3

イメージセンシング技術の新しい応用

3-3 遊びの目

～インテリジェントイメージセンサのアミューズメント分野への応用～

田中 健一[†], 正会員 鹿毛 裕史^{††},倉持 宗一郎^{††}, 正会員 久間 和生^{††}

キーワード

Artificial Retina, Image Sensor, Optical Flow, On-chip Image Processing, Amusement

1. ま え が き

半導体の技術革新に伴い、ゲーム機やおもちゃの電子化が進み、また、マイコンを搭載した製品の処理能力が向上するに連れて、これまで専用の画像処理ボードやASIC、DSPなどの独壇場であった画像認識システムを、汎用マイコンを搭載した製品へも採用することが可能になってきた。とりわけ、コストの制約が厳しいアミューズメント分野へ画像認識技術を導入するためには、処理速度、サイズ、消費電力、コストなど、従来の画像認識システムの構成をそのままの形で小型化するだけでは解決できない問題が多くあり、何らかの工夫が必要であった。

2. 画像処理機能を持つイメージセンサ

画像認識システムの演算量は、1画素当たりの演算量が増えるほど、また、処理すべき画素数が増えるほど、多くなる。一般的に、後者の割合が全体の演算量に与える影響が大きいため、イメージセンサの画素数を削減することはシステム性能に与えるインパクトが大きい。低画素数のイメージセンサでデジタルカメラのような高精細の画像は望むべくもないが、画像中の物体の動きを検出するようなセンサ機能を実現することは可能である。このような目的で開発された32×32画素の人工網膜LSI¹⁾は、低画素数という特徴に加えて、演算量の多い画像の前処理をオンチップの画像処理機能で分担するという人工網膜の設計思想²⁾に基づいたインテリジェントなイメージセンサで、これを使うことにより、小型で安価な画像認識システムを実現することができる。特に、動き検出用途を主たる応用ターゲットとし、高いフレームレートと画像を圧縮して後段の処理プロセッサへ送ることができる構造に特徴がある。センサの主なスペックを表1に示す。以下に本センサが持つオンチップ画像処理機能の概要を紹介する。

解像度	32×32
画素サイズ	56μm×56μm
画素エリア	1.8mm×1.8mm (1/6インチ光学系)
チップサイズ	3.8mm×3.1mm
電源電圧	5.0V
消費電力	15mW (Typ.)
最大フレームレート (@Clk=2MHz)	350FPS (2次元画像) 5,000FPS (射影データ)
蓄積時間設定	16μs～1s (@Clk=2MHz)
ゲイン設定	0dB～34.5dB (Step=1.5dB)
飽和出力電圧	2.0V
感 度	20V/lx・s
ダイナミックレンジ	66dB
FPNレベル	2mV
機 能	2次元画像データ出力 射影データ出力 全画素平均値データ出力 輪郭抽出 (2次元画像/射影画像)

表1 32×32画素人工網膜LSIの仕様

(1) 2次元画像データ出力機能

縦方向に32画素、横方向に32画素の2次元アレイ状に画素が並んだイメージセンサであり、通常のイメージセンサとしての2次元画像の撮像機能を持つが、動き検出に特化するため、高速フレームレートでの画像撮像が可能となっている。最大フレームレートは350fpsである。

(2) 全画素平均値データ出力機能

全画素の出力値の平均に相当する値をセンサ内で演算し、出力することができる。画像を撮像する際に必要となる、自動露出機能などへの応用を想定したものである。

(3) 射影データ出力機能

センサ内の演算機能を用いて縦・横方向の射影データ (各画素の輝度値を一定方向に加算して得られる総和) を出力することができる。この変換により、32×32=1,024画素の画像データを32+32=64データに圧縮することができ、8ビットマイコンなど小規模な処理回路との組合せでも、高速な画像処理システムの構築が可能になる (最大5,000fps)。

(4) 9分割射影データ出力機能

画像中央部の30×30画素の領域を“井”の字形に9ブロックに分割し、その中から中央と4隅の合計5ブロックの縦・

†三菱電機株式会社 先端技術総合研究所

††三菱電機株式会社 システムLSI事業統括部 AR事業推進プロジェクトグループ

"Intelligent Image Sensor Utilization for Amusement Application" by Ken-ichi Tanaka (Advanced Technology R&D Center, Mitsubishi Electric Corp., Tokyo), Hiroshi Kage, Souichirou Kuramochi and Kazuo Kyuma (Business Promotion Project of Artificial Retinas, Mitsubishi Electric Corp., Tokyo)

横方向射影を独立に出力する機能を実装した。これにより、通常の全画面に対する縦・横方向射影のみでは困難であった、画像面内での回転、拡大/縮小方向の動きを検出することが可能となった(最大約3,000fps)。

(5) 空間微分・時間微分データ出力機能

動き検出アルゴリズムの一つであるオプティカルフロー計算に必要な空間微分・時間微分データの演算回路をチップ内に実装した。

これまでに述べた32×32画素の人工網膜LSIは動き検出など、機械による処理を前提としたセンサ用途に特化したイメージセンサであるが、人工網膜LSIのその他の品種として、より画素数の多いQCIF³⁾やCIFサイズ⁴⁾のものも開発されている。これらのチップでは「絵を見る」用途が主となるため、人工網膜側で行う画像処理としては輪郭強調・色調補正などの、人が見る「画像」を作るためのオンチップ画像処理機能に力を入れた設計となっている。

3. 画像処理アルゴリズム

パソコンやゲーム機など、ユーザフレンドリーなインタフェースを実現するためには、ユーザの操作による動きをリアルタイムで検出して、入力情報へ変換する必要がある。人工網膜LSIの持つオンチップ画像処理機能を利用すれば、高速処理が要求されるこれらのパソコンやゲーム機などのシステムへの画像認識機能の導入が、低コストで可能となる。前章で紹介した人工網膜LSIの持つオンチップ画像処

理機能の中から、本章では全画素射影出力および9分割射影機能を用いた画像認識アルゴリズムについて述べ、その有用性を明らかにする。

図1(a)は、画像全体に含まれるグローバルな動きを検出するために、全画素の縦方向および横方向の射影出力データを用いることを示す図である。(1)は入力画像、(2)と(3)はそれぞれ(1)の画像に対応する、縦方向に並ぶ画素および横方向に並ぶ画素の輝度値を加算した結果であり、これをそれぞれ縦方向射影、横方向射影と呼ぶ。現フレームと次フレームの間にT字形の模様が右下方向へ移動したとすると、(2)の射影データの形は右方向へ、(3)の射影データの形は下方向へ平行移動するので、これらの射影データから縦横それぞれの方向への1次元的な動きを検出することができる。すなわち、(2)の縦方向射影からは横方向の動きを、(3)の横方向射影からは縦方向の動きを検出することができる。さらにこれらを組合せて2次元ベクトルを構成すれば、T字形の模様が右下へ向かって移動しているという、2次元面内での並進的な動きを検出することが可能である。このような原理に基づいて画像から高速に2次元的な動き(オプティカルフロー)の検出を行うアルゴリズムを開発し、これを用いたジェスチャ入力インタフェース⁵⁾を開発した(後述のダンスゲームシステムはその実用例)。本例での射影出力データを利用した動き検出アルゴリズム⁶⁾は、前章でも述べたように N^2 個のデータを $2N$ 個に削減することができるため、処理データ数の観点から格段に少ない計算量での実行が可能である。これは入力インタフェースのような高速処理が要求される事例では特に重要なポイントである。実用面でいえば、本技術はマウス等の情報機器インタフェースや、あるいは近年積極的に各メーカーによって携帯電話への搭載が検討されているモバイルカメラにおいても、ボタンやキー入力以外の入力インタフェースとして利用することができる。

次に、もう1つの射影出力機能である9分割射影機能について図1(b)を用いて説明する。(1)は入力画像(32×32画素)の例であり、点線枠は、画素エリアを縦横それぞれ3分割してできる9個の部分画像領域について、周辺4個と中央1個の合計5個の領域について、図1(a)で示したものと同様の射影データがこれら5個の領域すべてについて出力される。(2)は周辺4領域からの射影出力、(3)は中央の1領域からの射影出力である。これらの射影出力データから、対応する部分画像領域毎に、図1(a)で示したものと同様の手続きで動きベクトルを検出することができる。

さらに、これら5個の動きベクトルを統合して、カメラが3次元空間内を移動する際に検出される3種類の動き、すなわち、(1)撮像面に並行な2次元平面上での並進、(2)視軸を中心とした回転、(3)奥行き方向に動く際に現れる拡大あるいは縮小、の検出が可能となる(図1(c))。本章で紹介した9分割射影出力を利用することで、マウス等で必要

な2次元平面内の並進運動だけではなく、回転や奥行き方向を含む3次元空間での動きを検出できるので、フライトシミュレータ等のゲームインタフェースや、あるいは3次元空間を進む自律走行車をナビゲートするための制御入力としても適用が可能であり、幅広い応用が期待できる。

4. アミューズメントへの応用

本章では、第2章と第3章で紹介した人工網膜LSIおよび画像認識アルゴリズムのアミューズメント分野における実用化事例について紹介する。1つは低画素数のセンサを用いてボディアクションを検出し、ゲーム入力へ利用する例、もう1つは携帯電話のモバイルカメラを用いた似顔絵生成・表情変形の例である。

4.1 ジェスチャ入力型インタラクティブゲーム

図2(27頁に掲載)に示すアーケードゲームシステムは、(株)ナムコ殿のダンスゲーム「恋のバラバラ大作戦」⁷⁾のシステム構成図である。このシステムには32×32画素の人工網膜LSIとレンズを一体化した人工網膜センサモジュールが搭載されており、ゲームプレーヤのアクションをオプティカルフロー(各画素における動きベクトル)として検出し、この情報を基に身体各部の動きを解析する。ゲーム画面に現れるインストラクタのダンス動作とプレーヤの動きとの一致度を照合することで、ゲームを進行させる(ダンスが上手であればスコアが上がり、下手であればゲーム終了等)ことができるシステムである。本システムでは、赤外線センサ等の補助センサは一切使用しておらず、プレーヤの動きをすべて画像による動き解析で実現しており、システムの小規模化にも貢献している。

同様の技術は、家庭用ゲーム機のゲームソフト⁸⁾としても商品化されている。これは家庭用ゲーム機への初の実用化例であり、多くの消費者にこれまでのジョイスティックやキーボードとは違った、画像認識技術による新しいゲームインタフェースを提供することができる。

4.2 似顔絵生成と表情変形

近年、携帯電話への搭載が相次いで発表されるイメージセンサの中で、撮像だけではなく画像処理機能を活用した事例として「似顔絵生成技術」⁹⁾について紹介する(図3(27頁に掲載))。携帯電話のユーザがカメラを利用する際にもっとも多い被写体として想定されるのが顔である。撮像した顔を相手に送る際に遊びのセンスを付け加えることに着目し、人工網膜LSIで撮像した顔画像から自動的に顔領域を切出し、さらにデータ量削減のためにこれを二値化し、肌色で色づけした線画似顔絵を生成する技術を開発した。これに加えて、目や口等の顔パーツを検出し、これらの位置情報を利用してパーツの入れ替え(漫画似顔絵)や、泣き笑い等の表情変形を施す技術も併せて開発した。

本技術の特徴は、①動き解析技術による顔領域切出し、②人工網膜LSI利用による線画化の高速実行、の2点である。携帯電話ユーザがこれを利用する際、単に顔画像をメールに添付するだけではなく、メールの内容に沿って、例えば、楽しいメールには笑った似顔絵を添付する等、付加価値の高い利用が期待できる。本技術を搭載した携帯電話はすでに市場へリリースされており、特にここ数年マーケット規模が急速に拡大する携帯電話分野へ画像認識技術が採用された初めての例でもあり、今後はアミューズメントなど、他の分野への展開も期待できる。

5. む す び

これまで産業分野での利用が中心であった画像認識技術を、アミューズメント分野へ適用したいいくつかの事例を紹介した。これを可能にした技術的なポイントはオンチップ

画像処理機能を有するハードウェアとその処理機能をサポートするアルゴリズム(ソフトウェア)の組合せであった。より高度な機能を実現するための画像認識システムの開発が今後も進められると予想されるが、その際には、ハードウェアとソフトウェアのバランスがとれたシステムソリューションを考えることが重要だと考える。

(2001年11月05日受付)

〔文 献〕

- 1) 船津, 倉持他: “射影演算機能を持つ人工網膜LSI”, 映像学誌, 55, 10, pp.1329-1336 (2001)
- 2) K.Kyuma, E.Lange, J.Ohta, A.Hermanns, B.Banish and M.Oita: “Artificial Retinas - Fast, Versatile Image Processors”, Nature, 372, 6502, pp.197-198 (1994)
- 3) <http://www.melco.co.jp/oldnews/1999/0713-b.htm>
- 4) <http://www.melco.co.jp/news/2001/0322-b.htm>
- 5) H. Kage, W. T. Freeman, Y. Miyake, E. Funatsu, K. Tanaka and K. Kyuma: “Artificial Retina Chips as On-Chip Image Processors and Gesture-Oriented Interfaces”, Optical Engineering, 38, 12, pp.1979-1988 (1999)
- 6) W. T. Freeman, D. B. Anderson, P. A. Beardsley, C. N. Dodge, M. Roth, C. D. Weissman, W.S.Yerazunis, H. Kage, K. Kyuma, Y. Miyake, and K. Tanaka: “Computer Vision for Interactive Computer Graphics”, IEEE Computer Graphics and Applications, 18, 3, pp.42-53 (1998)
- 7) <http://www.namco.co.jp/home/pr/2000/jul/press-04.html>
- 8) <http://www.konamityo.com/cgi-bin/police/intro/intro.cgi?returl=/police/index.html>
- 9) <http://www.melco.co.jp/news/2001/0424.htm>

田中 健一 1981年、京都大学大学院工学研究科(航空工学専攻)修了。同年、三菱電機(株)中央研究所に入社。宇宙機用電気推進システム、宇宙ロボット、ニューラルネットワークの研究開発、および人工網膜LSIの事業化に従事。現在、同社先端技術総合研究所に所属。1998年、映像情報メディア学会技術振興賞受賞。

鹿毛 裕史 1990年、京都大学大学院情報工学修士課程修了。同年、三菱電機(株)入社。以来、先端技術総合研究所において人工網膜LSIを利用した画像認識アルゴリズムの研究・開発に従事。1998年4月より、同社AR事業推進プロジェクトグループに所属。正会員。

倉持宗一郎 1995年、東京大学大学院計数工学専攻修士課程修了。同年、三菱電機(株)に入社。以来、先端技術総合研究所において人工網膜LSIの研究・開発に従事。1998年4月より、同社・AR事業推進プロジェクトグループに所属。

久間 和生 1977年、東京工業大学大学院理工学研究科博士課程修了。現在、三菱電機(株)システムLSI事業統括部AR事業推進プロジェクト部長。東京工業大学客員教授などを兼務。専門は光エレクトロニクスとニューラルネットワーク。IEEEとOSAのフェロー。映像情報メディア学会開発賞、大河内記念技術賞などを受賞。正会員。