

平成21年 3月 31日現在

研究種目：基盤研究（C）
 研究期間：2007～2008
 課題番号：19500153
 研究課題名（和文） リアルタイム・ジェスチャ認識集積回路に関する研究
 研究課題名（英文） Research on real-time gesture recognition system
 研究代表者
 肥川 宏臣（HIKAWA HIROOMI）
 関西大学・システム理工学部・教授
 研究者番号：10244154

研究成果の概要：

画像として取り込んだ手の形を分類するハードウェア指文字認識システムについて、回路簡略化のためにアルゴリズムの改良、回路設計、機能シミュレーションによる動作確認を行った。さらに、指の動きを伴うジェスチャ認識の基礎実験、自己組織化マップ（SOM）をベースとしたハイブリッドネットワークによる認識率向上、画像の大きさの違いによる認識率の悪化を防ぐために画像の正規化手法に関する研究を行い、それぞれ有効性を確認した。

交付額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	1,200,000	360,000	1,560,000
2008年度	1,300,000	390,000	1,690,000
年度			
年度			
年度			
総計	2,500,000	750,000	3,250,000

研究分野：総合領域

科研費の分科・細目：知覚情報処理・知能ロボティクス

キーワード：画像情報処理

1. 研究開始当初の背景

ハンドジェスチャは、従来のぎこちない human computer interaction (HCI) に対して、より有効なインタフェースとして期待されている。ハンドジェスチャを適切に認識することで、人間同士が普通に用いている自然な動作をHCIとして使うことができるようになることを期待できる。

2. 研究の目的

本研究の目的は、手の身振りによるジェスチャを直接認識するハードウェアの設計を行い、FPGAに実装、その有用性を明らかにすることである。なお、FPGAはField Programmable Gate Arrayの略で大規模デジタル回路をユーザがプログラムできる集積回路である。また、本研究の最終的な目的をジェスチャ認識回路の集積化としているが、これにより、小型化が可能となるとともに、低消費電力な

システムとすることができる。現在、研究されているジェスチャ認識システムは、ソフトウェアによるものがほとんどである。ジェスチャ認識は大量の画像データを瞬時に計算しなければならず、かなりのコンピュータの能力を消費してしまう。そのため、本来は、より大きなシステムの中のヒューマンインタフェースを行うサブシステムとなるべきジェスチャ認識がコンピュータに大きな負担となる。これに対して、提案手法ではジェスチャ認識システムを集積化することで、低消費電力で小型なサブシステムとすることができるため、コンピュータの能力を本来のアプリケーション機能に集中させることが可能となる。

3. 研究の方法

まず、手の認識システムのハードウェア化を行う。この段階では手の形状認識のみを行うポストチャ認識回路となる。このポストチャ認識回路に動画像から切り出した静止画を入力することで、時系列的にポストチャ変化シーケンスが得られることになる。

さらに、このシーケンスからどのようなジェスチャかを判断する形状変化を認識する専用回路を開発する。これらのポストチャ認識回路と手の形状変化認識回路を融合することで、手の身振りによるジェスチャを認識する専用ハードウェアを実現する。

4. 研究成果

【アルゴリズム】

手の形を認識するシステムに実装する認識アルゴリズムを図1に示す。前処理において求めた入力イメージの水平／垂直方向のヒストグラム振幅スペクトルを特徴ベクトルとして用いる。そして、この特徴ベクトルが所属するクラスを分類ネットワークにより検索することで対応する指文字の認識を行う。

振幅スペクトルを用いることで、扱うベクトルの次元数を大幅に削減することができるように、入力画像内の手の位置ずれの影響を無くすることができるという特徴を持つ。しかし、このシステムでは振幅スペクトルの変換に discrete Fourier transform (DFT) を用いている。しかし、DFTの計算には正弦・余弦関数の生成と多数の乗算を必要とするため、ハードウェア化には適していない。この問題に対応するため、正弦、余弦関数を3値化した簡易DFTを用いる。

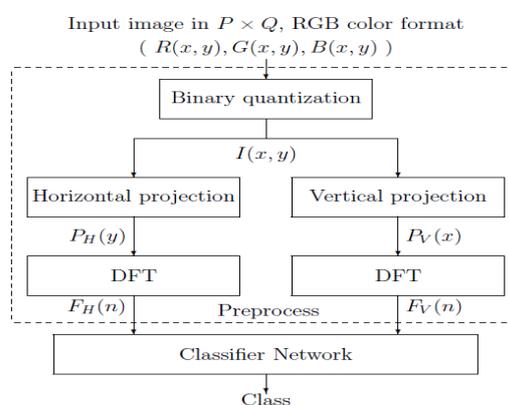


図1 手の形認識アルゴリズム

分類ネットワークは入力された文字の形を認識する。ここで認識する手の形をクラスと呼ぶ。認識は、あらかじめ求めておいた各クラスを代表するベクトル（参照ベクトル）と入力ベクトルの距離を求め、最少距離となる参照ベクトルを持つクラスを認識結果とする。本アルゴリズムで使用する分類ネットワークでは、回路化が容易になるよう、簡易ベクトル距離評価方法を用いる。この方法では、クラス s に属する特徴ベクトルの各要素があらかじめ設定した「範囲」内にあるかどうかによる距離評価を行う。通常のベクトル距離計算に必要な減算、乗算、平方根などの計算が不要でハードウェア化に適している。

【シミュレーション】

この認識アルゴリズムの有効性を調べるために日本語の指文字のうち動きを伴わない41種類の指文字の認識実験を行った。まず、図2に特徴ベクトルの次元数 D と認識率の関係を示す。次元数が22次元で最良の認識率が得られた。入力画像が 128×128次元なので、大幅な次元数の削減が可能であることが示された。

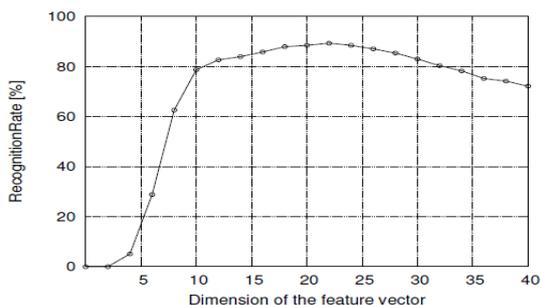


図2 ベクトル次元数と認識率の関係

簡易化したDFT の使用によりシステムの認識率の低下、特に本システムの特徴の1つである手の位置に対するロバスト性の悪化が懸念される。そこで、入力画像と学習画像における手の位置ずれに対するロバスト性を調べた。同じ手の画像の位置だけを変えた画像データを各クラス100枚ずつ用意した。データは3グループから構成される。手の位置を画像の左上のランダムな位置としたLTグループと右下に置いたRB グループ、LTとRTデータを同じ割合で含むMIXの3種類である。そして、通常のDFTと簡易DFTを用いた認識システムの両方の認識率を求めた。実験はLT, RB, MIXの3種類の学習データを用いて3度学習し、平均認識率を求めた。テストデータは、LTとRB両方のデータを含むが、RBグループの割合を変えて認識実験を行った。RBグループの割合に対する認識率を図3 に示す。図3(A)に示すように、従来のDFTを用いたシステムでは、学習データの種類に関係なく平均認識率は100%となった。

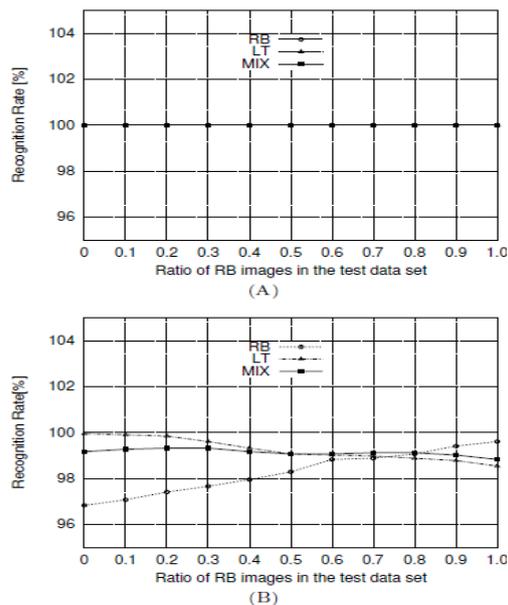


図3 位置の違いと認識率の関係、
(A) 通常のDFT を用いたシステム、
(B) 簡易DFT を用いたシステム。

そして、提案するDFTを用いたシステムの平均認識率は図3(B)のようになった。認識率は、手の位置について学習データとテストデータ間に相関があり、従来のDFTを使ったシステムに比べ、ロバスト性が若干悪化している。RBデータで学習した場合、テストデータに含まれるRBデータの割合が大きくなるほど認識率が良くなっている。逆に、LTデータで学習した場合、RBデータの割合が小さくなるほど認識率が良くなる。これらの場合、最悪の平均認識率は97%となった。しかし、MIXデータを用いて学習した場合、認識率の低下は1%に抑えられていることがわかる。

次に、同一クラスの画像でも位置や形が若干異なる、41種類の指文字カラー画像データを使って、認識実験を行った。その結果、従来のDFTを用いた場合の認識率は93.61%、簡易DFTを用いた場合92.49%の認識率となった。

【ハードウェア設計】

以上から、認識アルゴリズムの有効性が確

認められたため、ハードウェア設計を行い回路規模の評価を行った。全体のブロック図を図4に示す。入力信号はピクセルのRGBカラー信号とそのx, y座標である。まず、2値化回路により背景画像が取り除かれ、赤い手袋をした手の部分のみが2値化される。次のメモリと加算回路により、水平・垂直ヒストグラムが求められる。TDFT22は簡易DFTにより特徴ベクトルを求める回路である。特徴ベクトルは次の分類ネットワークに送られる。

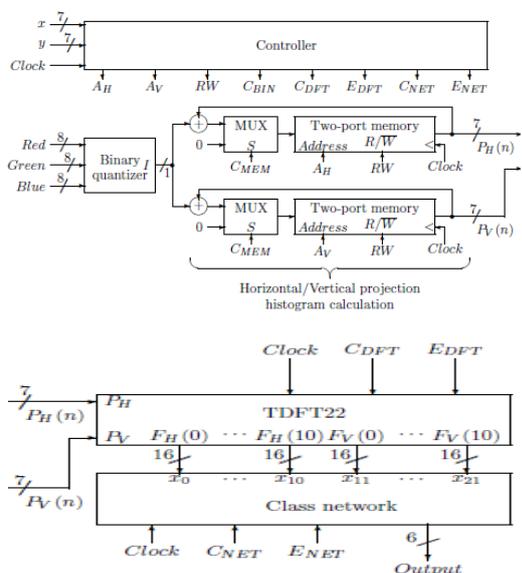


図4 手の形認識ハードウェアの構成

図5に分類ネットワークの構成を示す。ここでは2種類の分類ネットワークを提案している。図5(A)はすべてのクラスの評価値を同時に求める並列処理による分類ネットワーク、図5(B)は順番に評価値を求める直列処理による分類ネットワークである。並列型は処理が速いが、直列型は回路規模が小さくなるという特徴がある。

図4に示すシステムをハードウェア記述言語であるVHDLで記述し、機能シミュレーションによる動作確認後、論理合成を行った。その結果、図5(A)の並列分類ネットワークを

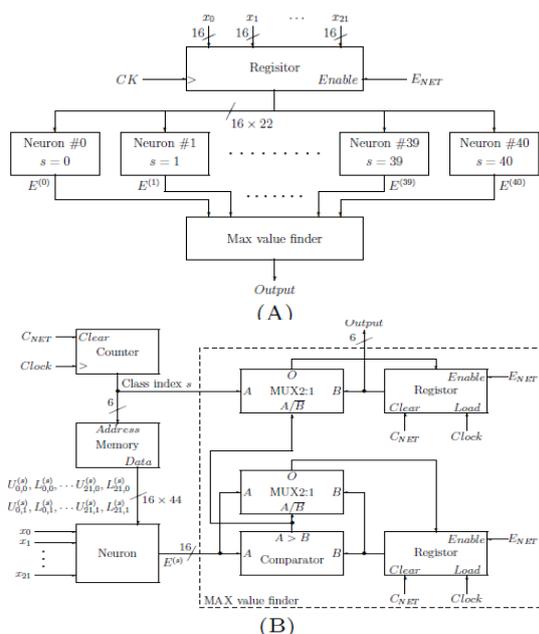


図5 分類ネットワーク, (A) 並列ネットワーク, (B) 直列ネットワーク

用いた場合 282,341 ゲート相当, 図5(B)の直列ネットワークを用いた場合 89,695 ゲート相当となり、直列ネットワークを用いることで大幅な回路規模の縮小が可能であることがわかった。また、動作速度としては、並列型が 24.5MHz, 直列型の場合 52MHz のクロック周波数での動作が可能となった。並列型の場合 1 入力につき 1 クロックでの認識が可能であるのに対して、直列型では 41 クロック必要なため、認識時間は並列ネットワークのほうが短くなる。

【ジェスチャ認識への応用】

動的な手の動きをカメラで撮影すると手の形が順次変化したものとして捉えることができる。一定の時間間隔でジェスチャ動画をサンプリングした静止画を本研究で提案するシステムに入力することで、手の形の変化を表す時系列データが得られる。これを認識することでジェスチャ認識を行う。図6に示す時系列変化を認識する JORDAN型ニューラルネットワークを用いた実験を行った。

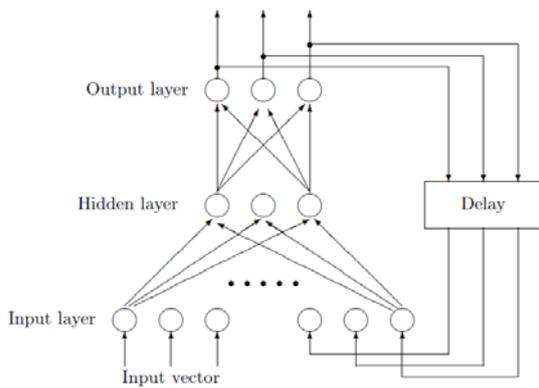


図6 Jordan型ニューラルネットワーク

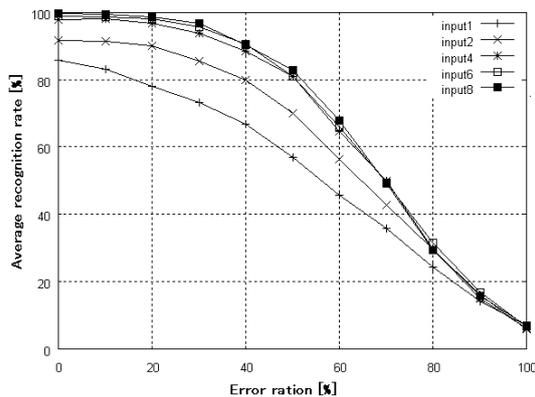


図7 ジェスチャ認識特性

図7に12種類のジェスチャに対する認識率と静止画認識のエラー率，ジェスチャ速度の関係を示す。エラー率が50%でも速度が遅ければ80%以上の認識率が得られることが示されている。

【SOMとHebb学習ネットワークによる分類ネットワーク】

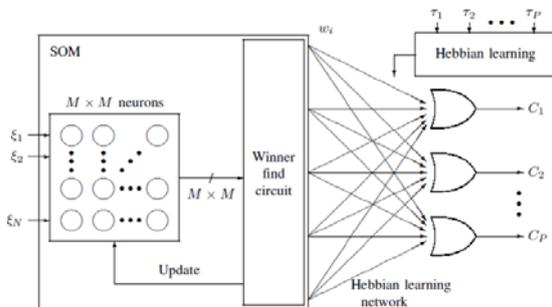


図8 自己組織化マップとHebb学習ネットワークによる分類ネットワーク

図8に示す分類ネットワークによる認識率の改善を行った。その結果，ニューロン数12×12個のSOMを用いた場合，認識率を97%以上に向上させることができることを確認した。

【大きさ補正】

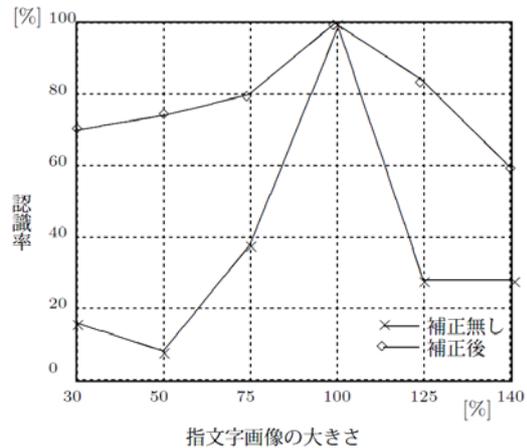


図9 大きさ補正の効果

学習した画像と新たに入力された画像の大きさが異なると学習能力が極端に悪化してしまう。これに対応するために画像の正規化を行う手法を導入した。その結果を図9に示す。学習に用いた手の大きさを100%として入力する手の画像の大きさを変えた場合の認識率の変化を示している。正規化を行うことで，手画像の大きさ変化に対するロバスト性が向上していることがわかる。

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 9 件)

① 肥川 宏臣, 藤村 紘匡, 佐藤 大輔, ”ハードウェア向き指文字認識アルゴリズム”, 電子情報通信学会論文誌 D Vol. J92-D, No. 3, pp. 405-416, 2009. 3 査読あり

② H. Hikawa, H. Fujimura, ”Hardware design of Japanese Hand Sign Recognition System,” ICONIP 2008, Revised Selected Papers, LNCS 採録済み 査読あり

③H. Hikawa, H. Fujimura, “Simplified DFT for Hand Posture Recognition System,” Proceedings of 2008 International Symposium on Nonlinear Theory and its Applications (NOLTA 2008), pp.112-115, 2008.9 査読あり

④ H. Fujimura, Y. Sakai, H. Hikawa, “Japanese Hand Sign Recognition System,” Neural Information Processing, 14th International Conference, ICONIP 2007, Kitakyushu, Japan, November 13-16, 2007, Revised Selected Papers, Part I, LNCS 4984, p. 983 ff. 査読あり

⑤H. Hikawa, K. Kugimiya, “A New Hardware Friendly Vector Distance Evaluation Function for Vector Classifiers,” Neural Information Processing, 14th Int. Conf., ICONIP 2007, Kitakyushu, Japan, November 13-16, 2007, Revised Selected Papers, Part II, LNCS 4985, p. 137 ff. 査読あり

⑥H. Hikawa, K. Harada, T. Hirabayashi, “Hardware Feedback Self-Organizing Map and Its Application to Mobile Robot Location Identification,” Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol.11, No.8, pp.937-945. 2007.10 査読あり

⑦H. Hikawa, T. Miyanishi, K. Tamaya, “Performance Comparison of SOM Based Hybrid Hardware Classifiers,” Proc. of 2007 International Joint Conference on Neural Networks (INNS-IEEE IJCNN 2007), 2007.8 査読あり

⑧H. Hikawa, S. Matsubara, “Pseudo RBF Network for Position Independent Hand Posture Recognition System,” Proceedings of 2007 International Joint Conference on Neural Networks (INNS-IEEE IJCNN 2007), 2007.8 査読あり

⑨松原重喜, 肥川宏臣, “ハードウェア実装向きデータ分類アルゴリズム,” 電子情報通信学会論文誌, Vol. J90-A, No.8, pp.646-654. 2007.8 査読あり

〔学会発表〕(計 7 件)

①佐藤大輔, 肥川宏臣, “大きさの変化にロバストな指文字認識システム,” 電子情報通信学会 機能集積システム研究会, FIIS-09-250, 2009.3.6 (長崎美術館)

②小野尾 彰, 肥川宏臣, “自己組織化マップハードウェアの自動生成について,” 電子情報通信学会 ニューロコンピューティング研究会, NC2008-65, 2008.11 (佐賀大学)

③釘宮香織, 今村仁美, 肥川宏臣, 二村祥一, “全方位カメラを用いた位置認識システムの特性評価,” 電気関係学会九州支部第 61 回連合大会, 02-1P-12, 2008.9 (大分大学)

④藤村紘匡, 肥川宏臣, “位置のずれにロバストな指文字認識システム,” 電子情報通信学会 機能集積システム研究会, FIIS-07-222, 2007.10.26 (兵庫県立大学)

⑤釘宮香織, 肥川宏臣, “レンジチェック関数を用いたベクトル要素距離の重み付けに基づく距離計算手法,” 電気関係学会九州支部第 60 回連合大会, 05-1P-03, 2007.9.18 (琉球大学)

⑥藤村紘匡, 肥川宏臣, “レンジチェック回路を用いた指文字認識システム,” 電気関係学会九州支部第 60 回連合大会, 05-1P-02, 2007.9.18 (琉球大学)

⑦釘宮香織, 藤村紘匡, 肥川宏臣, “ベクトル要素距離の重み付けに基づく類似検索手法,” 電子情報通信学会 機能集積システム研究会, FIIS-07-208, 2007.6.22 (大分大学)

6. 研究組織

(1) 研究代表者

肥川 宏臣 (HIKAWA HIROOMI)
関西大学・システム理工学部・教授
研究者番号 10244154