計21回行なった。また、正準角の算出においては、部分空間の累積寄与率が95%以上になるよう基底ベクトルを用いる場合と、比較する部分空間の次元数が一定になるように上位5軸の基底ベ

(注1): 例えばユーザ12がテストセットの場合CとGをクラスタ $\alpha$ と呼ぶ  
(注2): 例えばDに対してはEをクラスタ $\beta$ と呼ぶ

表1 クラスタ $\alpha$ と $\beta$ の追跡誤差の比較

Table 1 Difference of Tracking Error Using Cluster  $\alpha$  or  $\beta$

test set	第1層		第2層	
	$\alpha$	$\beta$	$\alpha$	$\beta$
1	8.14	12.08	6.78	12.86
2	8.87	10.99		
3			5.77	5.92
4			6.93	7.10
5	9.20	18.88	10.16	16.27
6	9.07	9.94	7.21	12.29
7	5.82	7.98	6.11	9.03
8	6.16	7.60	6.18	9.89
9	6.41	7.89	6.38	7.37
10	7.46	9.61		
11	8.27	11.61	9.30	10.39
12	6.27	9.01	8.28	8.51
13			10.85	12.70
14	5.54	13.71	7.04	9.61
15				
16	4.87	5.47	4.98	7.31
17	4.26	5.72	4.67	5.44
18	7.23	7.89		
19			6.77	8.24
20	6.63	10.67	5.39	8.10
21	4.51	5.71	4.97	5.35

表2 適切なクラスタ(クラスタ $\alpha$ )を選択した割合

Table 2 Rate of Selecting Appropriate Cluster (Cluster  $\alpha$ )

	第1層	第2層
寄与率一定	81.0%	71.4%
次元数一定	100.0%	85.7%

クトルを用いる場合の2通りの方法で計算した。

表2は、適切なクラスタを選択した割合を示す。本実験ではテストセットがもともと含まれていたクラスタ $\alpha$ を選択できた場合、選択成功とみなした。表2にある通り、次元数をそろえた場合は、第1層で100%、第2層で85.7%選択に成功している。失敗した3ケースの内2つはユーザ2と5のケースである。以上のケースは図7に示すように、クラスタEを選択すると成功とみなされる。しかしクラスタEは要素数の2のクラスタであるためleave-one-outで選択することは困難と考えられる。以上のことからクラスタに十分なデータが含まれていれば、本手法を用いることで不特定のユーザに対し、より追跡精度の高いクラスタを選択することができると言える。

一方、寄与率が一定のときに識別に失敗したケースでは、よりデータ数が多いクラスタを選択していた。これはデータ数が大きいクラスタの方が空間の表現能力が向上し、より小さい最小正準角が見つかったものと考えられる。追跡においては、十分な次元数で空間を表現するため、寄与率一定で計算すべきである。また、今後データベースに含まれる顔データの数を増やしていけば、次元数は一定になっていくと考えられる。ただし、識別プロセスにおいては、データ数が十分でないクラスタの次元数に合わせて評価していく必要がある。

## 6. 議論

### 6.1 最適なクラスタサイズの検討

追跡にどの階層のクラスタを用いるか検討するため、ユーザ12に対する各階層(クラスタA,C,G,L,Q,T)の追跡誤差を、



表 3 各クラスタ階層における追跡誤差  
Table 3 Tracking Error of Each Level Cluster

cluster	A	C	G	L	Q	T
誤差	6.41	6.27	8.28	8.43	11.75	16.44

5.2と同様に leave-one-out 法で求め表 3 にまとめた。表 3 によると、クラスタ C-T ではクラスタサイズが大きくなるにつれ精度が向上しているが、クラスタ A はクラスタ C と比べ誤差が大きくなっていくことが分かる。以上のケースでは、不特定人物に対し全データを用いた AAM に比べ、クラスタリングを行なった AAM がより高い精度での追跡に成功しており、不特定人物の追跡における本手法の有効性を示していると言える。

ここで図 7 に示した各クラスタの次元数を比較すると、T から C にかけては子クラスタの次元数が線形に増加しているが、C と A の間では次元数の増加が鈍っており、部分空間の表現能力の上昇よりも局所解の増加による精度低下が大きくなったことが上記の原因と考えられる。ただし、他のユーザのケースも同様のテストを行なったが、B, C の階層が必ずしも最適のクラスタとはならず、A のクラスタの誤差が最小になるケースが多く見られた。これはデータ数が不十分のため、上記で述べた精度の逆転が起こらず最大クラスタの A の誤差が最小となったと考えられる。しかし、データ数を増やしていけば、表 3 に示す精度の逆転は発生し、最適なクラスタサイズは部分空間の次元数によって推定できるものと考えられる。

### 6.2 初期追跡誤差の影響

本手法は相互部分空間法の入力に AAM で切り出した顔画像を用いているため、背景に関係なくユーザの顔のみを切り出すことが可能であるが、図 6 の (a) を用いた初期追跡で失敗した場合、ユーザが不適切なクラスタに分類されてしまう可能性がある。しかし、式 (4) に示すようにテキストベースのマッチングにより追跡を行った上で部分空間に圧縮しているため、ずれなどのノイズには耐性があると考えられる。また、オンラインシステムに実装する際は、精度の低下に合わせてクラスタ階層を変える等の方法でこの問題は解消できるものと考えられる。

### 6.3 AAM と相互部分空間法の統合

4.2 で述べたように、今回は精度低下を防ぐためトラッキング用の部分空間 (AAM) と、ユーザ識別用の部分空間は分けて用いた。クラスタ専用 AAM を構成する際に、各 Shape を全平均 Shape に正規化することで、AAM 同士を直接比較したクラスタリングやクラスタ選択が可能となる。ただし、この場合追跡精度の低下が起こりうるため、Sung らが提案する Active Shape Model の併用等が必要になると考えられる [19]。

### 6.4 顔向きによるクラスタリング

精度を向上させるためより詳細にクラスを分けるというアプローチも考えられる。今回は最小データ単位として 1 ユーザ毎に部分空間を求めたが、顔向きの軸で分類することも考えられる。特にカメラと顔向きの角度が大きい場合、Warp による変形では顔向き変化が十分に吸収できないため、誤差を生じやすい。そこで顔向き毎にモデルを構成することで、より高い追跡精度が得られると考えられる。

## 7. おわりに

本稿は、不特定人物の追跡に際し、相互部分空間法による類似度を基に顔データのクラスタリングを行ない精度の向上を図った。クラスタ毎の追跡精度を leave-one-out 法により評価し、得られたクラスタの妥当性を示した。本手法の識別精度を評価し、不特定のユーザをより精度の高い追跡を行うクラスタに分類できることを示した。また、全データから学習したモデルに対してより高い追跡精度を持つクラスタの事例を示し、本手法の有効性を示すと共に、クラスタ次元数から最適なクラスタサイズを求める方法について検討した。今後は、クラスタ (AAM) の選択手法を改良するとともに、オンラインシステムに組み込んだ上で多数の被験者による評価実験を行なっていく予定である。

謝辞 本研究の一部は文部科学省科学研究費補助金特定領域研究 No.18049046 の補助を受けて行った。

### 文 献

- [1] 佐竹, 小林, 川嶋, 平山, 水口, 小嶋, 松山: “インタラクティブな情報提示システムのための非装着・非拘束な視線推定”, 情処研報 2007-HCI-125, No. 99, pp. 9-16 (2007).
- [2] 水口, 浅野, 佐竹, 小林, 平山, 川嶋, 小嶋, 松山: “Mind probing: システムの積極的な働きかけによる視線パターンからの興味推定”, 情処研報 2007-HCI-125, No. 99, pp. 1-8 (2007).
- [3] T.F.Cootes, G. Edwards and C.J.Taylor.: “Active appearance models”, Proc. ECCV, Vol. 2, pp. 484-498 (1998).
- [4] 平山, 川嶋, 西山, 松山: “表情認識: 顔パーツ間のタイミング構造に基づく表情の記述”, ヒューマンインタフェース学会論文誌, 9, 2, pp. 201-211 (2007).
- [5] 岡田, 武村, 尺長: “確固有テンプレート追跡による実時間顔認識の検討”, 情処研報 2007-CVIM-160, No. 87, pp. 153-160 (2007).
- [6] M.Turk and A.Pentland: “Eigenfaces for recognition”, Journal of Cognitive Neuroscience, 3, 1, pp. 71-86 (1991).
- [7] K. K. Sung and T. Poggio: “Example-based learning for view-based human face detection”, PAMI, Vol. 20, pp. 39-51 (1998).
- [8] A. Pentland, B. Moghaddam and T. Starner: “View-based and modular eigenspaces for face recognition”, Proc. CVPR (1994).
- [9] 福井, 山口: “形状抽出とパターン照合の組合せによる顔特徴点抽出”, 信学論, J80-D2, 8, pp. 2170-2177 (1997).
- [10] R. Gross, I. Matthews and S. Baker: “Generic vs. person specific active appearance models”, Image and Vision Computing, Vol. 23, pp. 1080-1093 (2005).
- [11] 菅野, 佐藤: “顔変形を伴う 3 次元頭部姿勢の単眼推定”, Proc. MIRU, pp. 311-317 (2007).
- [12] L. Wiskott, J. Fellous, N. Krüger and C. Malsburg: “Face recognition by elastic bunch graph matching”, PAMI, Vol. 19, pp. 775-779 (1997).
- [13] S. Baker and I. Matthews: “Lucas-kanade 20 years on: A unifying framework”, IJCV, Vol. 56, pp. 221-255 (2004).
- [14] S. Sclaroff and J. Isidoro: “Active blobs”, Proc. ICCV, pp. 1146-1153 (1998).
- [15] 前田, 渡辺: “局所的構造を導入したパターン・マッチング法”, 信学論, J68-D, 3, pp. 345-352 (1985).
- [16] 福井, 山口: “一般化差分部分空間に基づく制約相互部分空間法”, 信学論, J87-D-II, 8, pp. 1622-1631 (2004).
- [17] 宮本: “クラスター分析入門”, 森北出版 (1999).
- [18] <http://www.r-project.org/>.
- [19] J. Sung, T. Kanade and D. Kim: “A unified gradient-based approach for combining asm into aam”, IJCV, Vol. 75, pp. 297-309 (2007).