

口唇の動きを利用した動画像による顔認証システム

理学専攻・情報科学コース
玉田竜子(指導教員: 金子 晃)

1 はじめに

近年, 生体特徴を用いた生体個人認証技術が注目を集めている. その背景として, ID カードの偽造やパスワード盗難などの犯罪が巧妙になっていることや, 従来の身分証明書やパスワードなどの記憶による認証は紛失や忘却の危険性があることから, より利便性の高い認証方法が求められていることが挙げられる.

本研究では顔認証技術に焦点をあてて検討していく. 顔認証は人間が他人を認識するのに日常的に用いられる自然な方法であり, 非接触型なので人からの抵抗感が少ない. 顔画像は遠くから取得できるため, 登録・認証処理の利便性が高い. このような長所により現在活発に研究が行われている. また, 同様の理由から人物監視システムの要素技術として期待されている. セキュリティ用途以外にも, 顔をキーとした画像検索やユーザーの顔を認識する擬人化エージェントなど, 広範囲で応用が利くのが顔認証の大きな特徴である. 一方顔認証の短所として, 他人の顔写真などを用いて簡単になりすまし攻撃を受けるといことが挙げられる. 例として, 2009年には, 顔認証方式のたばこ自動販売機で, 小学生の男児が成人と誤認証され, たばこを購入した事例も判明している. また, 顔認証の精度を下げる要因として, 一卵性双生児の識別の困難さや, 加齢や傷病, 眼鏡やマスク, 撮影時における照明変動の影響を受けやすいことなどが挙げられる.

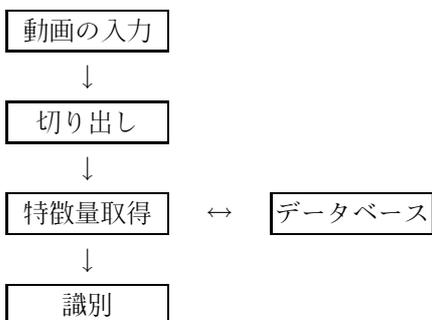
前述のように, 簡単になりすまし攻撃を受ける顔認証だが, 顔は表情変化や開口によりパターンを変化させることができるため, 動的特徴を認証に利用できる.

2 システム概要

本研究では, 口唇の動きを利用して研究室内への入室審査のような少人数の人間を識別するシステムを提案する. 低価格化が進んだ USB 接続の Web カメラを使用したことで, USB ポートのあるパソコンなら簡単に接続でき低価格で認識環境を整えられる.

2.1 システムの流れ

以下のようなシステム流れで識別を行う.



2.2 動画の取得

動画像の取得・加工には OpenCV のライブラリ関数を利用した. 動画の入力には Web カメラを利用し, 登

録者にはカメラに向かって自分の名前を言ってもらった.



2.3 特徴量の取得

入力された動画像を YCrCb 表色系に変換し, OpenCV のライブラリ関数 cvFindContours を利用して輪郭の抽出を行った. その輪郭の口唇の上下左右の4点の動きを追跡し, その差分を特徴量として取得した.

2.3.1 YCrCb 表色系

この表色系は, テレビなどの放送系のメディアで用いられるもので,

Y=色の輝度

Cr=輝度と R 成分の色差

Cb=輝度と B 成分の色差

という3つの成分によって色を表現する方法である. 「輝度」とは人間の視覚的に明るく見える度合いである. 光の三原色の場合赤が一番明るく見え, 青が一番暗く見える. このような視覚的效果を考慮した明るさが輝度である. また, 「色差」とは, 輝度と色成分との差分値. 緑の色差成分は上記の3つの情報があれば求められる.

RGB⇒YCrCb の変換は以下のようになる.

$$Y = 0.2989R + 0.5866G + 0.1145B \quad (1)$$

$$Cb = -0.1687R - 0.3312G + 0.5000B \quad (2)$$

$$Cr = 0.5000R - 0.4183G - 0.0816B \quad (3)$$

さまざまな色変換を試した結果, YCrCb 表色系が最も唇輪郭の抽出に適していたので使用した.

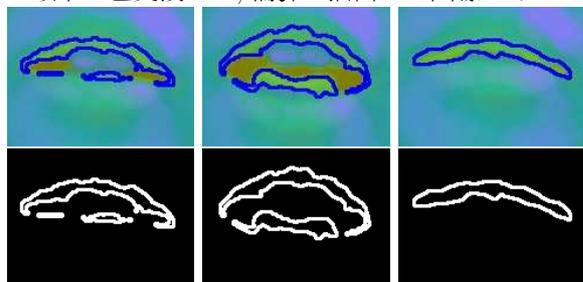
2.3.2 cvFindContours

cvFindContours は OpenCV の輪郭抽出関数である. この関数は次のようなつくりになっている.

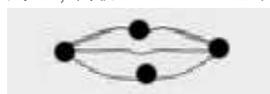
```
int cvFindContours(CvArr* image, CvMemStorage* storage, CvSeq** first_contour, int header_size=sizeof(CvContour), int mode=CV_RETR_LIST, int method=CV_CHAIN_APPROX_SIMPLE, CvPointoffset=cvPoint(0,0));  
image... 輪郭を抽出したい入力画像 (1チャンネルで8ビットの階調値を持つ画像である必要がある)  
storage... 得られた輪郭の情報を保存する場所を指定  
first_contour... 最初の輪郭の保存場所を示すポインタ  
header_size... ヘッダのサイズを定義
```

mode... どのような輪郭をどのようなデータ構造で格納するか
 ・ CV_RETR_EXTERNAL... 一番外側の輪郭だけを抽出したい場合
 ・ CV_RETR_LIST... すべての輪郭を抽出したい場合
 method... 輪郭を近似する手法を定義
 offset... 得られた輪郭データを、定義した値だけずらして保存. 特に指定がない場合は (0,0)

以下が色変換をし、輪郭を抽出した画像である.



このような画像から以下のように4端点を時系列で追跡し、特徴ベクトルを取得した.



2.4 識別手法

システムに登録された人物であるかそうではないかを識別する.

入力されたものの主成分を求め、登録されたデータそれぞれとのクラス毎のユークリッド距離が最小のものと、登録されたデータの重心とのユークリッド距離が最小なものをそれぞれ抽出した.

2.4.1 主成分分析

主成分分析とは、多くの量的変数が存在する場合に、それらの間の相関構造を考慮して、低い次元の合成変数（主成分）に変換し、データが有している情報をより解釈しやすくするための方法である. 主成分分析には「相関係数行列から出発する方法」と「分散共分散行列から出発する方法」の2種類がある. 今回は、前者を利用した.

主成分分析の一般的な解析の流れは以下の様になる.

- (1) 相関係数行列 R の第1固有値（最大固有値） λ_1 に対応する固有ベクトルから第1主成分 z_1 を求める. 次に R の第2固有値 λ_2 に対応する固有ベクトルから第2主成分 z_2 を求める. 同様にして、第 k 主成分を求める ($k = 3, 4, \dots, p$).
- (2) それぞれの主成分の寄与率および累積寄与率を求める. 「固有値が1以上」ないしは「累積寄与率が80%を超える」を目安として主成分を選択する.
- (3) 因子負荷量を求める. 固有ベクトルや因子負荷量の値を参考にして、選択した各成分の意味について考察する. また、因子負荷量を散布図にプロットし、変数の分類を行う.
- (4) 主成分得点を散布図にプロットし、サンプルの特徴付けや分類を行う.

2.4.2 クラスタ間の距離

クラスタ間の距離にはいくつかの方法がある. ここでは、最短距離法と重心法を用いた.

• 最短距離法

2つのクラスタに属する対象のうち、最も近い対象間の距離をクラスタ間の距離とする方法である.

• 重心法

各クラスタの代表点を重心とし、重心間の距離をクラスタ間の距離とする方法である. 重心とは、各変量の平均値の座標にあたる点である.

2.4.3 ユークリッド距離

$(x_{ik})_{k=1}^n, (x_{jk})_{k=1}^n$ をそれぞれ i 番目, j 番目の対象データとすると、ユークリッド距離は、

$$d_{ij} = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2} \quad (4)$$

である. ユークリッド距離は、日常的に用いる距離と同じである.

2.5 実験結果

登録サンプル数は1人につき5個取った.

| 登録者 (人) | テスト回数 (回) | 最短 (%) | 重心 (%) |
|---------|-----------|--------|--------|
| 3 | 16 | 84 | 81 |
| 5 | 20 | 78 | 60 |

なりすまし (他人の名前をかたる) に関しては5回テストをして、詐称できたのは0回であった. 5回中3回は発言者本人とされた.

2.6 まとめと今後の課題

今回の実験では登録者5人と少なかったため、さらにサンプル数を増やし実験を行う必要があるが、現時点での結果では、少人数であれば口の動きだけで個人識別をすることが可能であることが分かった. 最短距離法と比べ重心法の結果があまりよくなかったことから、登録データの中に異常値が入っていたことが考えられる. よって、動画像入力時のノイズの除去をよりよく行う必要がある. なりすましに関しては、テスト回数が少ないが少人数では詐称が不可能である可能性がある. 今後なりすましデータの収集も行い、システムの精度の向上のために OpenCV のライブラリ関数を使い、簡単な顔認識機能を合わせるにより実用的なシステムとしたい.

参考文献

- [1] Federico Matta, "Person recognition using facial video information", Journal of Visual Languages and Computing 20(2009), pp.180-187.
- [2] Maycel-Isaac Faraj, Josef Bigun, "Audio-visual person authentication using lip-motion from orientation maps", Pattern Recognition Letters 28(2007), pp.1368-1382.
- [3] Gray Bradski and Adrian Kaebler, "Learning OpenCV", O'REILLY, September 2008.