

# 知的画像認識技術と脳型 LSI 実装

Intelligent Image Recognition Technology and Brain-like LSI Implementation

森江 隆 石川聖二

## Abstract

車載用及びロボット視覚用の知的画像認識で必要となる HOG, SIFT などの基本アルゴリズムを紹介するとともに、筆者らが開発した画像認識手法を紹介する。現在の画像処理はより並列的・階層的になり、脳での視覚処理モデルに近づいているといえるが、人の高い知覚機能に近づくには更なるブレイクスルーが必要である。そのために、脳型画像処理技術とそれを実現する集積回路及びナノ構造の利用を含めた脳型デバイス開発の必要性を述べる。

キーワード：画像認識，脳型視覚処理，スパイクニューロン，ナノ構造デバイス

## 1. ま え が き

視覚は五感のうちで最も情報量の多い感覚であり、それを工学的に実現する画像認識に関する研究には非常に長い歴史がある。その有望な応用先の一つとして、自動車の知能化に関連して、運転者支援のための車載カメラによる自動車・歩行者認識が注目されている。また、ロボットの分野でも、工場の生産ラインのような規格化された環境で働く作業用ロボットだけでなく、家庭のリビングルームのような非定形環境で人間と共存する自律知能ロボットにも注目が集まっている。ここでは、複雑な環境での画像認識処理が重要な役割を果たす。

本稿では、筆者らの研究を中心に、車載用及びロボット視覚用の知的画像認識技術を簡単に紹介するとともに、そのような認識システムに今後必要となる脳型処理とそれを実行する集積回路の開発例を紹介し、将来展望を述べる。

## 2. 車載用画像認識技術

固定された監視カメラの画像から人を検出・認識する技術は既に実用化されている。また、見通しの良い道路の脇を歩いている歩行者を見つけ出すなど、比較的条件的の良い状況での認識もそれほど難しくない。実際、レーダやステレオカメラを用いた認識システムは既に実用化されている。しかし、自動車に搭載されたカメラから、繁華街の路上の多くの歩行者の中から、車道に飛び出そうとしているような危険な行動をとろうとしている人だけを検出することは極めて難しい。

市街地での車載カメラによる歩行者認識は、日照・街灯の照らし具合の変化などに頑健であること、電柱・立て看板など歩行者以外の種々雑多な物体と区別できること、様々な服装によらず人として認識できること、ガードレールに人の一部が隠されるなどのオクルージョンにも頑健であることなどが要求される。そこで、歩行者全身の輪郭特徴を取り出す手法として、画像の部分的な小領域ごとの輝度勾配方向のヒストグラムを求めて、それを特徴量とする HOG (Histograms of Oriented Gradients) と呼ばれる手法が提案され、有効性が示されている。この特徴を SVM (Support Vector Machine) などのよく知られた識別器で処理することで、歩行者識別が可能になる<sup>(1)</sup>。識別器に与える学習データを大量に用意することで、歩行者検出の精度を向上できる。

歩行者を検出できたら、次はその挙動を認識する必要がある。これを実現するために、タンらは人物シルエット

森江 隆 正員 九州工業大学大学院生命体工学研究科脳情報専攻  
E-mail morie@brain.kyutech.ac.jp  
石川聖二 正員 九州工業大学大学院工学研究院機械知能工学研究系  
E-mail ishikawa@cntl.kyutech.ac.jp  
Takashi MORIE, Member (Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, Kitakyushu-shi, 808-0196 Japan) and Seiji ISHIKAWA, Member (Graduate School of Engineering, Kyushu Institute of Technology, Kitakyushu-shi, 804-8550 Japan).  
電子情報通信学会誌 Vol.94 No.6 pp.459-463 2011年6月  
©電子情報通信学会 2011

トを構造化固有空間で分類する方法を提案した<sup>(2)</sup>。その一例を図1に示す。この手法を用いるには挙動が識別できる程度にシルエットを取り出す必要があるが、前景と背景を画素レベルで分離することで、これを実現する手法の開発も行っている。

一方、上述のHOGと、AdaBoost及びSVMの2段

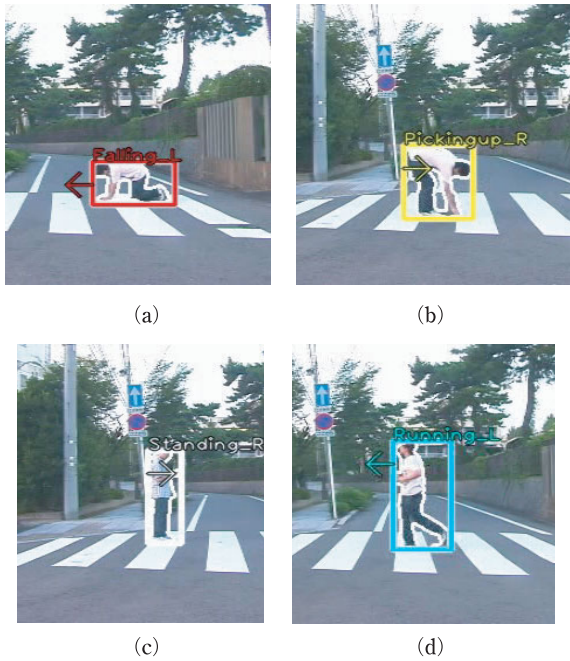


図1 人物挙動認識 モーションデータベースから構築した構造化固有空間法。(a)「転倒」、(b)「拾う」、(c)「立ち止まる」、(d)「走る」の各挙動を認識している。

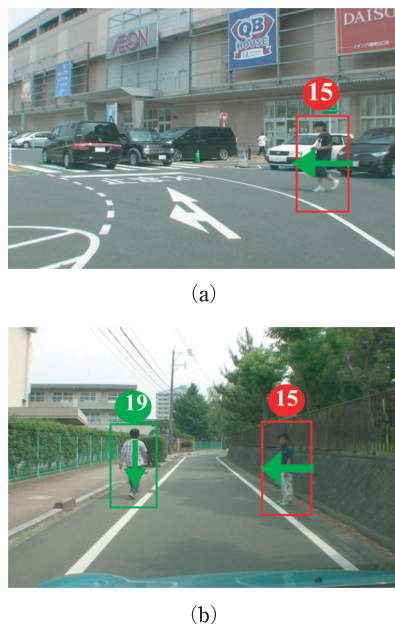


図2 人の身体向き検出 HOGと複数の識別器を組み合わせ、人の身体の向きを検出することができる。枠上部の数字はカメラからの推定距離(m)。

の識別器を組み合わせ、歩行者の身体の向きを検出する手法も提案している<sup>(3)</sup>。結果の一例を図2に示す。これにより、歩道に沿って歩いているか、道を渡ろうとしているかを識別できる。

HOGにおけるヒストグラムの利用では、小領域内の特定方向エッジの多寡を見るだけで、近隣エッジ間の関係を見ていない。そこで、HOGの局所特徴同士を組み合わせ、より高次の特徴量を得るように発展させた手法が盛んに研究され、検出精度を向上させている<sup>(4)</sup>。

### 3. 自律知能ロボット用視覚のための画像認識技術

ロボットがショッピングモールや家庭などの非定形環境で自律的・知的に行動できるようにするには様々な課題があるが、まず高度な視覚機能が必要となる。

車載用歩行者検出の場合は、歩行者か否かという一般物体認識技術が必要であったが、知能ロボット視覚ではそのような課題だけでなく、一度見た特定の物体を別の場面で探索するという特定物体認識技術が必要である。このような課題では、数多くの画像を用意して識別器に学習させる手法は有効でない。この課題には、DoG (Difference-of-Gaussian) フィルタ処理と勾配強度・方位のヒストグラムを用いることで、画像の拡大縮小・回転に不変な特徴量を算出するSIFT (Scale-Invariant Feature Transform) と呼ばれる手法を用いることが有効である<sup>(5)</sup>。

また、ロボットと人間とのインタフェースとして、音



図3 顔・腕姿勢認識システムの出力例 複雑な背景や、腕の一部遮蔽にも頑健な腕姿勢認識を実現した。

声だけでなく、身振り認識技術が重要である。身振りや姿勢を認識するには、腕姿勢などを知る必要があるが、歩行者の場合と同様に、様々な服装や背景のためにその認識は容易ではない。ここでは筆者らのグループが開発したステレオ視を用いた腕姿勢認識技術について紹介する<sup>(6)</sup>。

ステレオ視は、左右のカメラ画像において対応する特徴点の位置ずれ（視差）から、その物体までの距離を計算するが、一般の複雑な状況では、誤対応または特徴点が見つからないという問題が発生する。上記の応用では、ロボットと相対する人の顔をまず認識して、顔位置から肩位置を推定し、更に腕のモデルを仮定してマッチングを行う。ステレオ視の誤対応問題を避けるために、左右画像のマッチングで視差を求めることをせず、顔領域の視差を用いて左右画像と腕モデルとのマッチングを行う。これにより、背景や服装などの変化に頑健な腕姿勢認識を行うことが可能になる。認識結果の一例を図3に示す。

#### 4. 脳型視覚処理と専用 LSI の開発例

さて、以上で述べた一連の手法は、コンピュータビジョンからのアプローチであるが、他方で人の脳の視覚

機能をまねて、より高度な画像認識機能を実現しようとする流れがある。それは、人とのインタラクションが重要な車載用やロボット視覚において特に重要である。そのような応用では、人と全く異なった原理に基づく視覚ではなく、人と類似の視覚を持たせる必要があるからである。

人の視覚系をモデルとした画像処理手法の一つとして、福島のネオコグニトロンに端を発する畳込み型ネットワーク（Convolutional Neural Networks）がある。これは、初期視覚野の処理に対応したガボールフィルタにより、入力画像を細かいエッジ・勾配特徴に分解し、部分特徴を抽出して、それを高次視覚野に向かって階層的に統合していくモデルである<sup>(7)</sup>。この種のモデルを用いて歩行者検出を試みた例が報告されている<sup>(8)</sup>。

3. ままで紹介した従来のコンピュータビジョンによる手法が、SIFT や HOG の発展形のように多重解像度による特徴抽出や部分特徴の統合の流れにあることから、これらが脳型画像処理手法に接近してきているともいえる。実際、両者の実質的な演算処理の一部はかなり等価なものとも考えることもできる。今後、ハードウェアの計算能力が更に向上するにつれて、この傾向はますます強まっていくと予想される。

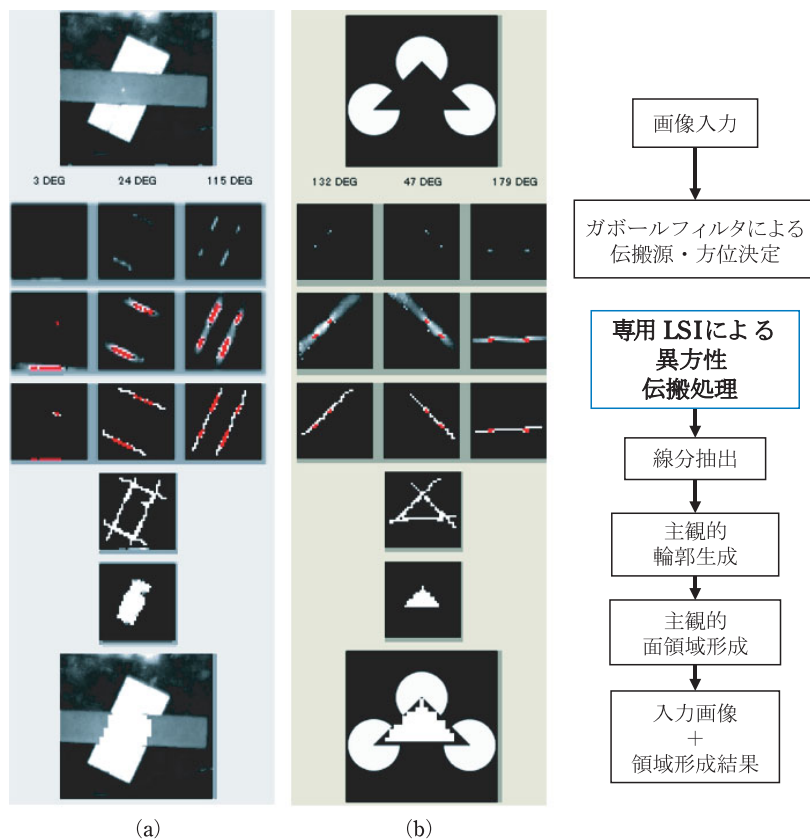


図4 主観的輪郭生成システムの実出力例 遮蔽輪郭生成（左）と、主観的輪郭生成（右）。カニツアの三角形（右上）を見ると、人は中央に三角形を知覚する。これを人工的に再現した。この機能はまた遮蔽物体の復元にも有効である。

車載用認識システムのために高性能なメディアプロセッサも各種開発されているが<sup>(9)</sup>、ナノメートル領域の集積回路はコストと消費電力増大の問題に直面している。一方で、インテリジェントカーや自律知能ロボットの視覚・頭脳を実現する知能処理用集積回路には、現状のコンピュータをはるかに超える演算性能が要求される。デジタル集積回路は脳の神経回路とは全く異なった方式・アーキテクチャで発展してきたが、精密な計算を必要としない脳型知能視覚処理については、脳の処理様式に学んでデジタル方式以外の方向性を研究していく価値があると考えている。

2000 年前後から、神経細胞のスパイクパルス発生メカニズムである非線形アナログダイナミクスを実現するアナログ集積回路及びスパイクパルス自身を扱う回路方式が活発に研究されるようになってきた。実用的には、神経回路の要点と思われるところのみを模倣して、大規模システムに適したモデル及び集積回路を開発することが重要である。その要点としては、スパイクパルスの同期性や非線形ダイナミクスが挙げられる。

更に、脳の視覚系のように、様々な画像処理機能を各部位で並列分散処理するアーキテクチャが考えられる。そこで、専用の各種処理ハードウェア (LSI チップ) を合体・システム化することで高度な視覚機能を実現する方向性を提案している<sup>(10)</sup>。

森江らは、このような観点に立って、既存のデジタル方式とは異なった情報処理方式を提案している<sup>(11)</sup>。それは、パルス幅やパルス位相に情報を持たせた時間軸情報処理である。この方式は、原理的に積和演算や非線形変換演算を低消費電力で実現することができ、また実際の神経回路で実行されている情報処理モデルを効率良く実現できる可能性がある。この方式に基づいて、脳の第一視覚野 (V1) の処理モデルとして知られているガボールフィルタや、V1-V2 野で行われていると考えられる主観的輪郭生成等を実現する異方性伝搬処理 LSI (図 4)、大局的画像分割・領域抽出を行う抵抗ヒューズネットワーク及び結合 MRF モデル回路等を設計・開発してきた。更に、スパイクタイミングを直接利用したガボールフィルタ回路モデルも提案している<sup>(12)</sup>。

これらの集積回路は、将来的には画素並列アーキテクチャという超並列性を生かして、イメージセンサと合体し、現在実用化が進んでいる TSV 等を用いた三次元集積技術を駆使して、高集積化・高性能化を図ることが可能と考えている (図 5)。

更に、脳型視覚処理モデル及び時間軸情報処理型集積回路は、既存のトップダウン型 LSI 製造技術だけでなく、自己組織化に代表されるボトムアッププロセス<sup>(13)</sup>によるナノ構造を用いた新機能デバイスの適用モデル・実現方式としても有効と考えている。これにより、超低消費電力・超高性能な脳型 (視覚) システムを構築でき

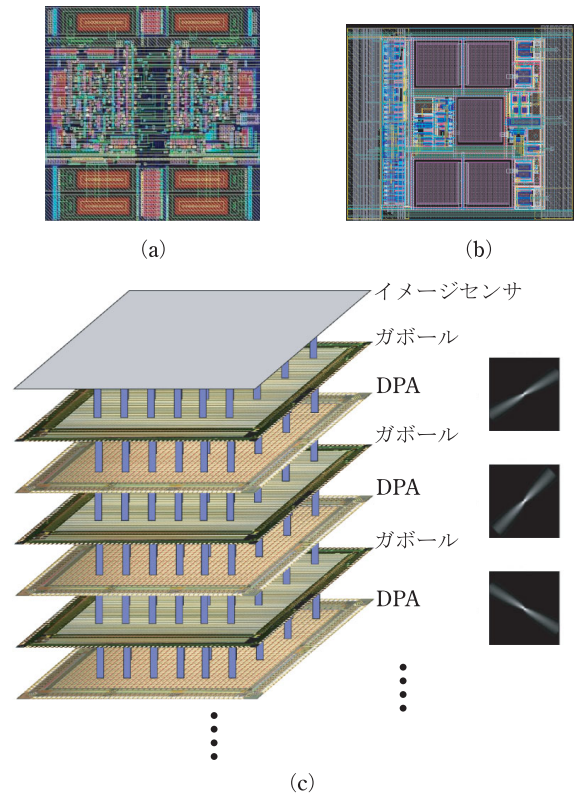


図 5 脳型視覚 LSI (a) ガボールフィルタ LSI 画素回路レイアウト, (b) 異方性伝搬処理 (DPA) LSI 画素回路レイアウト, (c) 脳型視覚を実現する三次元集積システムのイメージ図. 定形だが演算量の多い初期視覚機能は、画素並列方式による三次元集積システム化が有効である。

る可能性がある。その端緒として、スパイクニューラルネットワーク LSI の開発及びそれを実現するナノ構造の提案・開発<sup>(14)</sup>を行っている。

## 5. まとめと展望

車載用・自律知能ロボット用視覚のための画像処理手法を紹介し、既存のコンピュータビジョンの流れの一部が、脳の視覚処理様式に近づき、それを取り込みつつ発展する可能性を述べた。

また、より将来的な脳型視覚処理とそれを実行する LSI について述べた。人の脳に迫る視覚システムを実現するには、処理モデル・アルゴリズム研究だけでなく、時間軸情報処理型回路方式、及び三次元集積、ボトムアップ型ナノ構造作製などの製造技術を駆使して、総合的にアプローチしていく必要があると考えている。

工学の分野では、「空を飛ぶのに、鳥をそのままねずに、飛行機を考案した」という強力な成功体験がある。人の高度な視覚・思考系を工学的に実現するために、脳をそのままねずに、アルゴリズムのみを考案してデジタル計算機でシミュレーションするという手法が、飛行機と同じ成功に至るのかどうか、まだ誰にも分

からない。ただ、クオリア、意識といった人の知覚の本質に関わる問題は、単なるデジタル計算で実現できるとは到底思えないと感じるのは、筆者だけではないであろう。そのため、スパイクパルスによるダイナミクスとそのアナログ計算を実デバイスで実現する試みは、今後より重要な意味を持つてくるであろう。

**謝辞** 本研究の一部は、文部科学省知的クラスター創成事業（Ⅱ期）（現・地域イノベーションクラスタープログラム（グローバル型【第Ⅱ期】））、21世紀COEプログラム（#J19）、科研費（19300079, 22240022）及びJST/ERATO「複雑数値モデル」の支援を受けた。LSI設計は東京大学大規模集積システム設計教育研究センターを通し、日本ケイデンス株式会社、メンター株式会社の協力で行われた。

## 文 献

- (1) N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in IEEE Conf. on Computer Vision and Pattern Recognition, pp. 886-893, 2005.
- (2) タンジュークイ, 河野憲亮, 石川聖二, 金 亨燮, 四宮孝史, "モーションデータベースを用いた人の動作の高速認識," 画電学誌, vol. 36, no. 5, pp. 738-746, 2007.
- (3) Y. Nakashima, J.K. Tan, S. Ishikawa, and T. Morie, "Detecting a human body direction using a feature selection method," in Proc. of Int. Conf. on Control, Automation and Systems, no. FP07-1, pp. 1424-1427, 2010.
- (4) 三井相和, 山内悠嗣, 藤吉弘亘, "Joint 特徴量を用いた2段階 Boosting による物体検出," 信学論(D), vol. J92, no. 9, pp. 1591-1601, Sept. 2009.
- (5) D.G. Lowe, "Distinctive image features from scale-invariant keypoints," Int. J. Comput. Vis., vol. 60, no. 2, pp. 91-110, 2004.
- (6) M. Shimizu, I.R. Khan, Y. Kuriya, H. Miyamoto, and T. Morie, "Markerless arm posture estimation independent of environment," J. Signal Process., vol. 14, no. 6, pp. 475-481, 2010.
- (7) M. Riesenhuber and T. Poggio, "Hierarchical models of object recognition in cortex," Nat. Neurosci., vol. 2, no. 11, pp. 1019-1025, 1999.
- (8) M. Szarvas, A. Yoshizawa, M. Yamamoto, and J. Ogata, "Pedestrian

detection with convolutional neural networks," in Proc. IEEE Intell. Vehicle Symp., pp. 224-229, 2005.

- (9) S. Kyo, T. Koga, S. Okazaki, R. Uchida, S. Yoshimoto, and I. Kuroda, "A 51.2GOPS scalable video recognition processor for intelligent cruise control based on a linear array of 128 4-way VLIW processing elements," in IEEE Int. Solid-State Circuits Conf. (ISSCC), no. 2.6, pp. 48-49, 2003.
- (10) T. Morie and Y. Kim, "A subjective-contour generation LSI system with expandable pixel-parallel architecture for vision systems," in IEEE Int. Solid-State Circuits Conf. (ISSCC), no. 28.6, pp. 478-479, 2009.
- (11) 森江 隆, "アナログデジタル融合方式 LSI 技術による脳型視覚システム," 信学 FR 誌, vol. 1, no. 1, pp. 19-29, July 2007.
- (12) 山口雄一郎, 田中秀樹, 松坂建治, 森江 隆, "スパイク駆動方式ガボールフィルタ回路モデル," 日本神経回路学会第 19 回全国大会 (JNNS2009), no. P3-20, pp. 178-179, 2009.
- (13) C.H. Huang, M. Igarashi, M. Wone, Y. Uraoka, T. Fuyuki, M. Takeguchi, I. Yamashita, and S. Samukawa, "Two-dimensional Si-nanodisk array fabricated using bio-nano-process and neutral beam etching for realistic quantum effect devices," Jpn. J. Appl. Phys., vol. 48, 04C187, 2009.
- (14) M. Igarashi, C.H. Huang, T. Morie, and S. Samukawa, "Control of electron transport in two-dimensional array of Si nanodisks for spiking neuron device," Appl. Phys. Express, vol. 3, 085202, 2010.

(平成 22 年 12 月 31 日受付 平成 23 年 2 月 8 日最終受付)



もりえ たかし  
森江 隆 (正員)

昭 54 阪大・理・物理卒。昭 56 同大学院修士課程了。同年電信電話公社 (現 NTT) 入社。半導体プロセス・集積回路設計の研究に従事。平 9 広島大・工・助教授。平 14 九工大教授。画像処理、脳型集積システム・デバイスの研究に従事。博士 (工学)。



いしかわ せいじ  
石川 聖二 (正員)

昭 49 東大・工・計数卒。昭 55 同大学院博士課程了。工博。昭 54 九工大講師。平 6 同教授 (機械知能工学科)。平 8 ユトレヒト大客員教授。三次元画像計測、医学画像処理等の研究に従事。バイオメディカル・ファジィ・システム学会平 20 The Best Paper Award, 同年和文誌優秀論文賞各受賞。