

ラベルグラフマッチングを用いた顔画像による個人認証

小野 幸真 村上 真 米山 正秀

東洋大学工学部情報工学科

1 はじめに

人間の顔を機械に認識させることは、人間と機械のコミュニケーションや盗難や偽造の危険性がほとんどないセキュリティへの応用など多くの可能性を秘めている。

ここで、顔に限らず物体の認識を行ううえでは少なくともいくつかの特徴量を得て比較するといったことが必要になる。また、いくつかの特徴点を抽出するためには全体の構造が必要となる。そこで、局所の特徴量と大局の特徴量を同時に用いることが重要となる。本稿で扱うラベルグラフマッチングはこのことを簡潔にモデル化する1つの方法である。本研究の目的は、グラフマッチングと呼ぶこの照合過程において、先ほどの2つの特徴量をどのような配分で重視すれば高い認識率が得られるかを考察することである。

2 グラフマッチング

グラフマッチングに用いるグラフは、図1のようにバーテックスとエッジから構成される。つまり顔画像から抽出する特徴量を、局所的なものと大局的なものの2つに分け sparse なグラフで表現する。照合の際には、あらかじめ用意した教師画像から得たグラフと入力画像から得たグラフの特徴量の比較により行う。

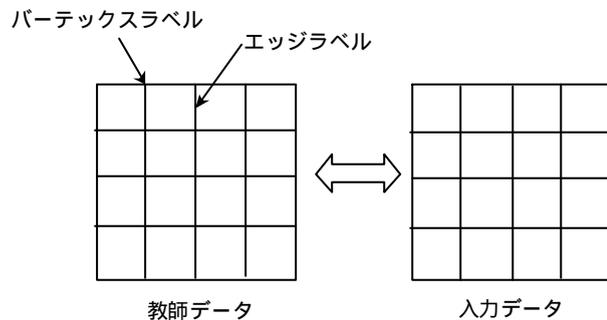


図1 グラフマッチング

2.1 バーテックスラベル

グラフの各頂点には、ガボールウェーブレット変換により得られる特徴ベクトルを局所特徴量として付加する。

ガボールウェーブレット変換 (GWT)

GWT は画像内の顔に拡大、縮小や回転などがある場合にもパラメータ j , q を変化させたガボールウェーブレットにより特徴抽出を行うことができる。

入力画像を I , 2次元ガボール核を y とすると GWT は式(1)のような畳み込み積分で表現される。

GWT の定義

$$G_{j,q}(x_0, y_0) = a^{-j} \iint I(x, y) y_q \left(\frac{x-x_0}{a^j}, \frac{y-y_0}{a^j} \right) dx dy \quad (1)$$

また2次元ガボール核 y は式(3)を用いて式(2)のように定義される。

$$y(x, y) = g_s(x, y) \left[e^{iu_0(x \cos(q) + y \sin(q))} - e^{-(u_0 s)^2} \right] \quad (2)$$

$$g_s(x, y) = \frac{1}{4ps} e^{-\frac{1}{4s^2}(x^2+y^2)} \quad (3)$$

ここで、教師データのある局所における特徴ベクトルを式(4)のように表す。

$$J_{j,q}(x_0, y_0) = \left| G_{j,q}(x_0, y_0) \right| \quad (4)$$

同様に入力データの局所特徴ベクトルを J^I と表すと教師データと入力データの局所特徴ベクトルの類似度は式(5)によって計算される。

$$S_v(J^M, J^I) = \frac{J^M J^I}{|J^M| |J^I|} \quad (5)$$

2.2 エッジラベル

i, j がエッジによって結合されているとするとエッジラベルは式(6)のように定義され、式(7)は教師データと入力データのグラフにおける同一局所間の差である。

$$\bar{\Delta}_{ij} = \bar{x}_j - \bar{x}_i, (i, j) \in E \quad (6)$$

$$S_e(\bar{\Delta}_{ij}^I, \bar{\Delta}_{ij}^M) = (\bar{\Delta}_{ij}^I - \bar{\Delta}_{ij}^M)^2 \quad (7)$$

つまり、式(7)は初期のグラフの構造が崩れるほど大きくなる。

2.3 グラフの照合

教師データと入力データのグラフの照合は、それぞれのグラフのバーテックスラベルとエッジラベルを用いた式(8)を評価することで行う。

$$C_{total} = I C_e + C_v \\ = I \sum_{(i,j) \in E} S_e(\bar{\Delta}_{ij}^M, \bar{\Delta}_{ij}^I) - \sum_{i \in V} S_v(J^M, J^I) \quad (8)$$

最急降下法により、 C_{total} が最小となるように、グラフ全体、もしくは各頂点を移動させる。図2に実際のグラフマッチング終了時の画像を示す。

An analysis of control parameter in labeled graph matching for human face recognition

Yukimasa Ono, Makoto Murakami, Masahide Yoneyama
† Dept. of Information and Computer Sciences, Toyo University

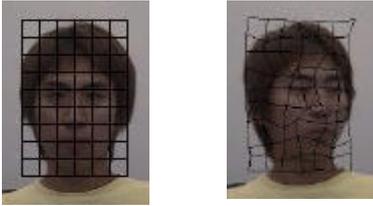


図2 教師画像とマッチング後の入力画像

2.4 パラメータ I の制御法

パラメータ I は認識過程において動的に変更することが可能であるが、一般的には簡単のために2段階に設定することが多い。認識初期においては大きな値を設定することで入力画像中の大域的に類似した部分、つまり顔に当たる部分を抽出し、顔領域の抽出後に小さな値に設定することでグラフを変形させ個人の認証を行う。

しかし、 I を動的に変更することで認識過程の収束速度を速め、領域抽出を速く正確に行うといった制御が考えられる。また、 I の値が大きいと、表情や向きの変化に対応できないが、小さすぎても全体の構造が大きく崩れてしまうために良い認識結果は得られないと考えられる。本論文においては、構造を保ちながら高い認識率が得られる I の最適値について考察する。

3 実験

I の値を変えて認識実験を行い、個人認識における I の最適値の分析を行った。1人の教師画像（正面画像）に対して、

- (i) 表情、姿勢の変化の無い同一人物の正面画像
- () 表情、姿勢の変化のある同一人物の画像
- () 教師画像とは異なる人物の正面画像を用意する。

表情変化を吸収し柔軟な認識を行うためには各教師画像に対して同一人物が入力されたときの C_{total} が他の場合よりも小さくしなければならない。

3.1 実験に用いた画像データ

今回は()に関しては上下左右方向へ回転したもの（回転の程度は3段階の12枚）()に関しては異なる人物2人分の画像を用意して行った。また、教師画像は3人分用意した。



図3 上下左右方向へ回転した顔画像の例

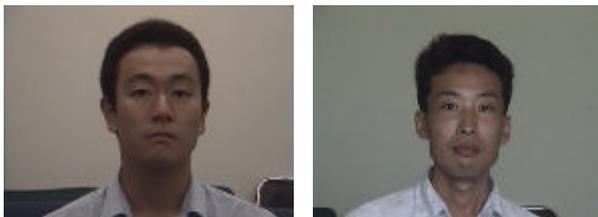
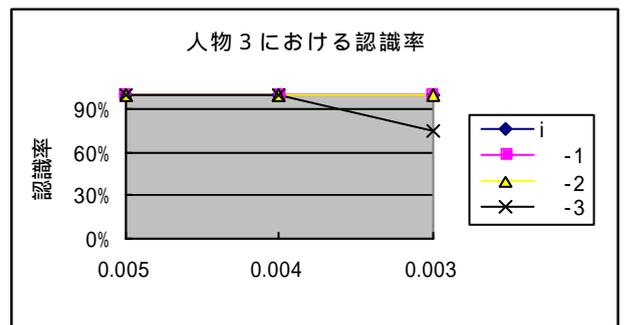
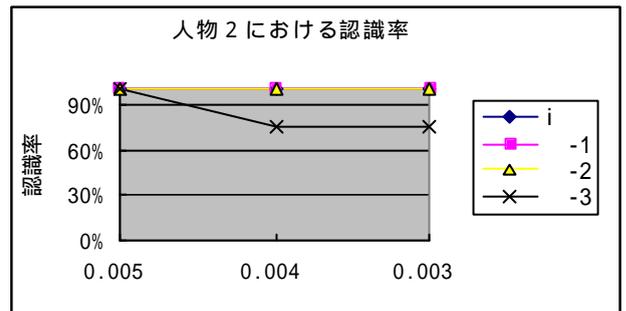
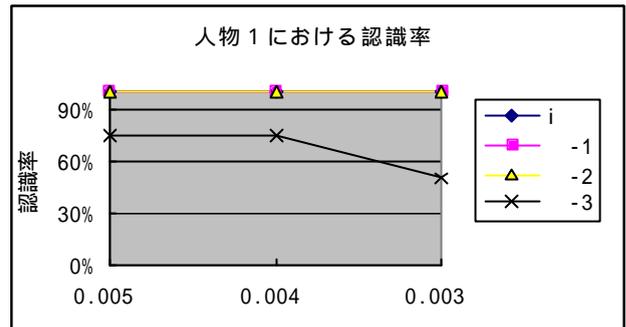


図4 別人として入力した顔画像の例

ここでガボールウェーブレットに与える各パラメータは $q = 0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ$ $j = 3, 4, 5$ $u_0 = 5p$ $s = 0.15$ とした。また使用する画像のサイズは 320×240 ピクセル、初期グラフの各エッジのサイズは20ピクセルとし頂点の数は 10×7 で行った。

3.2 結果

画像() (別人の画像) の C_{total} よりもコストが低くなった画像の割合を認識率として以下に示す。



-1は画像()回転の度合いが1のものを示す

3.3 考察

$I = 0.004$ のときには回転の度合いが大きなものから認識率が落ちている。よってこの条件下では I の値は $0.004 \sim 0.003$ が適当なのではないと思われる。

4 まとめ

今回の実験では入力するデータセットの数が少なかったために、データセットを増やして検証する必要がある。また今後の展望としては I を固定した値ではなく段階的に減少させることで認識率の向上が図れないか検討していくつもりである。