

トラジェクトリ特徴を用いた顔認証システムに関する検討

佐々木 大輔[†] 梶川 嘉延[†]

[†] 関西大学システム理工学部 〒564-8680 大阪府吹田市山手町 3-3-35
E-mail: †kaji@kansai-u.ac.jp

あらまし 我々はトラジェクトリ特徴を用いた顔認証システムを提案する。従来の顔認証技術として代表的な独立成分分析や非負行列因子分解では二次元の画像情報を扱っていた。しかし、これらの方法は顔向きや表情変化などの影響を強く受けるため認証精度の限界が懸念される。また認証対象が生体から直接提示されているか否かを認識することも困難となる。一方、我々の提案システムでは、人が発話する際に変動する部位の軌道を表現したトラジェクトリ特徴量を用いる。顔の変動部位である顔器官や表情筋の動きには個人差が生じると考えられるため、認証精度の向上を期待できる。また時系列情報を扱うため、生体検知の実現も期待できる。本稿では口唇領域に限定して、トラジェクトリ特徴を用いた認証実験を行い、その有効性について検証した。

キーワード 顔認証, トラジェクトリ特徴量, Dynamic Time Warping, AdaBoost

A Face Authentication System by Trajectory Feature.

Daisuke SASAKI[†] and Yoshinobu KAJIKAWA[†]

[†] Faculty of Engineering Science, Kanasai University 3-3-35 Yamate-cho, Suita-si, Osaka 564-8680, Japan
E-mail: †kaji@kansai-u.ac.jp

Abstract We propose a new face authentication system using trajectory features. Many conventional face authentication systems such as using Independent Component Analysis and Non-negative Matrix Factorization are generally based on 2-D information (still image). However, the authentication precision may be limited to be strongly influenced by a face direction and an expression change in these systems. Hence, these systems cannot make a sharp discrimination between human and other objects. On the other hands, the proposed system uses trajectory features which express the orbit of variations in the moment of utterance. Therefore, it is promising for person authentication and discrimination between human and other objects. Experimental results demonstrate that the face authentication system using trajectory features is effective in the actual system.

Key words Face authentication, Trajectory features, Dynamic Time Warping, AdaBoost

1. まえがき

近年、情報漏洩事件や不法侵入による窃盗事件等の被害件数の増加に伴い、巷ではセキュリティ強度を高めるとともに利便性を向上させるための仕組みに関するニーズが高まってきている。このような背景を受け、我々は顔認証について研究を行っている。顔認証は完全非接触であるため、認証対象者に与える心理的負担が少ないと考えられる。

顔認証は基底画像に基づく方法と特徴抽出に基づく方法に大別される。前者には独立成分分析 (Independent Component Analysis) [1], 非負行列因子分解 (Non-negative Matrix Factorization) [2] があり、後者には顔器官の特徴や配置に基づいて照合を行う方法であるガボールウェーブレットとグラフマッチングを組合せた手法 [3] がある。しかしこれらの技術は顔画像より抽出した特徴や基底画像に基づくため、顔向きや表情変

化などの影響を受けやすく認証精度の限界が懸念される。また認証対象が生体から直接提示されているか否かを確認するため、温度センサ [4] や複数のカメラ [5] を利用している認証システムもあり、利便性やコストの面でマイナス要因となる。そこで本研究では従来の認証技術に人間が発言する時の顔器官あるいは表情筋の動きを特徴に加えることで、生体検知及び認証精度の向上が期待できると考え、図 1 に示すトラジェクトリ特徴を用いた顔認証システムの構築を目指す。なお、本稿では図 1 に示す 1 対 1 認証部において、口唇領域の時間的変化を特徴として用いた認証方法を提案する。

2. ではシステムの概要について述べ、次に 3. では本稿の中心となる 1 対 1 認証部における提案手法について述べる。最後に 4. において、動画像を用いた認証実験を行うことにより 1 対 1 認証部における提案手法の有効性を確認する。

2. システムの概要

ここでは本システムの概要について述べる。図1に示すように、本システムは顔検出部、1対N認証部、1対1認証部より構成されている。まず顔検出部ではシステムに入力された被験者の発話動画の各フレーム画像に対し、顔検出[6]を行い顔画像を取得する。次に1対N認証部では、あらかじめ登録者の顔画像をICA変換することにより求めた基底画像を用いて顔画像の係数ベクトルを求め、照合によって類似度の高い候補を選出する。最後に1対1認証部では顔画像より顔器官を検出し、特徴パラメータを抽出する。特徴パラメータに基づきトラジェクトリ特徴を求め、1対N認証部で選出した候補の特徴データと照合を行うことで個人認証を行う。

3. 提案手法

ここでは1対1認証部において、口唇領域のトラジェクトリ特徴を用いた認証方法を提案する。図2に提案手法の概要を示す。図2に示すように、はじめに動画より手で切り出した口唇画像より特徴パラメータを抽出する。次にこれらの特徴パラメータに基づきトラジェクトリ特徴を求め、DTW(Dynamic Time Warping)及びAdaBoostを用いた認証方法により個人認証を行う。トラジェクトリ特徴ならびに認証方法の詳細について次項で述べる。

3.1 トラジェクトリ特徴

図3にトラジェクトリ特徴の一例を示す。トラジェクトリ特徴はフレーム毎に計測される特徴パラメータの時間的変化を空間上にプロットし、時間的変化を軌道として表現したものである[7]。トラジェクトリ特徴は読唇の研究分野で開発された特徴記述法であり、単語の各音の対応を視覚的に把握しやすい利点をもつ。しかしトラジェクトリ特徴は撮影時におけるフレーム速度に依存するため、30[fps]の場合においても情報が欠如し認証精度の低下が懸念される。そのため本稿ではB-Spline曲線を用いてトラジェクトリ特徴量に対して線形補間を施している。また図3に示すように、トラジェクトリ特徴量は複数の特徴パラメータを組み合わせたことが可能である。

続いて、本稿で扱うトラジェクトリ特徴として表現する口唇領域の特徴パラメータの種類について述べる。斉藤らの文献[7]では口唇領域及び口内領域の面積とアスペクト比の2パラメータを用いていた。本稿ではこれらの特徴だけでなく、図4に示す口唇領域を近似した楕円曲線より得られる長軸、短軸、離心率ならびに、それぞれの比を加えた13種類の特徴パラメータを扱う。なおここでの離心率とは楕円の中心から焦点までの距離と長軸長との比である。すなわち離心率は1のとき線分となり、0のとき真円となる。またアスペクト比は近似楕円における長短軸長比として算出している。

3.2 認証方法

3.1で述べたトラジェクトリ特徴は時系列ベクトルデータであり、そのベクトル長は各個人の発話時間に依存する。すなわちベクトル長は一定でないため、単純に類似度を算出することはできない。そのため本稿では伸縮性を有するDTW(Dynamic

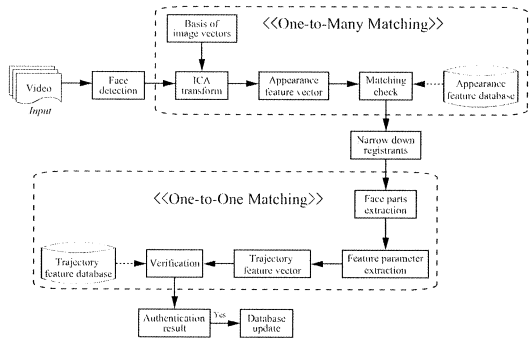


図1 システム概要

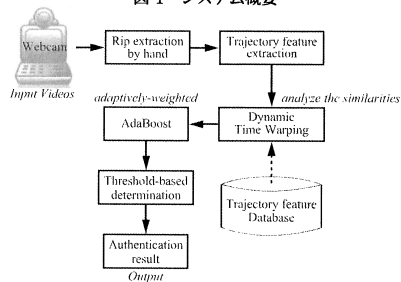


図2 1対1認証部における提案手法

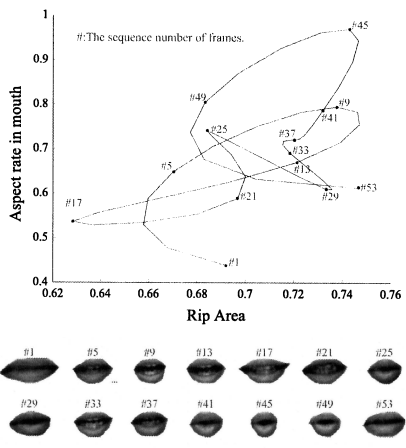


図3 2次元トラジェクトリ特徴の一例

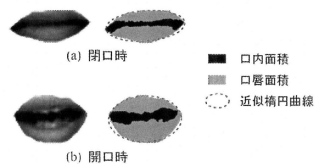


図4 口唇領域の特徴パラメータ

Time Warping)を用いて類似度を算出する。さらに膨大な数のトラジェクトリ特徴を用いるため、その中から有効な種類を選定するのが困難となる。したがって本稿ではアンサンブル学習で代表的なAdaBoostを用いて各トラジェクトリ特徴の重み

付けを行い、算出される尤度に対して閾値判定を行うことで個人認証を行う。

3.2.1 Dynamic Time Warping

DTW は音声認識の分野で開発された手法であり、動的計画法を用いて比較的少ない計算量でベクトル長の異なるデータ間の距離を求める方法である。口唇形状の動きは音声同様、微妙にスピードや軌道は異なる特徴を有するが、音声の周波数情報の 1 次元データであるのに対し、トラジェクトリ特徴は特徴パラメータを組合せる種類の数により次元数が変化する。そのため本稿では、高次元時系列データに適用できるように拡張した DTW を扱う。いま 3 次元時系列データが式 (1) で与えられているとする。

$$\begin{aligned} \mathbf{P} &= \mathbf{p}_1, \dots, \mathbf{p}_i, \dots, \mathbf{p}_M, \quad \mathbf{p}_i = (x_{p_i}, y_{p_i}, z_{p_i})^T \\ \mathbf{Q} &= \mathbf{q}_1, \dots, \mathbf{q}_j, \dots, \mathbf{q}_N, \quad \mathbf{q}_j = (x_{q_j}, y_{q_j}, z_{q_j})^T \end{aligned} \quad (1)$$

このとき、データ \mathbf{P}, \mathbf{Q} 間の距離 $D(\mathbf{P}, \mathbf{Q})$ は式 (2) で与えられる。

$$\begin{aligned} d_{i,j} &= \sqrt{(x_{p_i} - x_{q_j})^2 + (y_{p_i} - y_{q_j})^2 + (z_{p_i} - z_{q_j})^2} \\ L_{i,j} &= \begin{cases} d_{i,j} & (i, j = 1) \\ d_{i,j} + \min(L_{i,j-1}, L_{i,j}, L_{i-1,j}) & (i, j \geq 2) \end{cases} \\ D(\mathbf{P}, \mathbf{Q}) &= L_{M,N} \end{aligned} \quad (2)$$

なお $d_{i,j}$ は \mathbf{P}, \mathbf{Q} の対応する要素間の距離、 $L_{i,j}$ は i, j 要素までを対応づけたときの累積距離を表す。

3.2.2 AdaBoost

AdaBoost は「多数の弱識別機を適当に組み合わせてデータを学習させれば、強力な識別機が構成可能であるか」という問いの一つの解答として提案された。Boosting アルゴリズムの主な特徴は与えられた例題に対する重みの分布を学習の各ステップで適応的に変える事であり、対応したロス関数を逐次最小化することで得られる。いま Boosting に用いる弱識別機を $f(x) \in \{1, -1\}$ 、弱識別機の集合を \mathcal{F} とする。また定義関数を

$$I(R) = \begin{cases} 1 & R \text{ が真} \\ 0 & \text{その他} \end{cases} \quad (3)$$

として AdaBoost のアルゴリズムは以下のように書ける。

(1) 重みの初期値を $w_1(i) = \frac{1}{N}$ 、判別関数を $F_0(x) = 0$ とする。

(2) For $t = 1, \dots, T$

(a) 重みつき誤判定率

$$\epsilon_t(f) = \sum_{i=1}^N w_t(i) I(f(x_i) \neq y_i) \quad (4)$$

を最小にする弱識別機を選択する。

$$f_t = \arg \min_{f \in \mathcal{F}} \epsilon_t(f) \quad (5)$$

(b) 選択した弱識別機に対する信頼係数

$$\alpha_t = \frac{1}{2} \log \frac{1 - \epsilon_t(f_t)}{\epsilon_t(f_t)} \quad (6)$$

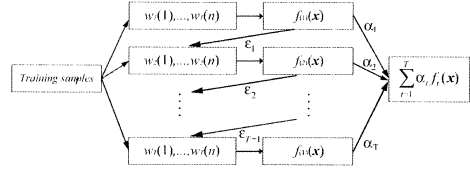


図 5 AdaBoost の概要

を計算し識別機を以下で更新する。

$$F_t = F_{t-1} + \alpha_t f_t \quad (7)$$

(c) 重みを以下のように更新する。

$$w_{t+1}(i) = \frac{\exp(-F_t(x_i)y_i)}{Z_{t+1}} \quad (8)$$

ただし Z_{t+1} は正規化定数であり、

$$Z_{t+1} = \sum_{i=1}^N \exp(-F_t(x_i)y_i) \text{ とする。}$$

(3) 判別関数 $F_T(x) = \sum_{t=1}^T \alpha_t f_t(x)$ の符号で判別する。

4. 個人認証実験

本稿では 3. で述べた提案手法を用いて (1) AdaBoost の弱識別機の個数を変化させた場合、(2) 尤度の判定閾値を変化させた場合、において認証実験を行い、提案手法の有効性について検討を行った。なお本実験では口唇領域の特徴パラメータ (13 種) のうち 3 つを組み合わせた 3 次元トラジェクトリ特徴を用いた。すなわちトラジェクトリ特徴は ${}_{13}C_3 = 286$ 個となる。なお実際のトラジェクトリ特徴は B-Spline 曲線により 0.5 間隔でフレーム補間を行っている。

続いて実験条件を表 1 に示す。以下に前処理、学習ならびに試験フェーズの流れについて述べる。前処理では動画像より口唇領域を手動で抽出しヒストグラム均一化を行った後、判別閾値法を用いて口唇画像ならびに口内画像の 2 値化を行う。学習フェーズでは学習用動画像の全組合せに対し提案手法を用いて、信頼係数ならびに判別関数を学習させ、強識別機を作成する。そして試験フェーズでは試験用動画像の全組合せに対し強識別機を用いて尤度を算出し、判定閾値に基づいて個人認証を行う。また本稿では提案手法の有効性を評価するための指標として誤判定率と誤り率の 2 つを用いる。誤判定率とは学習フェーズにおける誤判定の割合、誤り率とは試験フェーズにおける誤判定の割合である。

4.1 弱識別機の個数を変化させた場合

ここでは判定閾値を 0.5 に固定し、AdaBoost の弱識別機の数を 1 から 286 の範囲で変化させ認証実験を行った。実験結果を図 6 に示す。図 6(a) 及び (b) より弱識別機が 2 個以下では誤判定率が高く、誤り率は低くなっていることがわかる。誤判定率に関しては弱識別機の数に対して学習用組合せが多いため正確に判定できなかったからである。よって学習用組合せの数に適した数の弱識別機を準備する必要がある。誤り率に関しては、このときの誤判定率が高いことから偶発的に良好な結果が得られたと考えられる。

一方、弱識別機が多い場合では誤判定率は低くなり、誤り率

は比較的高くなることもわかる。これは学習フェーズにおいて未学習の組合せが試験用組合せに多く含まれているからである。したがって誤り率を軽減するためには、学習用動画像の登録者数を増やし頑健性を強化するか、またはロバストかつ個人差の生じる特徴パラメータを模索する必要がある。

4.2 判定閾値を変化させた場合

ここでは判定閾値を0.4から0.9まで0.1間隔で変化させ認証実験を行った。なお、各実験でのAdaBoostの弱識別機の数1から55の範囲で変化させた。判定閾値が0.4,0.6,0.8のときの実験結果を図7にそれぞれ示す。図7(i),(ii)において、弱識別機の数が少ない場合に、最も誤り率が低くなっている。これは、このときの誤判定率が高いことから偶発的に良好な結果を得られたと考えられる。また図7(ii)において、弱識別機が20個前後のとき、誤判定率及び誤り率が低くなっている。これは判定閾値を高くしたことで、尤度が0.5前後の未学習の拒否パターンを回避できたからである。

次に図7(iii)において、弱識別機が11個のとき、誤判定率と誤り率が低くなっている。これも上記同様、未学習の拒否パターンを回避できたからである。したがって判定閾値を高く設定することで他人受入率は低くなり、逆に本人拒否率は高くなる。しかし全体的に見てみると、誤判定率は0.76%、誤り率は5.3%が最も低い結果となり、トラジェクトリ特徴が個人認証を行う上で有効であるといえる。

5. むすび

本稿では読唇の研究分野において高認識率が得られているトラジェクトリ特徴を用いた顔認証システムを提案し、1:1認証部における認証方法の有効性を検証するため、認証実験を行った。なお実験は(1)弱識別機の個数を変化させた場合、(2)判定閾値を変化させた場合において行った。これらの結果、最も高い認証精度は(1)で約94%、(2)で約95%となった。これよりトラジェクトリ特徴には個人差が生じ、高い認証精度を有することを確認した。しかし実験結果においてトラジェクトリ特徴は高い認証精度を得られたものの誤認証は生じている。これはトラジェクトリ特徴の基となる特徴パラメータの種類が少ないからであると考えられる。したがって今後の課題として、さらに有効となる特徴パラメータを模索していく必要がある。また本稿では口唇領域に限定したが、今後は他の顔器官の特徴パラメータにおいても検討を行っていく。

文 献

- [1] Marian Stewart Bartlett, Javier R. Movellan and Terrence J. Sejnowski, "Face Recognition by Independent Component Analysis," *IEEE Trans. on Neural Networks*, vol. 13, no. 6, pp. 1450-1463, Nov. 2002.
- [2] D.D. Lee, H.S. Seung, "Algorithms for Non-negative Matrix Factorization," *Neural Inf. Process. Syst.* vol. 13, pp. 556-562, 2001.
- [3] Laurenz Wiskott1z, Jean-Marc Fellous, "Face Recognition by Elastic Bunch Graph Matching," *publ. CRC Press*, chapter 11, pp. 355-396, 1999.
- [4] 安達和隆, 亀山博史, 河田耕三, "顔認証装置および顔認証方法", 特開 2007-11456.
- [5] 早坂昭裕, 柴原琢磨, 伊藤康一, "顔の部分構造を用いた3次

表1 実験条件

被験者	14
学習用動画像	16
学習用組合せ	120
試験用動画像	34
試験用組合せ	561
フレームレート [fps]	30
発話単語の種類	2

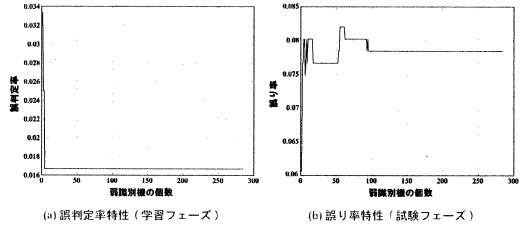


図6 弱識別機の個数を変化させた場合の実験結果

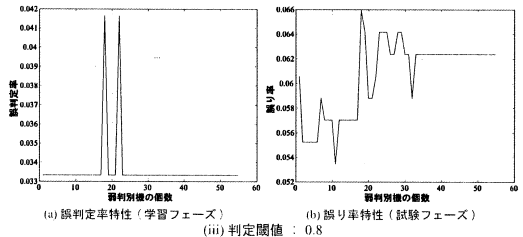
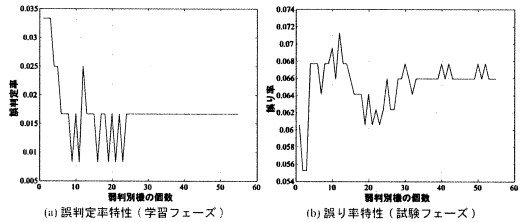
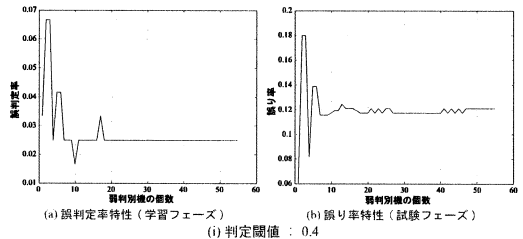


図7 弱識別機の個数を変化させた場合の実験結果

- 元顔認証の検討," 第22回信号処理シンポジウム, P4-4, pp. 741-746, Nov. 2007.
- [6] 佐々木大輔, 梶川嘉延, "モーショント特徴を用いた顔認証システムのための顔位置補正法に関する一検討," 第7回情報科学技術フォーラム, H-036, pp. 139-140, Nov. 2008.
 - [7] 斉藤剛史, 小西亮介, "トラジェクトリ特徴量に基づく単語読唇," 電子情報通信学会論文誌 D, vol. J90-D, no. 4, pp. 1105-1114, Apr. 2007.