

平成22年 5月31日現在

研究種目：基盤研究（B）  
 研究期間：2007～2009  
 課題番号：19300079  
 研究課題名（和文） リアルタイム高次統合視覚集積システムの開発とマーカレス身振り認識への応用

研究課題名（英文） Development of real-time high-level integrated vision systems and their application to markerless posture recognition

## 研究代表者

森江 隆 (MORIE TAKASHI)  
 九州工業大学・大学院生命体工学研究科・教授  
 研究者番号：20294530

研究成果の概要（和文）：脳の視覚情報処理機能に学んだ顔・身振り認識アルゴリズムを提案し、それをリアルタイムで実行する専用集積回路の統合システム構成法を開発した。脳の初期視覚処理モデルである、局所的な画像の空間周波数を特徴として抽出するガボールフィルタを、立体視やバイオロジカルモーションに基づく動き検出などに適用した。また、画素の状態を異方性拡散する集積回路によりリアルタイムで主観的輪郭を生成するハードウェア・ソフトウェア統合システムなどを開発した。

研究成果の概要（英文）：We proposed a face/arm posture recognition algorithm inspired by the brain visual information processing, and developed a system combining dedicated integrated circuits for real-time operation of the algorithm. As one of the brain primary visual information processing models, Gabor filtering, which extracts features of local image spatial frequencies, was effectively applied to stereo vision and moving detection based on biological motion. We also developed hardware/software integrated systems such as a real-time subjective-contour generation system using an integrated circuit that performs anisotropic diffusion of pixel states.

## 交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	4,100,000	1,230,000	5,330,000
2008年度	4,800,000	1,440,000	6,240,000
2009年度	5,200,000	1,560,000	6,760,000
年度			
総計	14,100,000	4,230,000	18,330,000

研究分野：総合領域

科研費の分科・細目：情報学・感性情報学・ソフトコンピューティング

キーワード：ニューラルネットワーク，集積回路（LSI），ガボールフィルタ，顔・物体・身振り認識，バイオロジカルモーション，大局的領域分割

## 1. 研究開始当初の背景

(1) 視覚モデルとハードウェア：人工視覚・画像認識については、今日に至るまでソフトウェア処理向けのアルゴリズム開発が主流

であり、常に計算量削減の努力が払われてきた。しかし、結果として、膨大な計算を神経回路という超並列ハードウェアで実行する現実の脳が実現できる機能には全く到達で

きていない。一方で、これまでに脳の機能をまねた LSI が各種発表されてきた。例えば、八木らは人工網膜チップを開発し、それと FPGA を組み合わせた初期視覚システムを構築している。また、動き検出を行うチップの開発も行われている。しかし、まだ人の持つ多彩な視覚機能のごく一部が実現できはじめたばかりである。この方面に関しては、研究代表者らはすでに V1 機能を実現するガボールフィルタ LSI を開発し、その機能を確認している。

一方、視覚系全体のモデルとして、ネオコグニトロンに代表される畳み込みニューラルネットワーク (Convolutional Neural Network; CoNN) が知られている。これは、強力で頑健な視覚認識モデルであるが、計算量が多いため、ソフトウェアのみの実装では実用に用いられることが難しかった。これについても、研究代表者らは顔検出を主な目的とした CoNN チップを開発している。

研究代表者はこれまでアナログとデジタル回路方式を融合したアナログ・デジタル融合アーキテクチャを提案し、それを用いて、既存のデジタル方式では困難な超並列 (画素並列) 動作型の様々な画像処理・視覚処理 LSI を開発してきた。

(2) 身振り認識技術：人間の身振りを認識する場合、従来の方法では条件の整った実験室内で特別なマーカーを用いる等の制約がある。一方、身体モデルを用いた単眼視による身体運動計測アルゴリズムが開発されており、手軽な身振り認識の道具として適用可能であるが、マニュアル・フィッティング操作が必要という欠点がある。立体視を用いた内部モデルとのフィッティング法が提案されているが、アルゴリズムレベルのものである。研究分担者 (宮本) は、より改良した内部モデルフィッティングによる身振り認識法を開発した。本研究ではこれを統合視覚集積システムと結合することにより、高性能な身振り認識システムを提案する。

(3) 心理・生理学実験によるバイオロジカルモーション認識メカニズム：人は、人間の体を 10 個程度の点に置き換え、それを体の動きに対応させて動かした場合にもその身振りを認識できることが知られている。この現象はバイオロジカルモーションと呼ばれ、人間が視覚的運動情報のみから身振り認識を行う能力を有していることを示している。しかし、この現象を工学的に応用する試みはほとんどなされていない。

## 2. 研究の目的

本研究では、脳の視覚系を真似た各種視覚処理用集積回路を改良・発展・統合し、立体視・動き検出などと組み合わせ高次視覚機能を実現するリアルタイム統合視覚集積シ

ステムを構築する。この視覚集積システムを用いて、何の付加装置 (マーカーなど) もなしに人の身振りを認識する方法を確立する。これは、身体内部モデルとのマッチングによる身振り認識アルゴリズムをハードウェアによりリアルタイム化・頑健化を図ることで実現する。さらに、バイオロジカルモーションを中心とする視覚心理実験および生理学実験により、より人に近い視覚認識システムへ向けた研究を進めるとともに、脳の視覚機能の解明に貢献することを目指す。

## 3. 研究の方法

(1) ガボールフィルタを用いた多重解像度マッチングによる立体視・動き検出アルゴリズムの開発：

ステレオ視では距離を算出するために、対応点探索 (ステレオマッチング) を行う必要がある。これには、一般に正規化相互相関

(NCC) を用いた窓相関法が用いられる。この方法は計算自体が単純でハードウェア化に向いているが、計算量が多い、窓内に輝度変化がないと誤認識する、物体の境界がぼけるといった問題がある。そこで、ガボールフィルタと窓相関法を組み合わせ、計算量と誤認識を減らすことにより、高速かつ高精度なステレオマッチング手法を提案した。

ガボールフィルタは、 $\sin$ ,  $\cos$  関数をガウス関数で局在化したカーネルを用いる空間フィルタで、画像の任意方向の局所的空間周波数を抽出する。水平方向の空間周波数を持つガボール特徴量を用いて窓相関法で計算した場合、対応点付近で相関が緩やかに変化する。視差の計算精度は落ちるが、数 pixel ごとに計算しても、対応点付近で高い相関が算出できるため、大まかな視差の推定が可能になる。そこで、ガボール特徴量を用いた窓相関法で大まかな視差の推定を行った後、通常の輝度を用いた窓相関法で視差計算を細かく行うことにより、高精度かつ高速な視差計算が可能になると考えられる。

また、窓内のガボール特徴量の平均値から、水平方向の輝度変化がない場合を容易に検出できるため、その場合は、相関計算を省略することで、誤認識を防ぐことができる。

(2) 身体内部モデルに基づく身振り認識アルゴリズムの開発：

ステレオカメラにより検出した顔の視差情報を利用して、背景や服装の環境に依存しない人腕の姿勢推定法を開発した。検出した顔の視差を腕モデルと人腕とのマッチングに利用するため、画像全体で視差を計算する必要がなく、腕姿勢の探索範囲を削減できる。マッチングさせる腕モデル近傍に配置した三種類のフィン内の特徴をマッチングの評価に使用する。左右画像のフィン内の二つの

特徴を比較するので背景が一様でなくても、光量や肌色に左右されることも少なく、人間の腕と腕モデルを一致させることができる。

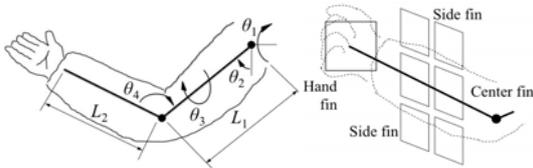


図1：腕のスケルトンモデルとマッチングフィン

モデルフィッティングに利用されるモデルには、人体近似モデルや円筒、楕円型モデルなどがあるが、本手法では、他と比べて計算量が少ないスケルトンタイプを採用した。この腕モデルは各関節角度  $\theta_1 \sim \theta_4$  を持つ (図1)。腕全体の姿勢推定を目的とするため手首関節は除外した。腕モデルの近傍には、三種類のフィン (Center, Side, Hand) を配置した (図1)。

Center フィンは腕や服を含むよう配置し、色やテクスチャ情報を評価する。Side フィンは腕や服の輪郭を含むように配置し、エッジ情報を評価する。Hand フィンではエッジ情報を評価する。Side と Hand フィンではエッジ上の色情報の類似度も評価に含める。これらフィンでの評価を重み付けして組み合わせ、腕全体の評価値を計算する。

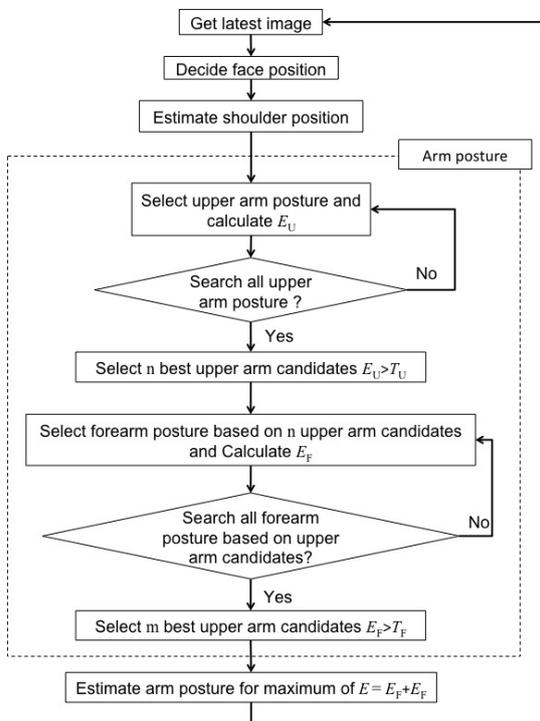


図2：腕姿勢決定フローチャート

探索範囲内の各腕姿勢で評価値を計算し、評価値が最大となる各関節角度  $\theta_1 \sim \theta_4$  を人腕の姿勢とする (図2)。

(3) バイオロジカルモーション知覚研究と応用：

バイオロジカルモーションを工学的に利用しようとする、関節などを検出する必要があり、そのこと自体が極めて難しい処理になる。そこで、手足の動きを検出して身振り認識などに応用することを検討した。動き検出には、通常ブロックマッチング法に比べて、頑健な特徴抽出ができる時空間ガボールフィルタを用いた。

(4) 統合視覚集積システムの開発：

自然画像について強力な特徴抽出が可能なガボールフィルタを実行する集積回路 (LSI) とその他の画像処理 LSI を組み合わせ、高速・低電力・省容積の視覚システムを構築する。これまで開発した専用視覚処理 LSI は、LSI 内部では PWM 信号を用いたアナログの処理を行うが、入出力はデジタル信号で行う方式としているので、デジタル LSI (FPGA) により専用視覚処理 LSI を制御し、PC・カメラと接続して、システムを構築する。本研究では、専用視覚処理 LSI として、ガボールフィルタ LSI、および画像状態の特定方向への拡散を行う「異方性拡散 LSI」を用いて、主観的輪郭生成を実現する視覚集積システムを構築することとした。

4. 研究成果

(1) ガボールフィルタを用いた多重解像度マッチングによる立体視・動き検出アルゴリズムの開発：

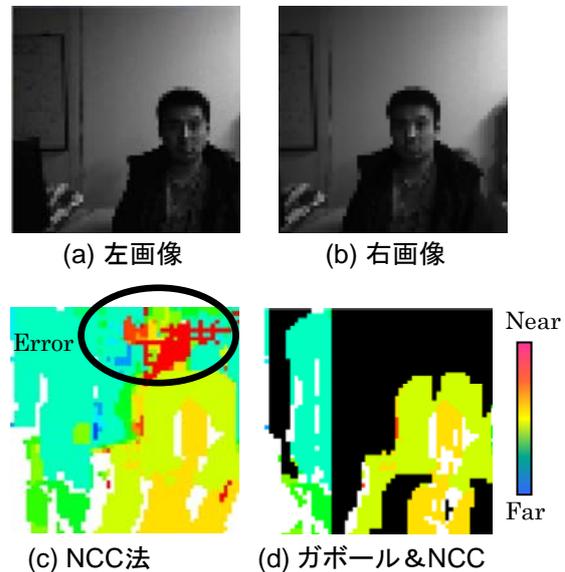


図3：ガボールフィルタを用いた立体視  
ガボールフィルタを用いるステレオマッチング手法を検討した。実験結果を図3に示す。(c)、(d)の結果から、窓内のガボール特徴量の平均値を用いることで、輝度変化の少ない領域で誤認識が防げることが確認できた。また、ガボール特徴量と NCC を組み合

わせた提案手法では、平均 50ms で視差を算出できた。NCC のみの場合は視差算出に 75ms かかったので、提案手法の高速性が確認できた。

以上より、水平方向のガボール特徴量から、横方向の輝度変化を検出することにより誤検出を低減し、ガボール特徴量と窓相関法を組み合わせることで、高速かつ高精度なステレオマッチングが実現できることを示した。

(2) 身体内部モデルに基づく身振り認識アルゴリズムの開発：

顔位置から推定した腕の位置を用いて、腕の姿勢を準リアルタイム (3-7 フレーム/秒) で検出するシステムを構築した。本システムはソフトウェアレベルでモジュール化され、ソフトウェア・ハードウェア (専用集積回路利用) のモジュールが交換可能であり、ガボールフィルタ LSI を組み込むことも可能である。このシステムと既開発の顔・物体認識システムを組み合わせ、特定の人物の身ぶりによって制御する簡易ロボットを開発した。

身振り認識の対象となる人物を決定するため、ステレオカメラから得られた画像に対して顔領域の検出 (顔検出) を行い、さらに予め登録しておいた顔と照合 (顔認識) する。顔検出にはオープンソースで提供されている OpenCV ライブラリに付属する Viola & Jones の方法、顔認識には既開発のエラスティック・グラフ・マッチング・プログラムを用いた。構築した認識システムの PC 表示結果の一例を図 4 に示す。



図 4：顔・身振り認識システム

通常の顔検出法に加えて、人物同士が重なったり、対象人物が動いたときに顔を見失わないように、テンプレートマッチングによる顔の追跡と顔照合用エラスティック・グラフ・マッチング機能を組み合わせて使用する手法を追加して、ロバスト性を高めた。

身振り認識を、大人から子供まで身体の大さにかかわらず精度良く実現するためには、対象ごとに体格 (顔から肩、肩から肘および肘から手先の長さ) を推定する必要がある。そのため、ステレオ視により推定した対象人

物までの距離、およびカメラの仰俯角から対象人物の身長を推定し、それによって顔位置から肩位置を推定する方式を開発した。推定結果の精度を検証するため、実験を行った。対象の身長を 169cm、俯角  $\theta = 0, 5, 10, 15$  度とし、各俯角でのカメラから対象までの距離を 90~230cm とした。結果を表 1 に示す。

表 1 俯角に応じた身長推定結果

	$H_{\text{mean}}$ [cm]	$H_{\text{STD}}$ [cm]
$\theta = 0$ [deg]	166.9	1.22
$\theta = 5$ [deg]	168.6	1.26
$\theta = 10$ [deg]	170.1	1.30
$\theta = 15$ [deg]	169.4	1.42

$H_{\text{mean}}$  は全ての距離で推定した身長の平均を表し、 $H_{\text{STD}}$  は  $H_{\text{mean}}$  からの標準偏差を表す。俯角が大きくなるに従い  $H_{\text{STD}}$  は増える傾向にあるが、身長 の 1% 程度である。この結果より、この推定手法は体格を推定するには十分な精度を有することがわかった。

次に、腕姿勢推定法の精度と複雑な背景への対応を検証するため、室内環境において腕姿勢推定実験を行った。十名の被験者が以下の条件で自由な腕姿勢をとった。計算量軽減のために推定対象は右腕のみとした。被験者の服装は、半袖や長袖、手袋などとした。被験者の姿勢は、カメラから約 1~2m 離れて直立や着席状態とした。

本実験では、Core 2 Duo 2.6GHz の MacBook Pro を用いた。ステレオカメラとして Bumblebee PointGrey (CCD = 1/3 inch, 基線長 = 12 cm, 焦点距離 = 3.8 mm) を使用した。画像は左右それぞれ 320x240 pixel で撮影した。画像中のエッジ方向は  $0, \pi/4, \pi/2, 3\pi/4$  とした。顔位置から肩までの距離  $L_{\text{sh}}$  は 0.325 m とし、腕モデルの各節の長さ  $L_1, L_2$  は  $L_1 = 0.2$  m,  $L_2 = 0.2$  m とした。

表 2：探索幅 4 および 1 度での認識精度

	$\theta_{\text{step}} = 4 \text{ deg}$			
	$\theta_1$	$\theta_2$	$\theta_3$	$\theta_4$
$\Delta_{\text{mean}}$ [deg]	-0.9	-0.8	-3.7	1.8
$\Delta_{\text{STD}}$ [deg]	9.7	6.1	10.3	10.4
Speed [fps]	2 - 3			
	$\theta_{\text{step}} = 1 \text{ deg}$			
	$\theta_1$	$\theta_2$	$\theta_3$	$\theta_4$
$\Delta_{\text{mean}}$ [deg]	-2.1	-2.2	-1.7	1.0
$\Delta_{\text{STD}}$ [deg]	7.8	5.1	9.7	8.1
Speed [fps]	$0.8 \times 10^{-3}$			

まず腕姿勢の認識精度を評価するために、探索幅  $\theta_{\text{step}} = 4$  および 1 度で実験を行った。データ数は 90 とした。結果を表 2 に示す。

$\Delta_{\text{mean}}$  は平均,  $\Delta_{\text{STD}}$  は標準偏差を示す.  $\theta_{\text{step}} = 4$  度ではある程度良好な精度が得られた. 1 度では精度は高いが, リアルタイムで動作しない.  $\theta_{\text{step}} = 1$  度の条件はリアルタイムでの結果を必要とせず, 高い精度が必要なアプリケーションに適用できる.



図 5 : 腕姿勢推定実験結果

次に, 様々な環境下での腕姿勢推定実験を行った (図 5). 表 2 の結果から探索幅は  $\theta_{\text{step}} = 4$  度とした. 本手法は肌色を評価しないため, 図 5 右下の被験者は黒い手袋をしているが, 腕姿勢を推定できる. このように背景, 服装, カメラからの距離にかかわらず, 腕姿勢を認識できた.

(3) バイオロジカルモーション知覚研究と応用:

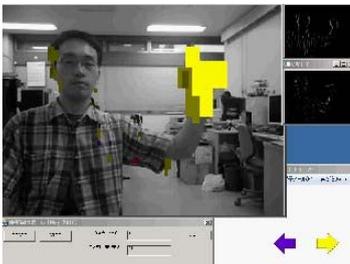


図 6 : 時空間ガボールフィルタによる腕振り検出

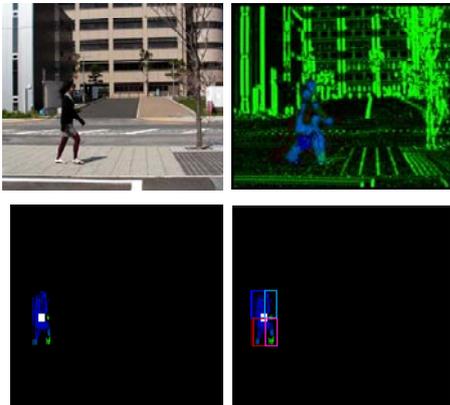


図 7 : 時空間ガボールフィルタによる歩行検出

バイオロジカルモーションの工学応用として, 時空間ガボールフィルタにより, 身振り認識に適用可能な腕の振りの検出 (図 6)

と, 足の動きによる歩行の検出 (図 7) を実現した. 前者については, 顔・身振り認識システムに統合した. 後者については, 身体の部分的な動きの周波数特性を用いて, 歩行者検出に適用した.

(4) 統合視覚集積システムの開発:



図 8 : ガボールフィルタ LSI と制御用 FPGA (下側基板)

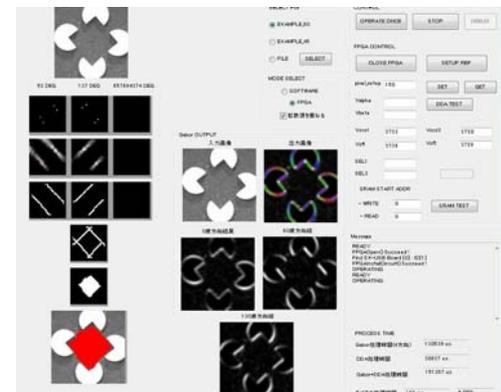


図 9 : 統合集積システムによる主観的輪郭生成の処理画面

ガボールフィルタのリアルタイム処理を実現するために, 既開発のガボールフィルタ LSI を制御し, カメラと PC から成る画像取得・処理システムと連携するデジタル回路を設計し, FPGA に実装した. これを, 既開発の顔・物体認識システムに組み込むことで, 処理の高速化を実現した (図 8).

統合集積システムとして, ガボールフィルタ LSI と他の視覚処理用 LSI を併せて FPGA で制御し, それらを PC と結合するシステムを構築し, 各種視覚 LSI を共通で制御するシステム構成法を確立した. FPGA に実装する制御用 HDL (ハードウェア記述言語) 記述法を体系的に整備し, 様々な専用視覚処理 LSI の制御を同一手法で実現できるようにした. その実現システムの一つとして, 異方性拡散 LSI とガボールフィルタ LSI を組み合わせた主観的輪廓生成システムを構築した (図 9). また, 入力画像を大まかに分割する大局的領域分割処理として, 領域ベース MRF モデルを研究し, その有効性を示すと共に, AD 融合方式により LSI を設計した.

## 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 13 件)

- ① 財津賢一郎, 金 永宰, 栗谷康隆, 森江隆, 画素並列型異方性拡散 LSI とそれを用いてエッジ補完を行う画像処理システム, 映像情報メディア学会技術報告 情報センシング研究会 (IST2009-94), Vol. 33, No. 56, pp. 17-20, 2009. (査読無)
- ② I. R. Khan, T. Morie, H. Miyamoto, Y. Kuriya and M. Shimizu, Real-Time Human-Machine Interaction System Based on Face Authentication and Arm Posture Recognition, Brain-Inspired Information Technology (Studies in Computational Intelligence), Vol. 266/2010, pp. 141-145, Springer, 2009. (査読有)
- ③ 高橋和志, 花沢明俊, 時空間ガボールフィルタによる歩行者検出, 第 8 回 ITS シンポジウム 2009, Peer-Review Proceedings, pp. 187-192, 2009. (査読有)
- ④ Y. Kawashima, D. Atuti, K. Nakada, M. Okada and T. Morie, Coarse Image Region Segmentation Using Region- and Boundary-based Coupled MRF Models and Their PWM VLSI Implementation, Proc. Int. Joint Conf. on Neural Networks, pp. 1559-1565, 2009. (査読有)
- ⑤ M. Shimizu, T. Yoshizuka, H. Miyamoto, Markerless Gesture Recognition under Cluttered Background, Proc. SCIS & ISIS pp. 1557-1561, 2008. (査読有)
- ⑥ I. R. Khan, T. Morie, H. Miyamoto, M. Shimizu, Y. Kuriya, A Prototype System for Secure Human-Machine Interaction Based on Face and Gesture Recognition, Proc. 34th Annual Conf. of the IEEE Industrial Electronics Society, pp. 1572-1577, 2008. (査読有)
- ⑦ I. R. Khan, T. Morie, H. Miyamoto, Face and Arm-Posture Recognition for Secure Human-Machine Interaction, Proc. IEEE Int. Conf. on Systems, Man and Cybernetics, pp. 411-417. 2008. (査読有)

[学会発表] (計 12 件)

- ① 森江 隆, 車載・ロボット視覚のための知的画像処理技術と脳型 LSI 実装, 電子情報通信学会 総合大会, 2010 年 3 月 18 日, 仙台.
- ② 高橋和志, 花沢明俊, 相対運動知覚における周辺情報の影響, 日本神経科学学会, 2009 年 7 月 10 日, 東京.
- ③ 松坂建治, 森江 隆, 大局的画像領域分割のための領域ベース結合 MRF モデルを実現する CMOS 回路, 電気学会 電子・情報・システ

ム部門大会, 2009 年 9 月 4 日, 徳島.

- ④ 鎌田卓治, 森江隆, ガボール特徴量を用いた相関法による高速・高精度ステレオマッチング法, 電子情報通信学会総合大会, 2009 年 3 月 17 日, 松山.
- ⑤ 宮本弘之, 栗谷康隆, 鎌田卓治, 清水昌樹, インディアク・ラスール・カーン, 花沢明俊, 森江 隆, 実時間顔・腕姿勢認識システムと影消去技術, 電子情報通信学会 総合大会, 2008 年 3 月 19 日, 北九州.

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

○取得状況 (計 1 件)

名称: 信号処理装置及び信号処理方法

発明者: 梅澤淳, 森江隆, 岩田穆

権利者: 九州工業大学

種類: 特許

番号: 4318936

取得年月日: 2009 年 6 月 5 日

国内外の別: 国内

[その他]

ホームページ等

<http://www.brain.kyutech.ac.jp/~morie/>

## 6. 研究組織

(1) 研究代表者

森江 隆 (MORIE TAKASHI)

九州工業大学・大学院生命体工学研究科・教授

研究者番号: 20294530

(2) 研究分担者

宮本 弘之 (MIYAMOTO HIROYUKI)

九州工業大学・大学院生命体工学研究科・准教授

研究者番号: 20336100

花沢 明俊 (HANAZAWA AKITOSHI)

九州工業大学・大学院生命体工学研究科・准教授

研究者番号: 10280588

(H20→H21: 連携研究者)

(3) 連携研究者

佐藤 能臣 (SATO YASUOMI)

九州工業大学・大学院生命体工学研究科・助教

研究者番号: 40548442

(H21: 連携研究者)