

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成25年 5月31日現在

機関番号：34315

研究種目：若手研究（B）

研究期間：2010～2012

課題番号：22700157

研究課題名（和文） 階層疑似 Fisher 行列による顔認識システム

研究課題名（英文） Face Recognition System using Hierarchical Pseudo Fisher Matrix

研究代表者

福水 洋平（FUKUMIZU YOHEI）

立命館大学・理工学部・助教

研究者番号：60467008

研究成果の概要（和文）：

顔検出および顔認識システムは、コンピュータの演算能力の拡大に伴って広く応用されつつあり、特に防犯カメラの高度化として不審者の発見や追跡に応用が期待されている。本研究は、顔検出、顔認識、並びにそれらをつなぐ顔器官検出のアルゴリズムに関するものである。ヒトの認知機構を参考にして、顔や人物の検出・認識性能が高く環境光などの外乱に堅牢な器官検出、検出並びに認識方式を導出し、その性能を実証した。また、その成果を人物行動の認識システムに発展させた。

研究成果の概要（英文）：

The increase in available computational power has led to the current widespread use of face recognition and detection systems. It is expected that many applications will be developed including, in particular, those designed to track and identify suspicious persons, which will enhance the functionality of security cameras. This study is concerned with algorithms that address face detection and face recognition, and face organ detection that connects them. We develop face organ detection, face detection, and recognition algorithms that show a high-level performance, as well as robustness against disturbances such as ambient light, referring to the mechanism that humans use to recognize each other. The performance results are presented. In addition, we extend the results of the study to recognition systems for personal behavior.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010年度	900,000	270,000	1,170,000
2011年度	900,000	270,000	1,170,000
2012年度	1,200,000	360,000	1,560,000
総計	3,000,000	900,000	3,900,000

研究分野：総合領域

科研費の分科・細目：情報学、知能情報学

キーワード：画像認識、機械学習、認知科学、脳型計算

## 1. 研究開始当初の背景

顔検出と顔認証のアルゴリズムは、このところ急速に進歩している。いずれも統計学習アルゴリズムの工学的応用によるところが大きい。

顔検出では、従来、肌色検出、平均顔検出等、Feature 特徴量を探索する手法が取られてきた。しかし、これらの手法では、顔の多様性（光、傾き、背景等）に対応するにおいて限界があった。そこで、より汎用的な特徴

量である、Haar 特徴量や HOG 特徴量 (Histograms of Oriented Gradients)等 を求め、これら特徴量に重みを付けて評価する (例えば 6 セグメントフィルタ) ことで、ロバスト性を向上させることが可能となった。次に、この重み付けを実画像での大量学習にて行うようになった。学習データにて、検出器特性を変える仕組みである。ニューラルネットワークや SVM (Support Vector Machine) 等と比較して、Boosting アルゴリズムの方が、各段に優れているというのが定評になっている。

顔認証研究は、1:1 認証から 1:N 認証へと移っている。1:1 認証は、本人と本人以外を識別する機能である、実用化も比較的容易である。Fourier 変換や Gabor 変換等を使って光環境耐性を上げることで製品化されている。1:N 認証になると格段に難しくなる。例えば 1000 人登録では、その 1000 人全てを識別しなければならない。重要な点は、顔画像情報を、個人特定に重要な特徴量だけに削減することである。現在、主成分分析や独立成分分析等による特徴空間の次元削減にて主要情報抽出を行っているが、N=10 人程度の Entertainment 用途に留まっている。最近、革新的手法が 2 つ提案されてきている。AAM (Active Appearance Model) と、LDA (Linear Discriminant Analysis:線形判別分析) である。AAM は、3次元顔モデルを使って任意方向の顔を認証する上で有望な手法であり、この手法を応用して、唇形状から発話母音を識別するシステムを構築、ほぼ 100% の認識精度を達成している。本研究テーマに関連するのは、線形判別分析である。これは、顔の多様性 (光、傾き、経年等) に対応するために、個人の複数枚の顔をグループとして扱って、主成分軸を決めて行く特徴量抽出手法である。この手法による実空間から特徴空間への射影を Fisher 行列と呼んでいる。

## 2. 研究の目的

本研究では、それぞれが独立成分に近いローカル検出器およびローカル認証器 (ローカル下位脳) と、その結果を多数決で総合判断する上位認知器 (上位脳) からなる、階層顔認知システムを構築し、このモデルを使って人間の顔認知機能を解明する。また並行して、モデル改良を進め、1:N 認証の高精度化を図り実用に供する。さらに、モデル構造を逐次更新する適応顔認知システムを LSI 化し、集積回路とシナプス回路網との対応を研究する。

顔検出で明らかにする第一としては、Haar や HOG 以上のより効率的な特徴量を検討し、どこまで特徴量を減らせるかを明らかにする。特徴量の機能と数の関係を明確にするこ

とで、人間が顔検出するときの刺激細胞の性質を明らかにすることに繋がる。第二には、カスケード型の Boosting 検出器の分化システムを明らかにする。例えば横顔検出で、ひとつのカスケード検出器の性能を、2 つに分化した検出器システムが上回る場合での、学習データ多様性との関係を明らかにする。また、カスケード並列システムが、多様なロバスト性 (横顔、斜め顔、眼鏡、髭等) に対応する場合、1 つのカスケード検出器にどの程度のロバスト耐性 (多様性役割) を持たせるのが効率的かを明らかにする。これは、人間が学習していく過程を明らかにすることに繋がる。そして第三としては、一度構築した検出器が、環境に適応して進化するプロセスをモデル化する。検出器の分化と退化機構を組み込み、学習と検出を一体化する。これは、人間の環境適応の仕組みを明らかにすることに繋がる。

顔認証で明らかにする第一としては、我々の認証モデル (擬似 Fisher 行列の階層構造) の有用性を認証精度の点から検証する。例えば、「光多様性擬似 Fisher 行列」と、「表情多様性擬似 Fisher 行列」とからなるシステムをその汎行列にて構築し、両方のロバスト性が強いシステムとする。これにより、人間の顔認証のロバスト性を、多くの専門脳 (擬似 Fisher 行列) の総合処理 (多数決処理) にてモデル化できるので、脳の高度な認証機構の解明に繋げることができる。第二には、我々の認証モデルの能力と、人間の認証能力とを定量比較する。例えば、光ロバスト性に着目し、学習時の顔データの数と Fisher 行列の次元数とが、1:N 認証の N 人を伸ばす上でどのように関与しているかを体系化し、これに照らして人間モニターの顔認証能力とを比較する。これは、人間脳の認証細胞の容量推定に繋がる。そして第三としては、人間の繰り返し学習の機構をモデルに導入し、その仕組みを脳に照らして明らかにする。例えば、人間は、常時逢っている人の認証精度が向上する。その人の多様な顔を学習しているからである。ここでは、構築された汎擬似 Fisher 行列が、逐次更新されるモデルで対応する。これは、人間のシナプス活性化機構の解明に繋げることができる。

顔検出および顔認証のいずれにおいても、世界トップクラスの精度を持ったシステムを構築してきている。この性能は、人間の能力と定量的に比較できるレベルにある。また、いずれも、LSI にて実装できる規模のものである。すなわち、この工学モデルを使って、(1)人間の顔認知機構の解明、(2)顔認知記憶細胞容量の定量化、(3)顔認知演算細胞能力の定量化に貢献できる。

### 3. 研究の方法

本研究は次の手順で進めた。

(1)顔検出や顔認知の前段階として、顔器官の検出に特化し、検出と認識の間に介在してその性能を改善するアルゴリズムを導出する。

(2)Boosting 手法に基づいた検出器についてアルゴリズム導出と性能評価を通じて外乱堅牢性や特徴量の機能を明らかにする。

(3)顔認知について、Fisher 行列を用いた LDA による 1:N 認識システムの性能評価を通し、LDA 手法の有用性を明らかにする。また、環境外乱に堅牢な特徴量を用いた認識システムを導出する。

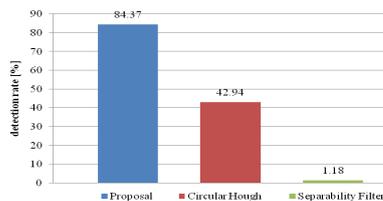
### 4. 研究成果

本研究は以下の成果を得た。

(1)検出・認知に至る前段階としてのアルゴリズムとして、①眼鏡画像除去アルゴリズム並びに②瞳検出アルゴリズムを導出した。

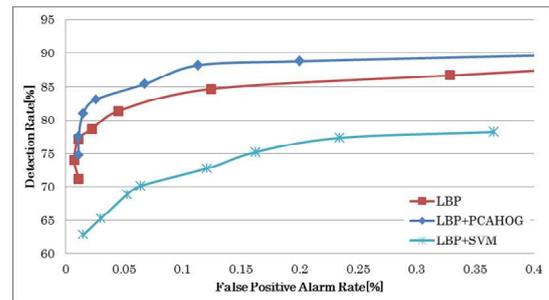
①眼鏡画像除去については、検出や認知の対象人物が眼鏡を着用していることによって顔検出並びに顔認知に大きな支障を来すことが以前から知られていた。そのため、本研究では眼鏡を着用した人物と着用していない人物、双方の入り交じった顔画像データベースより PCA を用いて主成分（基底）を求めておき、色空間処理により眼鏡特徴を弱化させた入力画像に対して眼鏡特徴を含まない固有顔で再構築することにより、眼鏡を着用した顔画像より眼鏡を除去することに成功した。この結果、HOIP データベースに対する瞳検出の成功率が従来法による 82.82%から 92.44%へ向上した。また、Valid データベースを用いた 1:1 顔認識により等価エラー率を従来法の 8.51%から 7.86%へ低減させることに成功した。

②瞳検出アルゴリズムについては、顔検出と顔認知の間を取り持つ顔器官検出手法の一つであり、左右の瞳の位置関係より顔の大きさや角度などが得られるため、顔認知に欠かせない情報を提供する。これまで円形 Hough 変換や分離度フィルタを用いたものが用いられてきたが、本研究ではエッジ保存型平滑フィルタである Mean Shift と局所特徴量に対して敏感に作用する Gabor フィルタを組み合わせた瞳検出アルゴリズムを導出した。この結果、FERET データベースを用いた性能評価の結果、次図に示す通り、従来手法に比して大きく検出性能を改善した。



(2)検出機能の研究として、①LBP 特徴量と Logit Boosting による人物検出アルゴリズム、②LBP, Boosting と Canny フィルタを組み合わせた環境外乱に頑健な顔検出アルゴリズム、並びに③Haar like 特徴量並びに HoG 特徴量による Real AdaBoost を用いた動物顔検出アルゴリズムを導出した。

①LBP 特徴量と Logit Boosting による人物検出アルゴリズムについては、画像中の人物部分を抽出するための手法であり、局所特徴量である LBP と、ロジスティック回帰を損失関数に取り入れた Logit Boosting を前段とし、膨大な LBP 特徴量を弱識別器でふるい分けるものである。さらに、この Booster によって得られた候補領域から HoG 特徴量を求め、PCA による主成分軸上への射影から Mahalanobis 距離により判定を行う後段の処理を組み合わせることによって、人物領域の抽出が可能となった。MIT CBCL の人物画像データセットおよび INRIA データベースを用いた評価の結果、次図に示す通り、前段のみのアルゴリズム（赤）および後段のみのアルゴリズムに SVM を適用したもの（水色）と比して、提案手法（青）は ROC カーブにおいて高い性能を示した。



また、処理時間についても次表の通り、前段のみの処理に若干のオーバーヘッドが加わるだけのものが得られた。

方法	提案	前段のみ	後段+SVM
時間(ms)	25.66	24.86	171.6
比率(%)	1.0	0.96	6.7

②LBP 特徴量を用いた Logit Boosting による強識別器に Canny フィルタを組み合わせ、入力画像のエッジ成分だけに反応する識別器出力を加えることにより、環境外乱に頑健な顔検出アルゴリズムを導出した。CASPEAL データベースによる評価の結果、環境光による影響が大きいデータについて次表の通り大きく検出率を改善した。

No.	1	2	3	4	5
従来	98.9%	84%	91.5%	95.3%	19%
提案	98.9%	88.8%	93.1%	99.2%	87.8%

③特徴量による検出性能への影響を比較するために、Haar like 特徴量および HoG 特徴量により猫顔の検出実験を行った。Real

AdaBoost を用いた検出器を構築し、双方の特徴量により学習させた識別器の性能を公開猫画像データベースを用いて評価したところ、次表の結果を得た。

	検出率	誤検出率	処理時間
Haar	63.2%	68.8%	5.58ms
HoG	72.5%	39.3%	67.6ms

結果として、猫顔のように毛並が個体ごとに異なる場合、人間顔でよく用いられてきた Haar like 特徴量よりも HoG 特徴量の方がよい性能を示した。これは、猫や動物の顔に限らず、人間顔においても髭や眼鏡などの多様な外乱に対する堅牢性を与えることを示している。

(3) 認知機能の研究として、①Fisher 行列による 1:N 顔認識システム、②LGDPHS オペレータによる年齢性別認知アルゴリズム、③LGDPHS オペレータによる表情認知アルゴリズム、並びに④時空間両軸への特徴点検出による行動分類アルゴリズムを導出した。

①PCA による固有顔と LDA による Fisher 行列を用いた 1:N 顔認識システムを構築し、表情多様性のある CASPEAL データベースに対して 1:N 顔認識における性能評価を行った。まず、本人性の評価にユークリッド距離を用いた場合を次表に示す。

	サイズ	20 次元	40 次元	50 次元
PCA	16x16	48.5%	59%	60%
	32x32	60%	64%	75%
LDA	16x16	59%	70%	72.5%
	32x32	61.5%	78%	84.5%

また、同様に相関係数を本人性の評価に用いた場合を次表に示す。

	サイズ	20 次元	40 次元	50 次元
PCA	16x16	56.5%	62%	64.5%
	32x32	62%	73.5%	81.5%
LDA	16x16	61.5%	73%	80%
	32x32	70%	84.5%	91%

以上の結果より、50 次元の LDA による Fisher 行列を用いて画像サイズを 32x32 画素とし、評価関数に相関係数を用いたとき、認識率が最もよくなった。

②局所特徴に着目する LBP 特徴量および LDP オペレータによる Boosting 検出器の性能改善に着目し、LDP オペレータに対して Gabor フィルタを組み合わせさせた LGDPHS オペレータを導出した。その上で、SVM により年齢および性別の認知推定アルゴリズムを導出した。FG-NET 年齢データベースを用いた年齢推定の評価結果を次表に示す。

手法	平均絶対誤差
提案手法	6.16
LBP+AAM	6.92
LGDP	7.89
AAM	7.07
AGES	6.77

また、FERET データベースを用いた性別推定の評価結果を次表に示す。

手法	分類率
提案手法	89.35%
LBP+AAM	85.87%
LGDP	88.91%
LGBP	84.57%
LBP	84.13%

いずれの分類においても LGDPHS オペレータによる分類が良好な結果を与えることを実証した。

③同じく、LGDPHS オペレータを用いて表情推定アルゴリズムを SVM 実装した。JAFFE データベースを用いた性能評価の結果、データベースに定義されている「怒り」「嫌悪」「恐れ」「幸せ」「中立」「悲しみ」「驚き」の各表情について、次表の結果を得た。

表情	個人非依存	個人依存
怒り	100.0%	100.0%
嫌悪	100.0%	100.0%
恐れ	40.0%	83.3%
幸せ	88.9%	91.7%
中立	100.0%	100.0%
悲しみ	100.0%	90.9%
驚き	55.6%	100.0%

個人非依存の場合について従来手法との性能比較を次表に示す。

手法	認識率(%)
提案手法	<b>82.81</b>
LGBP(Gabor + LBP)	67.19
LBP	53.12
LDP	54.69
LGBP based keypoints	77.62
LBP + MFA	65.71
LBP + Adaboost	65.71
HOSVD	66.0
DSF	62.78
SNE	65.7

また、個人依存の場合についての比較を次表に示す。

手法	認識率(%)
提案手法	<b>94.74</b>
LGBP(Gabor + LBP)	88.13
LBP	72.37
Radial encoded Gabor jets	89.67
Patch-based Gabor	92.9
Boosted-LBP	81.0
Gabor + FSLP	91.0
LDP	76.32
DCT	79.30
KCCA	77.05

いずれの場合にも LGDPHS オペレータが有効な性能を示した。

④行動分類システムとしては、時空間変動

に対して堅牢性が必要であるため、時空間の特徴点検出として、人物の動作を精度よくとらえることができるガボールフィルタをマルチスケール応用し、特徴ベクトルの算出に勾配特徴を用い、これを局所特徴量の位置情報を省いた Bag-of-Feature により特徴づける手法を導出した。KTH データセットを用いて行動分類を行った結果、次表の結果を得た。

行動	識別率
歩行する	87%
ジョギング	55%
疾走する	69%
ボクシング	95%
手を振る	81%
拍手する	92%

また、従来手法である Dollar の特徴点検出法と比しても 3%程度の性能向上が見られた。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 2 件)

- ① A. Higashi, Y. Fukumizu, T. Izumi and H. Yamauchi, "Expression Recognition Using LGDPHS Based Facial Key Parts," 画像電子学会誌, 査読有り, Vol. 41, No. 5, 2012, pp. 496-505.
- ② A. Higashi, Y. Fukumizu, T. Izumi and H. Yamauchi, "Age and Gender Estimation Using Global and Local Features with AAM and LGDPHS," 画像電子学会誌, 査読有り, Vol. 41, No. 3, 2012, pp. 262-269.

[学会発表] (計 3 件)

- ① 発表者名 : A. Higashi, Y. Fukumizu, H. Yamauchi  
 発表演題 : Expression Recognition using Local Gabor Directional Pattern Histogram Sequence (LGDPHS)  
 学会名等 : 2012 International Workshop on Nonlinear Circuits and Signal Processing  
 発表年月日 : 2012 年 3 月 6 日  
 発表場所 : ハワイ州 (アメリカ合衆国)
- ② 発表者名 : A. Higashi, T. Yasui, Y. Fukumizu, H. Yamauchi  
 発表演題 : Local Gabor Directional Pattern Histogram Sequence (LGDPHS) for Age and Gender Classification  
 学会名等 : 2011 IEEE Workshop on Statistical Signal Processing  
 発表年月日 : 2011 年 6 月 29 日

発表場所 : ニース (フランス共和国)

- ③ 発表者名 : M. Kamiya, Y. Fukumizu, H. Yamauchi  
 発表演題 : High Accuracy Eyes Detection Method by Using Gabor Filter and Dividing Emphasis Processing  
 学会名等 : SISA2010  
 発表年月日 : 2010 年 9 月 9 日  
 発表場所 : マニラ (フィリピン共和国)

#### 6. 研究組織

##### (1) 研究代表者

福水 洋平 (FUKUMIZU YOHEI)  
 立命館大学・理工学部・助教  
 研究者番号 : 60467008