

[招待論文] 脳機能に学ぶ画像認識集積システム

森江 隆[†] 岩田 穆^{††}

[†]九州工業大学大学院生命体工学研究科 〒808-0196 北九州市若松区ひびきの2-4

^{††}広島大学大学院先端物質科学研究科 〒739-8526 東広島市鏡山1-3-1

E-mail: [†]morie@brain.kyutech.ac.jp, ^{††}iwa@dsl.hiroshima-u.ac.jp

あらまし 脳の初期視覚系での特徴抽出モデルとして知られるガボールウェーブレット変換を画素並列動作で実現するLSIを核とする自然画像認識システム構築の試みについて述べる。まず、自然画像の特定領域を注視する処理として抵抗ヒューズを用いた大局的領域分割を行う。分割された各領域を個別に抽出し、ガボールウェーブレット変換を行う。得られたガボール特徴量を用いて、ダイナミックリンクアーキテクチャに基づく柔軟なマッチングにより認識を行う。特徴抽出までの処理をハードウェアで実行するために、画素並列で動作するLSIをパルス変調信号を用いたアナログ・デジタル(AD)融合回路アーキテクチャに基づいて設計した。本稿では特に、(1)パルス変調方式ピクセル回路を用いた大局的領域分割用抵抗ヒューズネットワークおよび抵抗ネットワーク型ガボールフィルタ回路とそのLSI設計例、(2)セルオートマトン型画像領域抽出アルゴリズムとそのFPGAへの実装例、(3)システム化の基盤となるLSI制御用FPGA搭載PCIボードの仕様について詳述する。

キーワード 自然画像認識, 顔認識, 領域分割・抽出, パルス変調方式LSI, 抵抗ヒューズネットワーク, ガボールウェーブレット変換, セルオートマトン, ダイナミックリンクアーキテクチャ

An Integrated System Mimicking Brain Functions for Image Recognition

Takashi MORIE[†] and Atsushi IWATA^{††}

[†] Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology

2-4, Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196 Japan

^{††} Graduate School of Advanced Sciences of Matter, Hiroshima University

1-3-1, Kagamiyama, Higashi-Hiroshima, 739-8526 Japan

E-mail: [†]morie@brain.kyutech.ac.jp, ^{††}iwa@dsl.hiroshima-u.ac.jp

Abstract This paper describes our ongoing research to construct a natural scene image recognition system using pixel-parallel VLSIs, one of which performs Gabor wavelet transformation known as a feature extraction model in the primary visual cortex of the brain. In the first step of the recognition process, coarse region segmentation is performed by a resistive-fuse network to pay attention to a specific region of the image. Each region is separately extracted and applied to feature extraction processing by Gabor wavelet transforms. Using the obtained Gabor features, flexible graph matching based on the dynamic-link architecture is performed. For hardware implementation of the feature extraction, we have designed pixel-parallel VLSIs based on the merged analog-digital circuit architecture using pulse-modulation signals. This paper describes in detail (1) a resistive-fuse network circuit for coarse region segmentation and a resistive network circuit for Gabor-type filtering using a pulse-modulation pixel circuit, and the results of their LSI design, (2) a cellular-automaton-type region extraction algorithm and its implementation using an FPGA, and (3) an FPGA board with a PCI-bus interface, which is used for controlling the image processing LSIs and for the basis of the system.

Key words natural scene image recognition, face recognition, region segmentation and extraction, pulse-modulation LSI, resistive-fuse network, Gabor-type wavelet transformation, cellular automaton, dynamic-link architecture

1. はじめに

近年、個人認証などの目的で顔認識技術が進展しており、また一般の情景画像から顔領域を抽出する技術も各種提案されている [1]。しかし、これらのほとんどは目、口などの部品配置や肌色などの顔特有の特徴を利用しており、一般的な物体の認識に適用できないものが多い。また、照明などの変化に口バストな認識技術も開発途上である。

一方、人は自然な情景の中から、人の顔を含むあらゆる物体を、どのような角度からでも、またどのような陰影がついていても、瞬時に認識することができる。残念ながら、最新の技術を持ってしてもこのような処理をリアルタイムで行うことは不可能である。これは現在の画像認識技術が、脳での視覚処理システムとは全く異なった情報処理様式とハードウェアアーキテクチャに依っていることに原因がある。

一般に従来から、脳の情報処理機能を解明し、それを工学的に応用しようとする試みが多くなされてきた。その際、研究のアプローチは大きく二つに分かれる。一方は、とにかく生理学的な知見に基づく忠実なモデルとハードウェアを作り上げ、それらをネットワークとして組み上げることで機能を実現しようとするアプローチである。他方のアプローチは脳が実現している情報処理機能を学び、ハードウェア構成自体を真似る必要はないとする考え方である。確かに、有機物で構成され、生化学・電気化学的反応で処理が進行し、3次元的なネットワーク構造を有する現実の脳と、かたや半導体中での電子輸送現象で動作が決定され、ほとんど2次元的な配置しか許されない現状のLSIとを比較すれば、同じ機能を実現するにも、同じ回路構成をとる必然性はないはずで、後者の立場の方が合理的に思われる。しかし、脳内での情報処理原理が完全には理解できていない現段階では、後者の立場で真に脳と等価な機能が実現できるという保証はない。そこで、脳型情報処理でポイントと思われる特徴に着目し、それを基にモデルとハードウェアを構築するというアプローチが賢明と思われる。

我々が構築を目指すシステムにおいては、必ずしも実際の脳・神経系を忠実に真似るアプローチはとらず、脳型処理モデル・ハードウェアのポイントとして超並列処理と（非線形）アナログダイナミクスを重視する。現状の集積回路でこれらを実現する手法として、パルス変調信号を用いた「アナログ・デジタル (AD) 融合アーキテクチャ」を提案している [2-4]。

本稿では脳の処理機能に学んだニューラルネットワークの枠組みによる複数の画像処理モデルを組み合わせ、自然画像から顔を含む任意の物体の認識を行うためのシステムについて述べ、AD融合アーキテクチャに基づき、セルラーネットワーク構造で画素並列動作を行う特徴抽出LSIを紹介する [5, 6]。

2. 自然画像認識のための情報処理モデル

人が意識的に物体の認識を行う場合、まず視野の中の特徴的な箇所注目する。ここでは輝度が周囲と顕著に異なる大まかな領域を順次注目していく。一般的にはこのような領域が必ずしも一つの意味ある物体を構成するわけではないが、ある状況

ではこの方法も有効に機能する。（離れた複数の領域を同一物体の一部と見なして認識する技術は、いくつかのモデルが提案されており、ハードウェア化は今後の研究課題である）。

我々の認識システムでは、まず自然画像の中から細かい濃淡を無視した大局的な領域を分割するために、抵抗ヒューズネットワークを用いる。例えば顔領域は目、眉、鼻、口などの細かい領域を消去し、均一な領域にして一体として分割する。このネットワークはアナログダイナミクスに従う簡単なセルラー構造をしているためにアナログ回路での実現が試みられてきたが [7-9]、この回路方式では非線形な特性を自在に実現することが難しく、大規模ネットワークでの制御性にも難がある。そこで、パルス変調方式を用いたピクセル回路を結合して等価な処理を実現する [5, 10]。このモデルの説明を第 3.1 節で、回路の説明を第 5 節で行う。

抵抗ヒューズネットワークは分割機能しかもたないため、分割された各領域を抽出する機能が必要である。これは、我々がすでに提案している非線形振動子ネットワーク [11-14]、またはそのモデルをデジタル的に簡略化したセルオートマトン型回路 [15] により実現できる。前者はアナログダイナミクスによる振動子の同期性を利用しており、領域境界のエッジ情報の欠けにも口バストであること、高次認識モデルへの展開が可能であることなど、高いポテンシャルを有しているが、専用LSIを設計する必要がある。一方、後者はデジタル回路で実現できるのでFPGAでも構成可能であること、回路構成が簡単であるため、上記の抵抗ヒューズネットワーク回路に組み込めることなどの利点を有している。本稿では後者の回路を第 6 節で説明する。

抽出した領域を認識するための特徴抽出については、人の視覚系の機能の一部を模擬したガボール・ウェーブレット変換を採用する。この変換は照明の変化などの影響を受けにくく、優れた特徴量として知られているが、膨大な演算量を必要とするため、これまで実用的な認識システムに採用された例はわずしかない [16, 17]。そこで、このタイプの変換をハードウェア的に実行するために、2層の抵抗ネットワークを用いてアナログダイナミクスを用いて演算する回路が提案されている [18, 19]。我々は、抵抗ヒューズネットワークと同様に、パルス変調方式を用いて画素並列で実行する回路を設計した [6]。ガボール変換回路の原理を第 3.2 節で、設計したLSIおよび測定結果を第 5 節で述べる。

ガボールウェーブレット特徴量を用いて認識を行うモデルとして、ダイナミック・リンク・アーキテクチャ (DLA) を採用する [20, 21]。これは局所の特徴のサンプリング点を最適化して歪みを吸収するマッチング法であり、物体の方向、変形に柔軟に対処でき、数多くある認識法の中でも最も精度の高い方法の一つとして知られている。現状の我々のシステムではこの処理はソフトウェアで行っており、ハードウェア化による高速化は今後の課題である。処理の概要を第 7 節で述べる。

以上のモデルに基づき、自然画像から顔を認識する例をまとめたものを図 1 に示す。

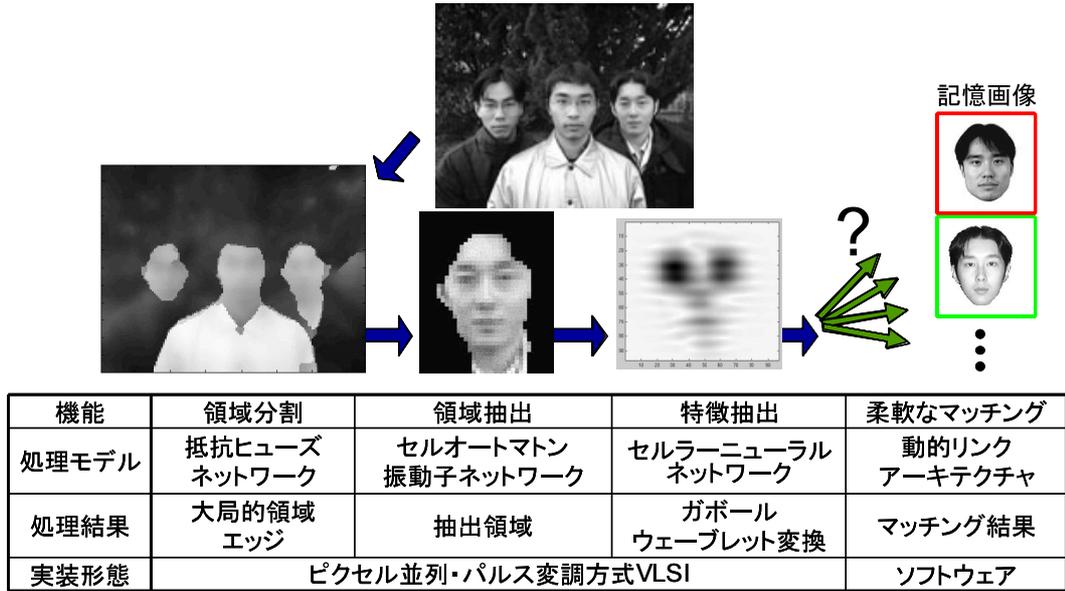


図1 我々の提案する自然画像認識処理

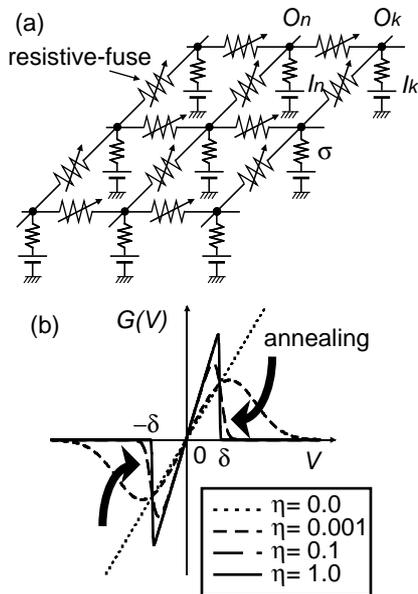


図2 抵抗ヒューズネットワークモデル

3. 領域分割・特徴抽出のためのセルラーネットワークモデル

3.1 抵抗ヒューズネットワークモデル

抵抗ヒューズネットワークは、元々、画像のエッジ情報を保存しながらノイズを除去することにより画像の再構成を行う画素並列処理モデルとして提案された[7]。我々はこれを微小な領域をノイズとみなして除去し、画像を大域的な領域に分割する処理に適用する。抵抗ヒューズネットワークはアナログ回路での実現が自然なことから、いくつかのアナログ抵抗ヒューズネットワークチップが提案されてきた[8, 9]。アナログ回路で構成した抵抗ヒューズネットワークを図2(a)に示す。ここで、各画素の輝度に比例した値を I_n に与え、ノード電圧 O_n を処理

結果とする。

このネットワークは以下の関数 E が最小になるように安定化する。

$$E = \sum_n \sum_{k \in N_n} \int_0^{O_n - O_k} G(V) dV + \frac{\sigma}{2} \sum_n (O_n - I_n)^2 \quad (1)$$

ここで、 σ は定数（抵抗のコンダクタンス）であり、 N_n は画素 n の隣接画素の集合を表す。 $G(\cdot)$ は図2(b)に示すような抵抗ヒューズの I-V 特性である。これは

$$G(V) = \left[\frac{1}{1 + \exp(-2\eta(\delta^2 - V^2))} \right] \frac{V}{R} \quad (2)$$

と与えられる。ここで、 η, δ および R は定数である。

この抵抗ヒューズ素子は $\eta = 0$ の時は線形抵抗になるが、 $\eta = 1$ の時にはほぼ完全な「抵抗ヒューズ」となる。すなわち、 $|O_n - O_k| > \delta$ で隣接セルとの結合が切れてしまい、エッジ検出ができるが、 $|O_n - O_k| \leq \delta$ のときは線形抵抗となり平滑化処理が行われる。したがって、細かい領域を平滑化しながら、大まかな領域の間をエッジとして切り出すことができる。図3に実画像を用いた処理結果を示す。パラメータを適当な値に設定すると、同図(d)に示すように顔領域の内部を平滑化して切り出すことができる。

この抵抗ヒューズネットワークの動作を、第4.4節で述べるPWMピクセル回路でのパルス幅を用いた離散時間ダイナミクスにより模擬する。すなわち、以下の最急降下法で E の極小値を求める。

$$O_n(t+1) = O_n(t) - \nu \frac{\partial E}{\partial O_n},$$

$$\frac{\partial E}{\partial O_n} = \sum_{k \in N_n} G(O_n - O_k) + \sigma(O_n - I_n) \quad (3)$$

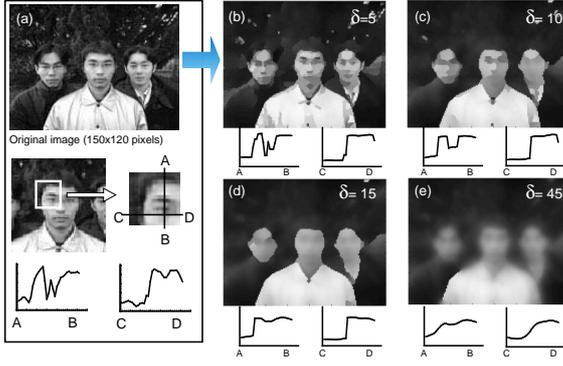


図3 抵抗ヒューズネットワークによる画像の大局的領域分割

ここで、 ν は定数である。この式 (3) の右辺はノード n から流れ出る電流を計算していることに他ならない。なお、ローカルミニマムに陥るのを避けるために、 $G(\cdot)$ の形を変えながらアンニリングを実行する。すなわち、図 2(b) に示すように、ゆっくりと η を増加させていくことにより、大局的な領域分割が実現できる。

3.2 ガボール型ウェーブレット変換回路 (ガボール型フィルタ)

ガボールウェーブレット変換は、入力 (画素の輝度値) とカーネル関数:

$$\psi_{\vec{\omega}_0}(\vec{x}) = \frac{1}{2\sqrt{\pi}\sigma} \exp\left(-\frac{\vec{\omega}_0^2 \vec{x}^2}{\sigma^2} + j\vec{\omega}_0 \vec{x}\right) \quad (4)$$

との畳み込みにより局所的 (空間) 周波数を抽出する。ここで、 $\vec{\omega}_0$ は $\psi_{\vec{\omega}_0}$ の周波数と方向を決めるパラメータである。この変換が特徴抽出に有効なのは、空間および周波数の両領域において ψ の広がりが最小になるからである。

しかし、窓関数としてガウス関数を用いたガボール変換は計算の効率化がしにくく、畳み込みをそのまま計算すると膨大な演算量になってしまう。そこで、窓関数を指数関数に置き換えて、セルラーニューラルネットワークの理論を用いて、抵抗ネットワークにより構成されるアナログ回路上で計算させる方法が提案された [18, 19]。このようにガウス関数に似た関数を窓関数に用いる変換をガボール型ウェーブレット変換と呼ぶ。

ガボール型ウェーブレット変換を実行する 2 次元のアナログ回路を図 4 に示す [19]。抵抗ネットワークが指数関数的な特性を有するローパスフィルタ (平滑化フィルタ) として働くことはよく知られているが、これを 2 層にし、たすきがけの層間結合を持たせることにより、各層でガボール型変換の実部と虚部の出力をそれぞれ同時に得ることができる。ここで、入力は電流源 I_n であり、出力結果は実部と虚部に相当する二つのノード電圧 $v^r(m, n)$, $v^i(m, n)$ で得られる。この回路で実現される畳み込みカーネルは以下のように、指数関数で減衰する正弦波である。

$$h_{\vec{\omega}_0}(\vec{x}) = \frac{\lambda}{2} \exp(-\lambda|\vec{x}| + j\vec{\omega}_0 \vec{x}) \quad (5)$$

ここで、 λ は減衰の定数である。これにより、空間周波数 ω_0 の成分が局所的に抽出される。

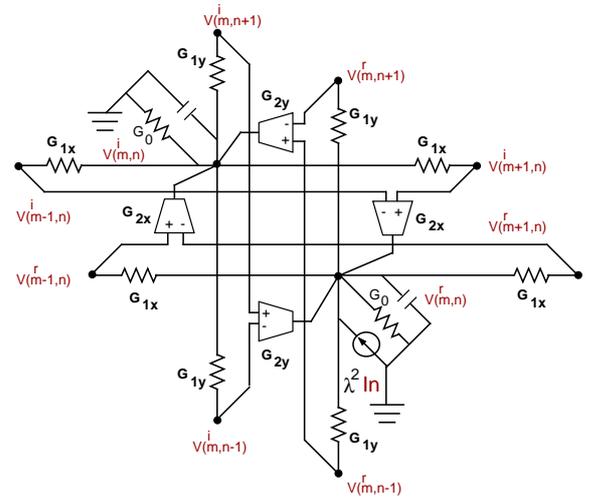


図4 2次元アナログセルラーニューラルネットワーク方式ガボール型ウェーブレット変換回路 (1 ピクセル相当)

我々の LSI 化のアプローチでは、このアナログ回路をそのまま用いるのではなく、回路のダイナミクスを離散化して、パルス幅変調 (PWM) 方式を用いて、繰り返し計算により解を得る。電圧の更新式は以下のとおりである。

$$\begin{aligned} \Delta V^r(m, n) = & -G_0(V^r(m, n) - V_0) \\ & +G_{1x}(V^r(m-1, n) - V^r(m, n)) \\ & +G_{2x}(V^i(m-1, n) - V^i(m+1, n)) \\ & +G_{1x}(V^r(m+1, n) - V^r(m, n)) \\ & +G_{2y}(V^i(m, n+1) - V^i(m, n-1)) \\ & +G_{1y}(V^r(m, n-1) - V^r(m, n)) \\ & +G_{1y}(V^r(m, n+1) - V^r(m, n)) \\ & +\lambda^2(I(m, n) - V_0) \end{aligned} \quad (6)$$

$$\begin{aligned} \Delta V^i(m, n) = & -G_0(V^i(m, n) - V_0) \\ & +G_{1x}(V^i(m-1, n) - V^i(m, n)) \\ & +G_{2x}(V^r(m+1, n) - V^r(m-1, n)) \\ & +G_{1x}(V^i(m+1, n) - V^i(m, n)) \\ & +G_{2y}(V^r(m, n-1) - V^r(m, n+1)) \\ & +G_{1y}(V^i(m, n-1) - V^i(m, n)) \\ & +G_{1y}(V^i(m, n+1) - V^i(m, n)) \end{aligned} \quad (7)$$

ここで、 $\lambda^2 = G_0 + 2G_{1x} + 2G_{1y} - 2(G_{1x}^2 + G_{2x}^2)^{1/2} - 2(G_{1y}^2 + G_{2y}^2)^{1/2}$, $G_{2x}/G_{1x} = \tan \omega_{0x}$, $G_{2y}/G_{1y} = \tan \omega_{0y}$ であり、 V_0 は 0 に相当する電位を表す。

この回路の利点は、 ω_0 の x, y 方向それぞれの成分 (ω_{0x}, ω_{0y}) の比 (すなわち、(6) および (7) 式の係数) を変えることにより任意の方向のガボール型変換が実行できる点にある。通常、顔認識などのためにガボール変換を行う場合、少なくとも 4 方向

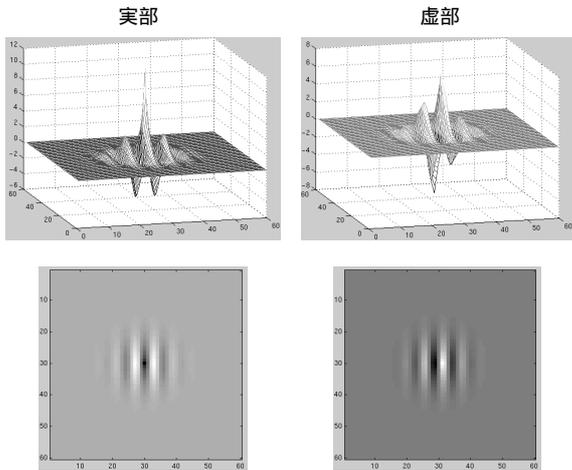


図5 2次元ガボール型ウェーブレット変換回路のインパルス応答

の変換が必要であるので、この特長の効果は大きい。

数値計算の結果得られたインパルス応答の例を図5に示す。

4. パルス変調回路技術

4.1 特徴

図2または図4に示したようなアナログ回路方式は、連続時間・連続状態のダイナミクスが実行できるので、マイクロ秒オーダで高速に結果が得られるという利点がある。回路ユニットのサイズも小さくできるので、超並列動作が可能である。しかし、実画像の認識に必要な 100×100 画素程度を並列に処理する大規模なアナログ回路を設計することは極めて難しい。それは、実際のアナログLSIには、予測困難な寄生素子や、ノイズ、クロストーク、素子のパラメータばらつきなどの様々な非理想的な要因が発生するからである。さらに、従来アナログ方式では、抵抗ヒューズ特性のような非線形変換を自在に実現することが難しい。

一方、パルス変調方式は、電圧または電流方向に2値の値をもち、時間軸方向にアナログ量を有するパルス信号を用いて、時間領域でアナログ情報処理を行う技術である。これはSi CMOS技術の微細化スケールにも対応できるので、大規模なアナログ情報処理回路を実現するのに適している。低電圧動作も可能で、高い制御性と既存のデジタルシステムとのマッチングもよい。ノイズやクロストークなどにも強いという利点もある。

また我々は、パルス幅/パルス位相変調(PWM/PPM)信号からアナログ電圧への変換過程で任意の時間波形を用いることにより、任意の非線形変換を実現する回路方式を開発している[4]。これを用いれば、離散時間・連続状態の非線形ダイナミカルシステムを高い制御性で実現することができる。したがって、この方式は抵抗ヒューズ特性の実現に用いることができる。以下に基本的な積和演算とこの非線形変換の方法について述べる。

4.2 積和演算

PWM信号を用いた積和演算は図6に示すような構成で実行

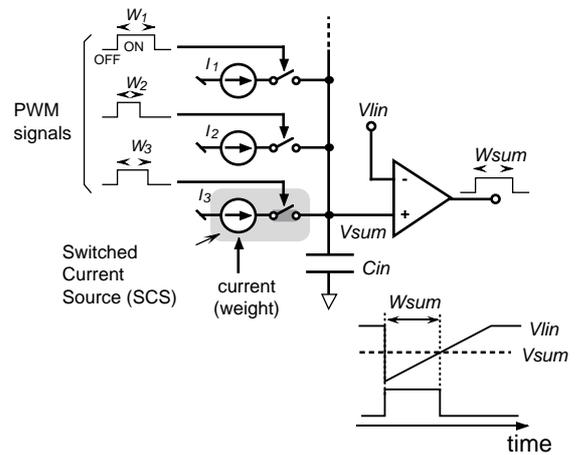


図6 PWM信号による積和演算の原理

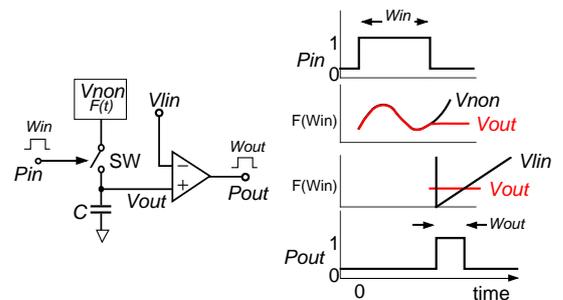


図7 PWM信号を用いた任意非線形変換の原理

できる。パルス幅 W_i ($i = 1, 2, 3, \dots$) を有するPWM入力信号が対応する電流源をスイッチし、電流 I_i がキャパシタ C_{in} に時間 W_i の間だけ流れる。キャパシタに蓄積される電荷量 Q_{sum} と端子電圧 V_{sum} は、それぞれ

$$Q_{sum} = \sum_i I_i W_i, \quad (8)$$

$$V_{sum} = \frac{Q_{sum}}{C_{in}} = \frac{\sum_i I_i W_i}{C_{in}},$$

と表される。こうして、電流源の電流量で重み付けられたPWM信号の加算結果がキャパシタの電荷量または端子電圧として得られる。電圧 V_{sum} は、線形に変化する参照ランプ電圧 V_{lin} とコンパレータで比較することにより、パルス幅 W_{sum} を有するPWM信号に変換されて出力される。

なお、我々が用いるモデルでは正負の加算を実行できないが、それには2通りの方式が考えられる。正・負の量をそれぞれ独立に加算して、最後に両者を減算する方式[22]、および単一のキャパシタに正負のスイッチト電流源を接続し、電荷を充・放電することにより実現する方式である。後者の方が加算のためのキャパシタが1個でよいので省面積になるが、加減算それぞれの電圧範囲は1/2になるので、精度は低下する。第4.4節で述べるピクセル回路では後者の方式を採用している。

4.3 任意非線形変換

PWM信号を用いた任意非線形変換の原理を図7に示す[4]。これは、任意の非線形電圧波形 $F(t)$ をPWMまたはPPM信号でサンプル・ホールドすることで実現するパルス-電圧変換で

ある。

PWM 信号から PWM 信号への非線形変換は次のようにして実現する。時間領域で任意の電圧波形 $F(t)$ を有する参照電圧 V_{non} がスイッチ SW を介してキャパシタ C に供給されているとする。スイッチ SW は入力パルス P_{in} により制御される。ここで、 P_{in} の立ち上がりで $V_{non} = F(0)$ とする。スイッチ SW はパルス幅の期間 W_{in} で導通する。入力パルス P_{in} の立ち下がりスイッチが開くと、キャパシタの端子電圧 V_{out} は $F(W_{in})$ に等しくなり、これを保持する。電圧 V_{out} は線形に変化する参照ランプ電圧 V_{lin} とコンパレータで比較され、電圧に比例したパルス幅 W_{out} を有する P_{out} が生成される。この結果、 $W_{out} = F(W_{in})$ が成り立ち、非線形参照電圧波形に比例した非線形関数で入出力 PWM 信号のパルス幅が変換される。

この方法が有用なのは、任意のアナログ入出力関数を作り出すのが難しいのに対し、時間の関数としての電圧波形 $F(t)$ を作り出すのが容易なことによる。それは、さまざまなアナログ発振回路で作り出すこともできるし、デジタル回路でルックアップテーブル法や関数生成法によりデジタル波形を作り出し、それを D/A 変換器によりアナログ波形にすることも実現できる。このデジタル方式の場合、任意の非線形波形を高い精度で生成でき、リアルタイムで波形を変更することもできる。これは非線形ダイナミクスを任意にリアルタイムで変更出来ることを意味する。デジタル波形をローパスフィルタで平滑化すれば、デジタル波形がかなり低ビット精度であっても、量子化の影響を受けずに、アナログ演算を行うことができる。現在では 100MHz 以上で動作する D/A 変換器も実現可能なので、6 ビット以上の時間分解能で 1MHz 以上で動作する任意非線形変換回路を構成することができる。アナログ波形生成器は多くの PWM 変換回路で共通に用いることができるので、システム全体に対する波形生成のオーバーヘッドはわずかなものにできる。したがって、この方式は大規模な並列動作型非線形変換回路の実現に適した方法である。

4.4 PWM ピクセル回路

抵抗ヒューズネットワークとガポール型フィルタ用抵抗ネットワークのような画素並列型アナログ処理回路においては、各ピクセル回路で、画素輝度に相当する入力値および隣接画素の状態値と自身の状態値との差を計算し、適当な関数変換（抵抗ヒューズ型の非線形関数変換または単純な線形変換）の結果に従って状態値を更新する処理を繰り返す。PWM ピクセル回路はそのような処理のために設計された回路である。すなわち、画素並列で動作し、隣接画素の情報を用いて PWM 方式の（非線形）演算により状態を更新する。PWM 方式は通常のアナログ回路での実現に比べて定常状態に達するのに多くの時間を要するが、極めて高い制御性が得られる。これはアナログのパルス幅を使ってダイナミクスを数値的に解くからである。

減算回路を組み込んだ PWM ピクセル回路の概略とタイミング図を図 8 に示す [10, 23]。まず、画素 n の輝度値に相当する電圧をキャパシタ C_{In} に蓄えておく。減算回路 SUB はアナログ入力電圧 in_1 と in_2 の差分に相当する PWM 信号を出力する回路である。 SUB はまず in_1 と in_2 の大小関係を判定して符

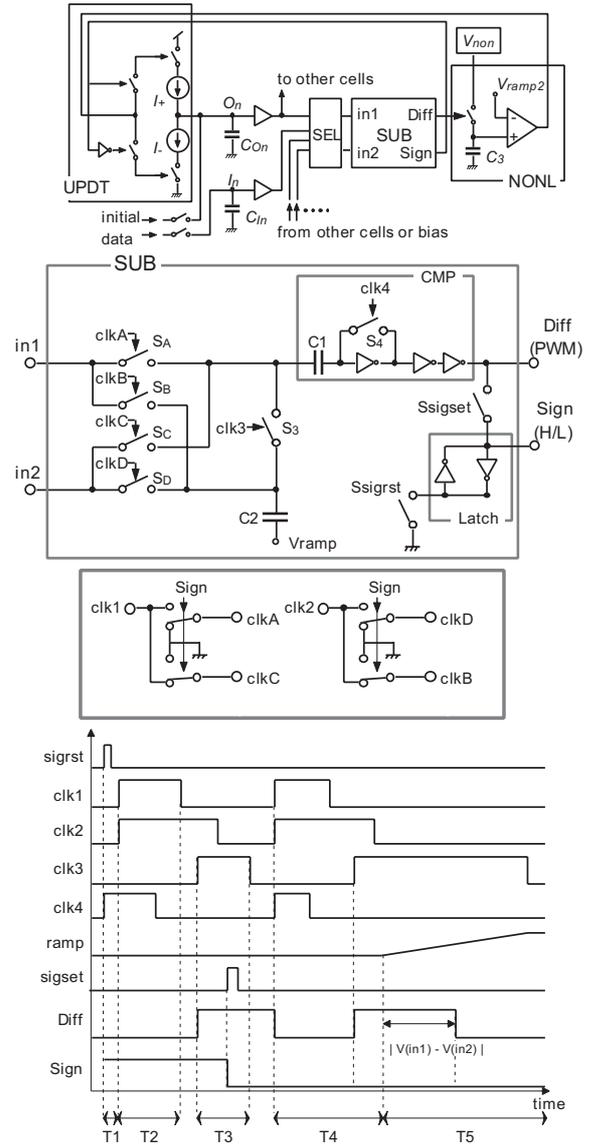


図 8 PWM ピクセル回路と減算回路 SUB の動作タイミング図 (処理シーケンスは $in_1 < in_2$ の場合)

号ビット $Sign$ を出力し、次に、 $|in_1 - in_2|$ に相当するパルス幅を有する PWM 信号を $Diff$ から出力する。その結果を非線形変換回路 $NONL$ に通して、回路の状態を記憶しているキャパシタ C_{On} の電荷を更新する。したがって、 O_n は次のような離散時間ダイナミクスにしたがって更新される。

$$O_n(t+1) - O_n(t) = \sum_{j,k \in N'_n} G_{jk}(O_j - O_k) + F_j(O_j - I_n) \quad (9)$$

ここで N'_n はセル n 自身を含む隣接セルを意味する。また、 $G_{jk}(\cdot)$ と $F_j(\cdot)$ は任意の奇関数である。

回路動作を以下に詳細に説明する。

(1) セレクタ SEL が計算すべき一対の信号の組を選択し、それを in_1 および in_2 とする。

(2) 期間 T_1 では、スイッチ S_{sigrst} を一時的にオンしてラッチ回路 $Latch$ をリセットすることにより $Sign$ が “High” となり、その結果 clk_A に clk_1 の信号が、 clk_D に clk_2 の信号が

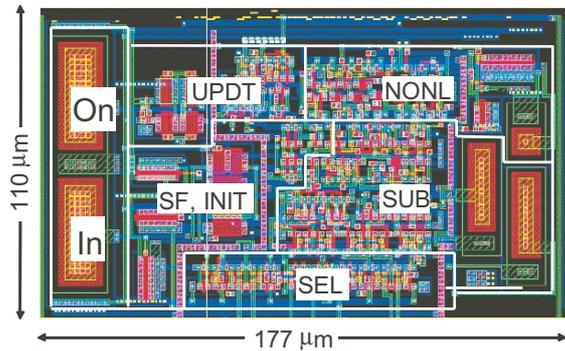


図9 PWM ピクセル回路のセルレイアウト図

それぞれ現れる。また、スイッチ S_4 をオンさせてコンパレータ CMP をリセットする。

(3) 期間 T_2 では、スイッチ S_A および S_D がオンし、キャパシタ C_1, C_2 にそれぞれ in_1, in_2 の電位をセットする。その後、スイッチ S_4 および S_A をオフする。

(4) 期間 T_3 では、最初にスイッチ S_3 をオンして C_1 と C_2 を直列接続し、コンパレータ CMP の出力ノードの電位により、 in_1 と in_2 のどちらの電位が高いかを判定する。すなわち、 $in_1 > in_2$ ならば $Diff = \text{“High”}$ であり、 $Sign$ の値はリセット時から変化しない。逆の場合は $Sign$ が反転する。その判定結果を S_{sigset} をオンすることによりラッチ回路 $Latch$ に取り込み、保持する。

(5) 期間 T_4 では $Sign$ により設定されたパスで T_2 および T_3 と同様の動作が繰り返される。

(6) 期間 T_5 では C_2 の下の端子にランプ波形 V_{ramp} を入力することにより、 C_1 と C_2 に蓄えられた2つのアナログ電圧の差分に相当する PWM 信号を $Diff$ に出力する。

(7) PWM 信号 $Diff$ は非線形変換回路 $NONL$ において非線形電圧源 V_{non} をスイッチングする。 V_{non} の電圧波形は対応する(非線形)関数 $G_{jk}(\cdot)$ または $F_j(\cdot)$ と同形である。その結果、キャパシタ C_3 は式(9)の対応する項の値を保持する。

(8) キャパシタ C_3 に保持された電圧は、線形ランプ電圧波形 V_{ramp2} と比較することにより再び PWM 信号に変換される。

(9) $Sign$ 信号で決定した充放電の向きにしたがって、電流源 I_+ または I_- が PWM 信号の時間だけ ON され、キャパシタ C_{On} の電圧値が更新される。

以上のシーケンスを繰り返すことで、式(3)や(6),(7)で示されたダイナミクスを実現することができる。

5. 試作チップと測定結果

PWM ピクセル回路を $0.6\mu\text{m}$ CMOS 技術を用いて設計し、VDEC のサービス(ルーム)で試作した。チップに搭載した PWM ピクセル回路の数は 2×20 である。ガボール型フィルタでは1画素当り2ピクセル回路を必要とするので、これで、抵抗ヒューズネットワークとガボール型フィルタそれぞれ20ピクセル分の1次元ネットワークを構成できる。ピクセル回路のレイアウト図とチップ写真を図9および10に示す。

この試作チップを用いて抵抗ヒューズネットワークを構成し

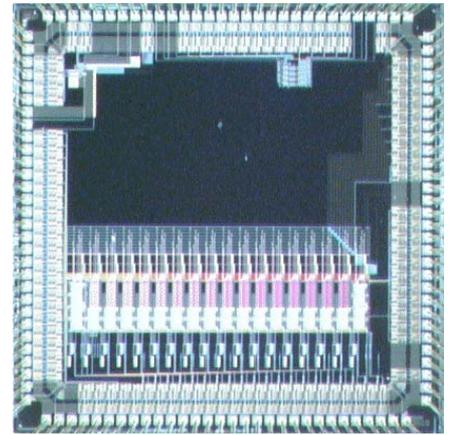


図10 PWM ピクセル回路のチップ写真 ($0.6\mu\text{m}$ CMOS, チップサイズ:4.6mm 角)

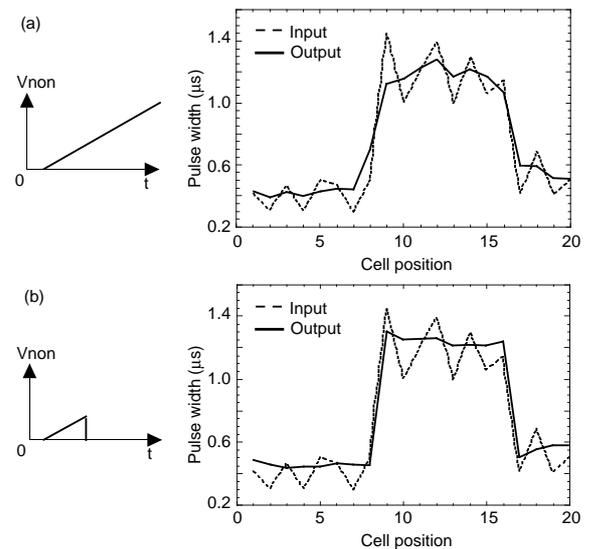


図11 抵抗ヒューズネットワークの測定結果

た場合の測定結果を図11に示す。線形抵抗に相当するネットワークでの結果(同図(a))に比べて、抵抗ヒューズネットワークの結果(同図(b))ではエッジを保存しながら細かいノイズを除去できているのがわかる。これらは更新を30回行った結果である。1回の更新は式(3)より3項(両隣りのセルと自分自身)より成り、各項の更新に $2.8\mu\text{s}$ 要したので、合計約 $250\mu\text{s}$ で計算が終了している。これを2次元画像に適用した場合は5項(上下左右のセルと自分自身)の計算が必要で、アンローリングを少なくとも3回行う必要があるため、処理時間は約 1.3ms となる。

図12にガボール型フィルタを構成した場合の測定結果を示す。約100回の更新で収束し、ほぼ数値計算通りのインパクト応答が得られた。これは周期 $(= 2\pi/\omega_0)$ が4ピクセルの結果であり、処理時間は約 $600\mu\text{s}$ であった。

以上の結果より、試作チップで我々の回路が正しく動作していることを確認した。

5.1 2次元ガボール型フィルタ LSI の設計と試作評価結果

前節で述べたピクセル回路を用いて、 12×24 画素を処理する2

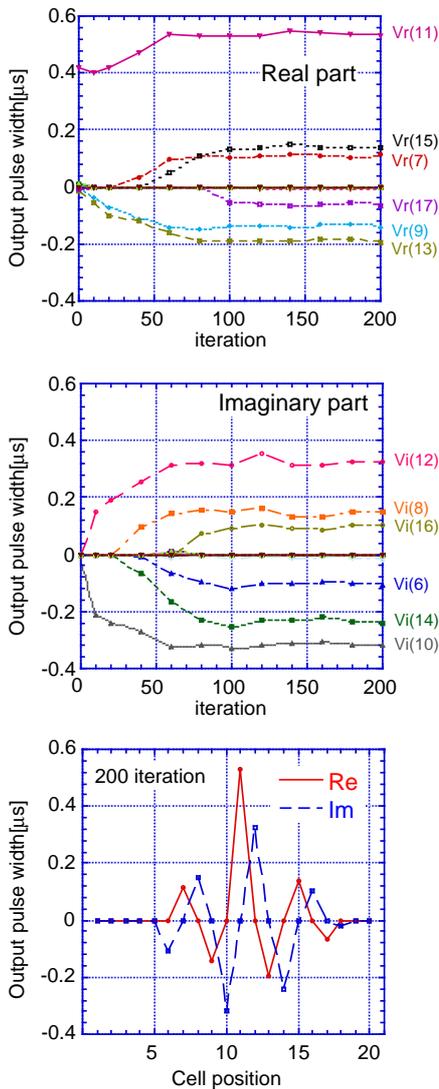


図 12 ガボール型フィルタの測定結果（インパルス応答: 初期値としてセル#11 にのみ値を入力）

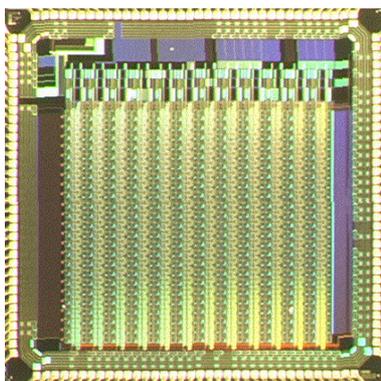


図 13 2次元ガボール型フィルタ LSI のチップ写真 (0.35 μ m CMOS, チップサイズ:4.9mm 角)

次元ガボール型フィルタ LSI を VDEC (ローム) 0.35 μ m CMOS 技術を用いて設計した。チップ写真を図 13 に示す。チップサイズは 4.9mm 角である。電源電圧 3.3V のときの消費電力は約 30mW であった。ガボール型フィルタとしての測定結果を図 14

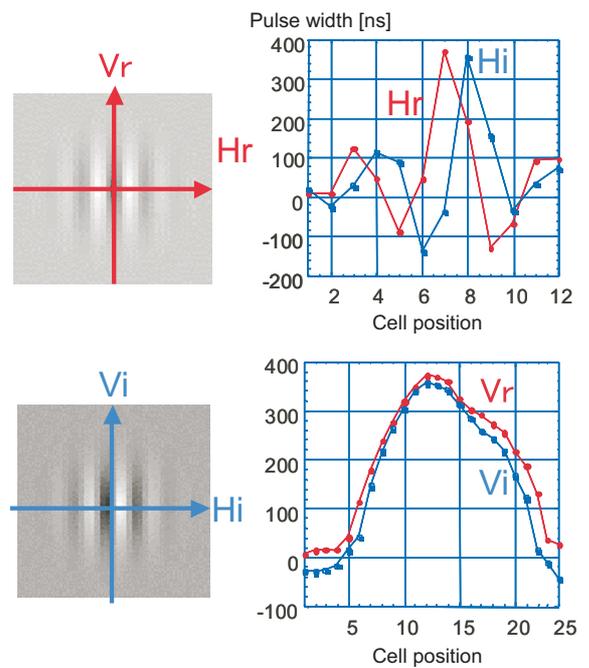


図 14 ガボール型フィルタの測定結果 (2次元インパルス応答の断面)。左に断面方向を示すための数値シミュレーション結果を添えた

に示す。2次元のインパルス応答の1次元断面を示している。ほぼ正しいインパルス応答が得られた。これは周期 ($= 2\pi/\omega_0$) が4ピクセルの結果であり、処理時間は約 2ms (更新回数 220回) である。

6. 領域抽出のためのセルオートマトンモデル

画像の分割抽出の一般的な手法として、エッジ情報を利用した「ラベリング」がよく用いられる。この手法ではメモリ空間に画素位置に対応するラベリング結果を保持するので、分割された領域を任意に処理することができるが、ラベル情報を保存するメモリが必要であることと、ラベル修正のために複数回の走査が必要であるために処理に時間がかかるという欠点がある。

本節では、「抵抗ヒューズネットワーク」により画像の領域分割後に得られるエッジ情報を利用し、エッジで囲まれた領域を画素並列で個別に抽出するアルゴリズムとそれを実現する回路構成を述べる。提案する画素回路は極めてコンパクトなので、画像分割用ピクセル回路内に組み込むことができるが、ここでは FPGA に実装した例を示し、現状の FPGA による画素並列処理の可能性を示した [15]。

6.1 領域抽出アルゴリズム

考案した領域抽出アルゴリズムを図 15 に示す。各画素のとり得る状態は、未発火・発火中・発火済の3状態とする。ある時点で発火した画素は、次のタイミングで隣接の未発火画素を発火させるといったセルオートマトン型アルゴリズムにより、エッジに囲まれた領域内で燎原の火のように発火状態を広げていく。

アルゴリズムの詳細を以下に示す。

(1) 初期値設定

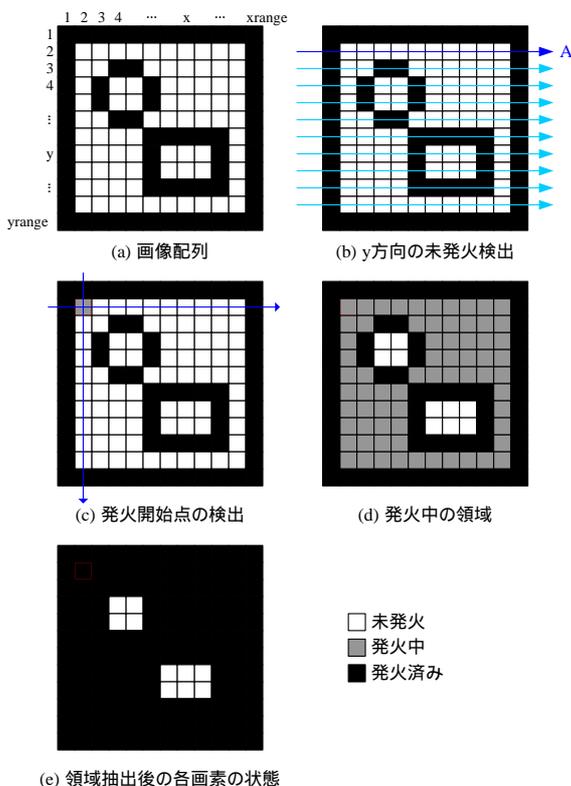


図 15 領域抽出アルゴリズム

各画素に、エッジであれば「発火済」、そうでなければ「未発火」という状態を持たせる。画像端については、図 15(a)に示すようにエッジとみなす。

(2) 発火開始点の設定

まず y 方向に「未発火」の画素が存在する行を検出する。図 15(b)において x 方向の線がその行を示す。その中で一番上の行を選択する。もしくは、 $y = 2$ から順に「未発火」の画素が存在する行を調べていってもよい。最初はエッジ以外すべて「未発火」なので図 15(b)の A の行が選択される。

その行を x 方向に走査していき「未発火」の画素を検出する。検出された画素の状態を「発火中」に変更する(図 15(c))。これを「発火開始点」と呼ぶ。

(3) 発火過程

未発火の各画素に対し、隣接画素が発火していればその画素自身も発火するという処理を繰り返し、発火開始点を拠点に発火領域を拡大していく。

(4) 発火過程終了判定

発火領域が広がらなくなった時点で、発火過程を終了する(図 15(d))。

(5) 発火領域の抽出

発火している領域情報を取り出し、画素の状態を「発火済」に変更する(図 15(e))。初回のプロセスでは背景が抽出される。

6.2 領域抽出回路

前節で提案したアルゴリズムをセルオートマトン型の画素並列動作で実現する。画素ユニット回路では、各画素が 3 状態を取り得るので、2bit のレジスタ ($b1, b0$) を用意し、1 時刻前の $b1$ 状態を記憶するために 1bit レジスタ ($b2$) を用意した。各画素

表 1 各状態に対応するレジスタ値

| 画素の状態 | $b1$ | $b0$ |
|-------|------|------|
| 未発火 | 0 | 1 |
| 発火中 | 1 | 1 |
| 発火済 | 0 | 0 |

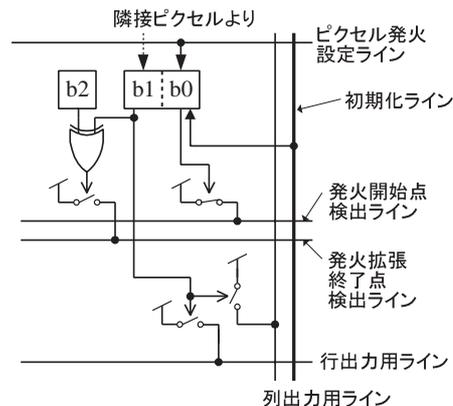


図 16 領域抽出用画素ユニット回路

回路が取る得る状態とその値を表 1 に示す。画素ユニット回路を図 16 に示す。

6.3 FPGA への実装、測定

上記のアルゴリズムを FPGA に実装した。用いたチップは ALTERA 社 APEX シリーズ EP20K 400EFC-672-1X (ゲート数 400K)、クロック周波数は 25MHz とした。30×30 画素相当の回路を実装した (ロジックセル使用量 56%, 内蔵メモリは使用せず)。

30×30 画素の画像を用いて測定をした結果を図 17 に示す。出力された領域は図 17(3)の斜線で定義される部分であり、発火領域を灰色で示している。発火している領域が正しく出力されていることがわかる。処理時間は 5 μ sec であった。25MHz で動作させているので 125 クロックで動作が完了していることになる。

この方式では、細かな領域がなければ(抵抗ヒューズネットワークで分割するとこれは成り立つ)、処理時間はおおよそ画素数の平方根に比例し、25MHz 動作では 100×100pixel の画像で約 15 μ sec がかかることが見積もれる。一般的な逐次処理型のラベリング処理では、処理時間は画素数に比例するので、画素並列方式により処理が大幅に(数 10 倍)高速化できることがわかる。また、領域が逐次的に抽出されてくるという特徴は、後段で個別領域の認識処理を行うのに適している。

提案した回路は 1 画素当たり 3 ビットの情報保持でよく、隣接結合と共通の行/列バスラインで構成されるので、極めてコンパクトである。以上のことから、この回路は画素並列型の領域分割/エッジ抽出回路(抵抗ヒューズネットワーク)に組み込むと効果的である。

7. ダイナミックリンクに基づく物体認識

100×100 画素の画像に 8×8 の格子を割り当て、格子点上のガボールウェーブレット特徴量(4 方向×5 周波数=20 個)に

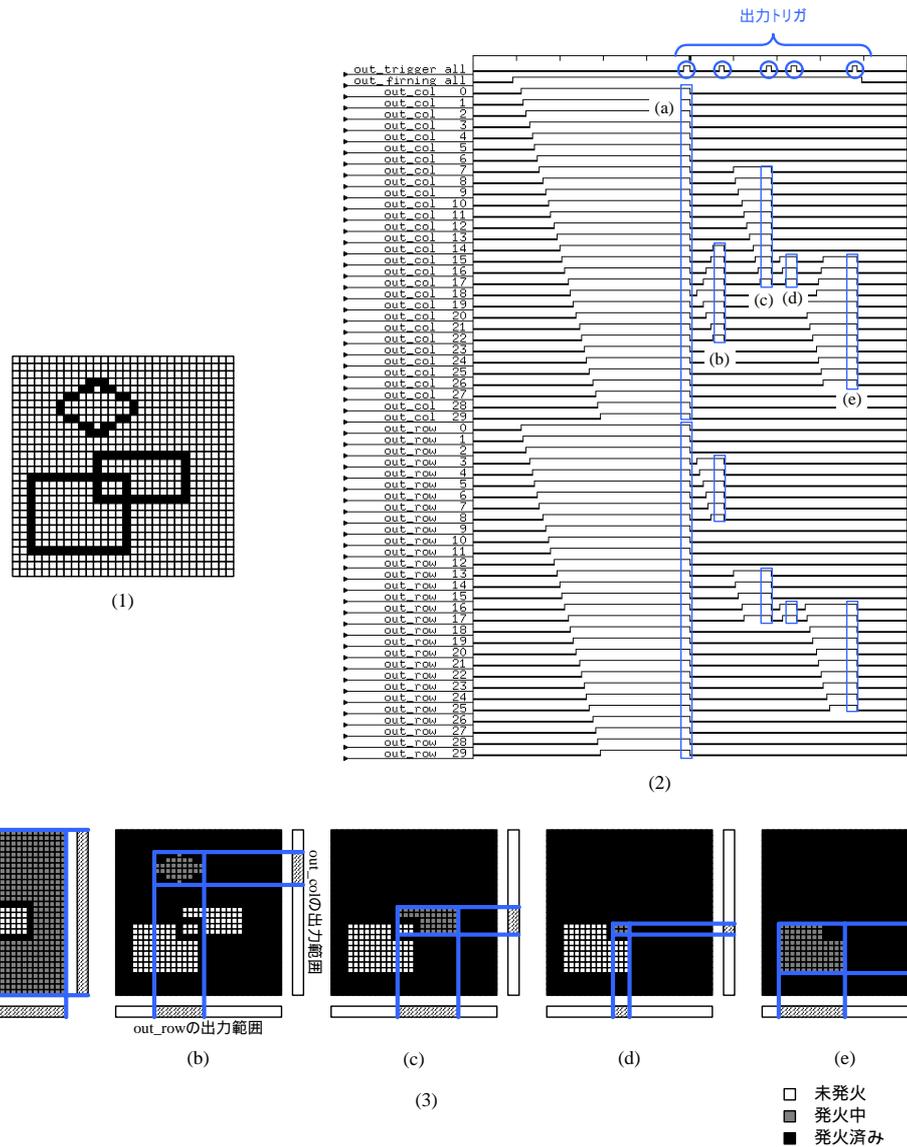


図 17 FPGA 実装結果, (1) エッジ情報, (2) 測定波形 (ロジックアナライザ出力), (3) 領域出力結果。測定波形の (a)-(e) は領域出力結果の (a)-(e) に対応している

ついて、入力画像と記憶画像をマッチングする。格子点を歪ませながら、歪み量と特徴量マッチング度の和で表現される評価関数が最良となる格子歪み位置を探索する。今回は顔認識に重点を置いたので、顔のパーツ（目、鼻、口）の特徴が良く捉えられるような格子配置をあらかじめ決定し、それを固定して用いた。

パソコン（Pentium4/1.5GHZ, OS:Linux）上で C 言語により実装した結果、マッチング時間に 1 記憶画像当り約 0.25 秒を必要とした。10 記憶画像からの認識処理画面を図 18 に示す。顔の向き、表情や照明などの変化に影響されにくい認識ができています。また、顔以外の本の表紙などでも正しく認識できることがわかる。

8. システム制御用 FPGA 搭載 PCI ボード

我々のパルス変調方式 LSI では、画素の輝度値や演算結果をパルス幅変調 (PWM) 信号で表現する。この信号はクロック同

期のデジタルカウンタを用いて、容易にデジタル信号と相互に変換することができる。また、LSI 動作に必要なアナログバイアス電圧および任意波形電圧は、ボード上に D/A 変換チップを搭載してデジタル的に制御できる。したがって、入出力制御を含めたチップ制御回路をデジタル方式で構成することが可能である。これらの回路は実用化レベルでは当然チップ内に組み込まなければならないが、研究段階では制御回路の仕様変更が容易なように外付けの FPGA チップを用いることが望ましい。

そこで、パルス変調方式 LSI を制御し、カメラからの画像取り込みやパソコンとのインターフェースを実現するための FPGA 搭載の PCI ボードを開発した^(注1)。開発した FPGA ボードの仕様、機能ブロック図、およびボード写真を図 19 に示す。パルス変調方式 LSI を制御するために、160 ピンの入出力ピンを有し、パルス幅を高精度にカウントするために内蔵 PLL 付きの FPGA を搭載している。

(注1)：開発元はワイエムシステムズ株式会社。



図 18 ダイナミックリングアーキテクチャに基づく画像認識。上段：入力画像（白線の枠内が認識領域），下段：10 個の記憶画像からの認識結果（太線枠）

9. ま と め

画素並列動作によるダイナミクスを利用して、自然画像の特徴抽出を行う複数の処理モデルとそれらを実時間で実行する LSI、およびそれらを組合わせた認識集積システムについて述べた。

処理モデルと LSI 実装については、抵抗ヒューズネットワークによる大域的領域分割回路・抵抗ネットワーク型ガボールフィルタ回路とパルス変調方式による LSI 実現、セルオートマトン型画像領域抽出アルゴリズムとその FPGA による実装結果を述べた。また、試作 LSI の制御用として、PCI バスを利用する FPGA ボードを開発した。現在、システムに組込可能な高ピクセル密度の画像処理 LSI の設計を進めるとともに、FPGA ボードを中心としてリアルタイム認識システムを構築中である。

謝 辞

本研究は広島大学大学院先端物質科学研究科機能集積システム工学研究室で行われた。LSI 設計を担当された学生諸氏および永田真助手（現在、神戸大学工学部助教授）に感謝いたします。

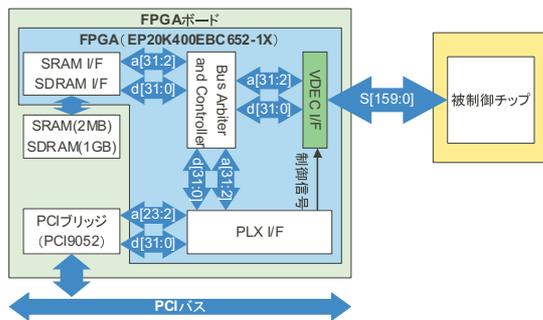
本研究は文部省科研費特定領域研究 (A) 「知的瞬時処理複合化集積システム」の補助を受けて行われた。チップ試作は東京大学大規模集積システム設計教育研究センターを通し、ローム (株) および凸版印刷 (株) の協力で行われたものである。

文 献

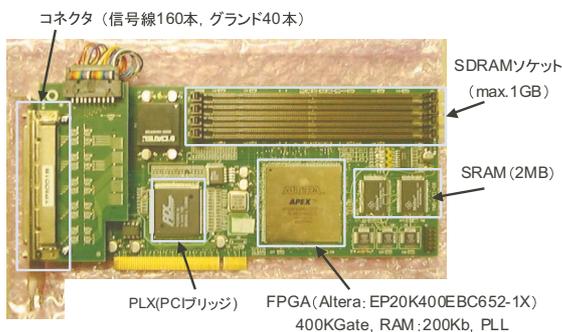
- [1] 赤松茂, “コンピュータによる顔の認識-サーベイ-,” 電子情報通信学会論文誌, vol. J80-D-II, no. 8, pp. 2031-2046, 1997.
- [2] A. Iwata and M. Nagata, “A concept of analog-digital merged circuit architecture for future VLSI’s,” IEICE Trans. Fundamentals., vol. E79-A, no. 2, pp. 145-157, 1996.
- [3] A. Iwata, T. Morie, and M. Nagata, “Merged analog-digital circuits using pulse modulation for intelligent SoC applications,” IEICE Trans. Fundamentals., vol. E84-A, no. 2, pp. 486-496, 2001.
- [4] T. Morie, S. Sakabayashi, H. Ando, M. Nagata, and A. Iwata, “Pulse modulation circuit techniques for nonlinear dynamical systems,” Proc. Int. Symp. on Nonlinear Theory and its Applications (NOLTA’98), pp. 447-450, Crans-Montana, Sept. 1998.
- [5] T. Morie, M. Miyake, S. Nishijima, M. Nagata, and A. Iwata, “A multi-functional cellular neural network circuit using pulse modulation signals for image recognition,” Proc. Int. Conf. on Neural Information Processing (ICONIP), pp. 613-617, Taejon, Korea, Nov. 2000.
- [6] T. Morie, M. Nagata, and A. Iwata, “Design of a pixel-parallel feature extraction VLSI system for biologically-inspired object recognition methods,” Proc. Int. Symp. on Nonlinear Theory and its Applications (NOLTA2001), pp. 371-374, Zao, Japan, Oct. 2001.
- [7] J. Harris, C. Koch, and J. Luo, “Resistive fuses: Analog hardware for detecting discontinuities in early vision,” in C. Mead and M. Ismail, editors, Analog VLSI Implementation of Neural Systems, pp. 27-55. Kluwer Academic Publishers, 1989.
- [8] P. C. Yu, S. J. Decker, H. S. Lee, C. G. Sodini, and J. L. Wyatt, Jr., “CMOS resistive fuses for image smoothing and segmentation,” IEEE J. Solid-State Circuits, vol. 27, pp. 545-553, 1992.
- [9] 永井宏昌, 坂井丈泰, 澤地利明, 松本隆, “CMOS フローティングゲート Resistive Fuse チップ,” 信学技報, NC96-113, 1997.
- [10] T. Morie, M. Miyake, M. Nagata, and A. Iwata, “A 1-D

| | |
|--------------|---|
| 搭載FPGA | Altera社EP20K400EBC652-1X(PLL搭載, $f_{clk}=(m/n \times k)f_{ref}$, $k, m, n=1, \dots, 16$) ロジックセル(ゲート)400K 動作周波数40MHz(PLLにより通信可能-100MHz), 別途発振器追加可能 |
| FPGAへのダウンロード | パネル部の専用コネクタからFPGA, EEPROMへダウンロード 動作中にFPGAへダウンロード可能(デバッグSWで切り替え) |
| FPGA使用済みリソース | ロジックセル 約19%(IP:SRAM I/F, SDRAM I/F, PLX I/F, コントローラ) RAM 0% |
| PCIボード | フルサイズ(スロットは1枚分, パネルは2枚分占有) |
| クロック用水晶発振器 | 40MHz(PLX, FPGA動作用) 別途設置可能 |
| 入出力コネクタ | ハーブピッチコネクタ100P(DX-100CV1)×2(内40Pはグラウンド) |
| 入出力 | 160P(双方向バッファ74LVTH16245A付き, 8P単位で切替), 3.3V, 2.3ns |
| SRAM | 2MB |
| SDRAM | 最大1GB(メモリスロット4本) |
| パソコン側 | Linuxドライバ付属 SRAM, SDRAMの2MB分メモリマッピング可能 SDRAMの残りの部分はバンク切り替えでアクセス |

(1) FPGAボード仕様



(2) FPGAボード機能ブロック図



(3) FPGAボード写真

図 19 FPGA ボード

ための画素並列領域抽出アルゴリズムと FPGA への実装,” 信学技報, ICD2001-42, 2001.

- [16] 大久保竜也, 安達澄昭, 岩尾博之, “顔認識技術を用いた入退室管理システム,” OMRON TECHNICS, 通巻 135 号, 2000.
- [17] 森本勝, 安達澄昭, 西村純一, “顔認識技術を用いた徘徊者保護支援システム,” OMRON TECHNICS, 通巻 137 号, 2001.
- [18] B. E. Shi, “Gabor-type filtering in space and time with cellular neural networks,” IEEE Trans. Circuits & Syst. I, vol. 45, pp. 121-132, 1998.
- [19] B. E. Shi, “2D focal plane steerable and scalable cortical filters,” Proc. Int. Conf. on Microelectronics for Neural, Fuzzy and Bio-inspired Systems (MicroNeuro), pp. 232-239, 1999.
- [20] M. Lades, J. C. Vorbrüggen, J. Buhmann, J. Lange, C. v.d. Malsburg, R. P. Würtz, and W. Konen, “Distortion invariant object recognition in the dynamic link architecture,” IEEE Trans. Comput., vol. 42, no. 3, pp. 300-311, 1993.
- [21] R. P. Würtz, “Object recognition robust under translations, deformations, and changes in background,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 769-775, 1997.
- [22] T. Morie, J. Funakoshi, M. Nagata, and A. Iwata, “An analog-digital merged neural circuit using pulse width modulation technique,” IEICE Trans. Fundamentals., vol. E82-A, no. 2, pp. 356-363, 1999.
- [23] 梅澤淳, 西島誠一, 三宅誠, 森江隆, 永田真, 岩田穆, “画像特徴抽出のためのパルス変調方式画素並列ガボールフィルタ回路,” 信学技報, ICD2001-41, 2001.

CMOS PWM cellular neural network circuit and resistive-fuse network operation,” Ext. Abs. of Int. Conf. on Solid State Devices and Materials, D-1-2, Tokyo, Sept. 2001.

- [11] H. Ando, M. Miyake, T. Morie, M. Nagata, and A. Iwata, “A nonlinear oscillator network for gray-level image segmentation and PWM/PPM circuits for its VLSI implementation,” IEICE Trans. Fundamentals., vol. E83-A, no. 2, pp. 329-336, 2000.
- [12] 安藤博士, 三宅誠, 森江隆, 永田真, 岩田穆, “抵抗ヒューズと振動子ネットワークを組み合わせた画像分割処理とその LSI 回路実現,” 信学技報, NC99-167, 2000.
- [13] H. Ando, T. Morie, M. Miyake, M. Nagata, and A. Iwata, “Image object extraction using resistive-fuse and oscillator networks and a pulse-modulation circuit for their LSI implementation,” Ext. Abs. of Int. Conf. on Solid State Devices and Materials, pp. 368-369, Sendai, Japan, Aug. 2000.
- [14] H. Ando, T. Morie, M. Miyake, M. Nagata, and A. Iwata, “Image segmentation/extraction using nonlinear cellular networks and their VLSI implementation using pulse-modulation techniques,” IEICE Trans. Fundamentals., vol. E85-A, no. 2, pp. 381-388, 2002.
- [15] 中野鉄平, 彦本里美, 森江隆, 永田真, 岩田穆, “画像認識の