

マルチモーダルインタラクションにおける個人識別について

マルチモーダルインタラクションに関する研究

小川 均

Abstract: マルチモーダルインタラクションにおいては、特別な入力装置が必要とせず、また、非接触で非拘束に個人認証することが望まれている。本研究では顔が傾いているときや顔の一部が障害物で隠れてしまっているときなど識別が困難と思われる画像を複数枚用いた顔の見え方にロバストな個人識別を行った。各画像の識別は登録画像と入力画像の相違をマハラノビス距離で算出し、識別結果による確率を Dempster&Shaferの理論を用いて統合を行った。

1. はじめに

本研究で提案するマルチモーダルインタラクションの目的は、システムを気軽に操作する環境を提供することである。すなわち、使用者からの入力、本人存在を示す顔、身振り、音声、および、キーボードやマウスからの記号が想定される。これらは、直接システムへの命令を示す場合があるが、多くの場合は、要求を示している。たとえば、命令は「画像を 1.2 倍に拡大せよ。」であるが、要求は、「画像をもっと大きくしたい。」である。要求から命令を実現する命令生成部では、要求に加えて個人の嗜好や使用できる環境の制約を考慮して命令を決定する。

一般に、コミュニケーションの方法や表現方法は相手によって多少とも異なる為、本マルチモーダルインタラクションは各個人に適応した対応を目指している。したがって、本システムにおいて最も重要な要素は個人識別である。

顔画像の入力にはカメラ以外の特別な装置は必要なく非接触で非拘束に認識が可能である。本研究では顔を使った個人識別システムを実現することを目指す。

最近の研究動向は商品化されているものも含めて正面顔画像に対してはほぼ認識が可能となっている。しかし顔が傾いていたり、正面を向いていなかったりする顔画像に対しては認識が行えていない。[1][2][3][4]また正面を向いていても顔の一部が隠れていると認識が行えない。そこで本研究では上述の画像に対しても認識が行えるように顔の見え方に対してロバストな個人識別を目指す。[5]

2. 識別方法

2.1 基本技術

本研究で用いる手法について説明する。

(1) 顔部位認識システム

顔画像を用いて個人識別するとき重要なのは、顔領域内で目や鼻、口など顔の構成部品の位置関係や特徴点を正しく把握する事である。本研究では情報処理振興事業学会 (IPA) が開発した、顔部品認識システムを使用する^[6]。

このシステムは1つの静止顔画像を入力として与えると、その画像から顔の輪郭や眉、目、鼻、口などの顔部品の輪郭とその座標を出力する。このシステムの処理の流れは図1の通りである。

部品切り出しを行った例を示す。図2 (a) の入力画像から部品を切り出した結果、図2 (b) に示すように部品の切り出しに成功した。白枠が、眉、目、鼻、口のそれぞれの領域で、点で囲まれているのがそれぞれの輪郭である。また黒い枠、紫の枠、緑の枠はそれぞれ頭部領域、肌色領域、顔領域を表している。

(2) マハラノビス距離

マハラノビス距離とは、パターン認識理論の一つであり、人間の判断をコンピュータに置き換えて、人間がより高度で間違いのない判断をするための有効な情報を提供することに応用されている。本研究では個人識別にこのマハラノビス距離を使用する。

パターン認識理論には、これまで様々な方法が提案されており、多くの問題に適用されている。しかし、今なお未解決の問題も多く存在する。マハラノビス距離はこうした未解決の問題に新しい解を与える可能性を持っている。例えば、従来コンピュータによる実現が困難であった

- ・ 製品の外観検査、パターンの判別、診断、官

能検査などに適用することができると考えられる。こうした広範な応用範囲の背景にあるのは、「均質なデータ群を基準としてそこからの距離を求める」という考えである。

マハラノビス距離は以下の式(1)で算出する。

$$D_M = (x - u^i)^T \Sigma^{-1} (x - u^i) \quad (1)$$

x :未知入力ベクトル u^i :平均ベクトル

Σ :分散共分散行列

本研究では顔画像から目・鼻・口など顔部品を切り出して特徴点を設定し部品間の位置関係を数値化する。それをを用いて登録した人物と入力画像との類似度を、マハラノビス距離を用いて計算してその値により入力的人物が誰であるかの個人識別を行う。

マハラノビス距離には、正常空間を作る基準データと認識の対象となるデータとの異なる度合いが大きくなると、その値が急激に大きくなるという性質がある。したがって、認識対象のマハラノビス距離が小さければ小さいほど、その対象は基準データと近い性質にあり、距離が大きければ遠いことになる。またマハラノビス距離を使った識別が他のパターン情報処理と比較して優れているのは、正常空間という考え方にある。これまでの多くの手法では、正常な状態も非正常な状態も同じ認識空間で処理していた。しかし、正常な状態というのはそのパラメータが比較的均質で締まった状態であるのに対し、非正常な状態はそのパラメータが様々な方向に大きくばらついている。均質性のあるデータ群から得た認識空間と、それが保証されていない認識空間とは信頼性や感度に大きな相違が出るということがいえるので、同一人の登録顔画像のパラメータで作成する正常空間は均質で締まった状態になっていると考えられる。認識対象である未知入力画像は、この正常空間とのマハラノビス距離によって識別されるので精度の良い識別が出来ると考えられる。

(3) Dempster&Shafer理論

本研究では上記のマハラノビス距離で識別を行うが、精度を上げるために識別には1枚ではなく連続して入力された複数枚の入力画像を使用する。その際に画像によるマハラノビス距離の値のバラツキを吸収してなおかつそれらの値を統合し一つの値にするためにDempster&Shafer理論を用いる。

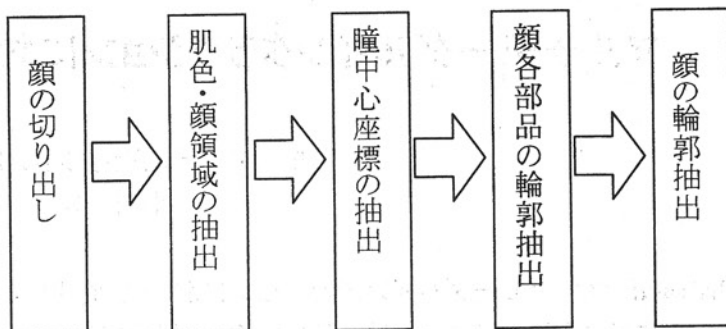


図1 顔部位認識システムの処理の流れ

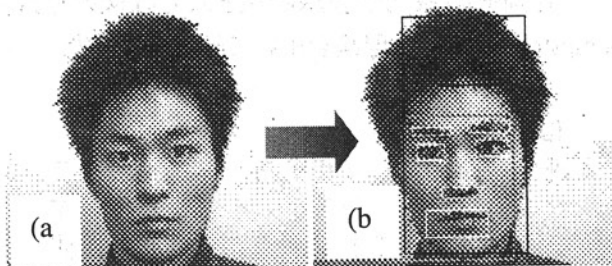


図2 顔部品切り出しの結果

Dempster&Shafer理論では、基本確率と上界確率を用いる。基本確率は $m(x)$ で表され、 $m(x)$ は集合を引数に持つ確率値で、引数 x の集合の要素である事象のどれかが成立する確率を表している。要素のどれが成立するかについては平等な可能性をもっている。事象ごとの成立する確率の違いは基本確率の組み合わせによって表現される。例えば、 $(m(a, b)=0.6, m(a)=0.4)$ の基本確率の組では0.6の確率で事象aか事象bが成立し、0.4の確率で事象aが成立することを表す。このことから事象aの成立可能性の最大値は1.0、事象bのそれは0.6となる。このような基本確率から計算される各々の事象が成立する可能性の最大値を上界確率という。この値はその事象の成立を否定する情報がない限り高い値に留まる性質がある。また複数の確率の値を統合でき補正計算を行い値を算出できるので、本研究ではまず識別結果のマハラノビス距離の数値を以下の変換式(2)を用いて確率に変換する。

$$P_k = \frac{\sum_{j=1}^n MD_j - MD_k}{\sum_{i=1}^n \left(\sum_{j=1}^n MD_j - MD_i \right)} \times (1 - \theta) \quad (2)$$

k :登録人物 n :登録人数

MD :マハラノビス距離の値 θ :不確実量

また不確実量を扱える事からあらかじめ求めておいた識別に使用する部品による信頼度の違いを不確実量の値とする。確率を統合する時にその値により重み付けを行うので信頼の高い顔部品を使用した識別結果のマハラノビス距離の確率が上昇する。

2.2 顔の見え方にロバストな個人識別

実際の識別の流れを示す。

①顔部品切り出し(登録データベース用、図4参照)

最初に、顔画像から顔の部品を切り出す。登録人物一人につき4枚～20枚の登録用の正面顔画像を用意する。IPAの顔部位認識システムを使用し目・鼻・口などの顔部品、顔の輪郭などをその座標とともに出力する^[6]。このとき以下の処理をすることで見え方に対してロバストにする。

- ・ 顔が傾いている画像に対しては画像を回転させて垂直にすることで切り出しの精度を上げる
- ・ 部品が正しく切り出されたかどうかを他の部品の配置との相対関係から判断し正しく切り出せた部品のみを②のパラメータ算出に用いる

②パラメータ算出(図4参照)

①で算出した顔部品の座標を用いて、顔部品の位置関係を数値化しパラメータとする。数値化はまず両瞳間、左(右)瞳-鼻先、左(右)瞳-唇の真中、鼻の両端間などの2点間距離を基準とする。そしてその距離を基準にしてその他の部品間の距離の比率をパラメータとする事で行う。パラメータは複数個(3～15個)算出する。

識別はマハラノビス距離で行うので、その算出に必要である各パラメータの平均と分散共分散行列を求めてデータベース(以下DB)にその値を登録人物ごとに格納しておく。このときあらかじめ考えられるあらゆる部品の位置関係を数値化しDBを作成する。全部品の座標を使用して作成したDB、口以外の部品を使用して作成したDB、鼻以外...というようにする。こうすることで識別の際に入力された画像の顔部品が障害物で隠れている(あるいは切り出しがうまくいかなかった)場合でも、隠れている部品以外を使用して作成したDBとの類似度を求めることができる。これにより決められた部品ではなく、識別に使う部品を入力された画像によって、または切り出しの結果によって変えられて、障害物で目や口など顔の一部が隠れてしまっている画像に対してロバストな識別処理が行えると考えられる。

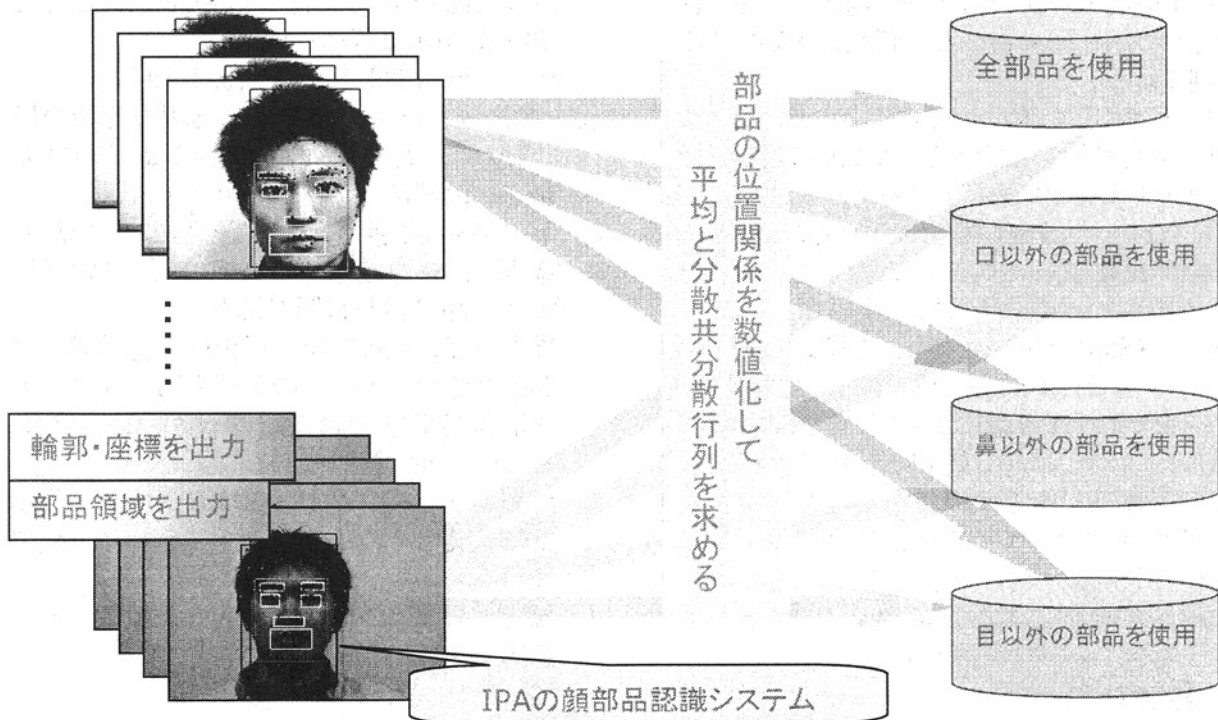
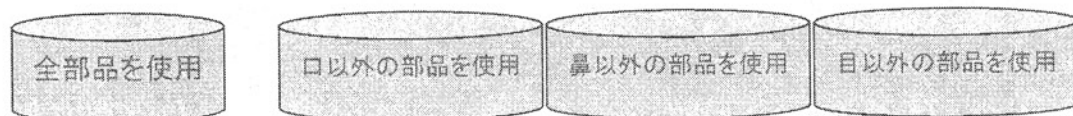


図4 識別の流れ(登録DB作成)



DB(平均と共分散行列)の登録人物との類似度をマハラノビス距離で算

さらにDempster&Shafer理論を用い複数枚の結果を統合



図5 識別の流れ(入力画像に対する識別)

③ 識別 (図5参照)

入力された(識別対象の)顔画像に対して切り出しが成功した部品で算出したパラメータと、登録画像で②で求めておいた同部品を使い算出したパラメータDBの登録人物との類似度を求める。入力画像の切り出し結果に応じてどのDBと比較を行うのかを決定する。識別はマハラノビス距離による類似度の算出で行う。その値により最も値の小さかった登録人物を入力画像の人物であると推測する。

ここで、切り出しが成功した部品が少ないときは識別の精度が低く、また正しく部品切り出しが行えた画像もより識別率を上げるために連続して撮影したほかの画像も識別に用いることを考える。つまり複数枚(3~6枚)の入力画像のマハラノビス距離の値を識別に用いる。ここでDempster&Shafer理論を使う。変換式(2)でマハラノビス距離の値を確率に変換してDempster&Shafer理論によりそれらの確率の値を統合し、1つの値にし最も確率の高い登録人を入力画像の人物であるとする事で個人識別を行う。障害物によって限られた部品のみしか識別に使えないようなときでも、このようにして複数枚の結果を統合して識別率の向上を図り顔の見え方に対してロバストな識別を行う。

3. おわりに

顔部品が障害物で隠れている画像など1枚のみでは誤識別する顔画像でも複数枚の識別結果を統合することにより識別率を向上させ

ることができた。このことから顔の見え方に対するロバストな個人識別を実装することが出来た。

参考文献

- [1] 島田英之, 磯部博文, 塩野充: “顔の両眼付近の部分画像を用いたゲート管理のための個人識別実験”, 電子情報通信学会論文誌D-II Vol. J77-D-II No. 9, pp. 1680-1690, 1994年9月.
- [2] 赤松茂: “コンピュータによる顔の認識の研究動向”, 電子情報通信学会誌 Vol. 80, No. 3, pp. 257-266, 1997年3月.
- [3] 栗田多喜男, 長谷川修: “顔画像からの個人識別”, 映像情報メディア学会誌Vo. 51, No. 8, pp. 1132-1135 (1997)
- [4] 橋本英樹, 福島邦彦: “選択的注意機構による顔の部分パターンの認識と切り出し”, 電子情報通信学会論文誌 D-II Vol. J80-D-II No. 8 pp. 2194-2202, 1997年8月.
- [5] 川出雅人, 細井聖: “実環境にロバストな顔認識技術とその応用”, オムロン 2002年3月.
- [6] 顔部品認識システムを開発した情報処理, 振興事業協会 (IPA), URL: <http://www.ipa.go.jp/>.