

ビジョンチップの動向

松本 隆*・小林 春夫**・八木 哲也***

ビジョンチップと呼ばれているパラダイムを模式的に表現すると図1のようになる。チップ上に光センサーアレイが埋蔵されており、画像はレンズを通して直接チップ上に照射され、それにより誘起される電流あるいは電圧を入力とし、センサーのごく近くに賢く工夫されたアナログ回路のアレイを用意し、入力電流、あるいは電圧をもとに並列処理を行う。処理された画像はアレイの各ノードに分布している電圧値である。もちろん、インプレメントされるチップは各々若干趣は異なるが、基本的にはこのようなパラダイムといってよいと思う。

このパラダイムは次のような特徴をもつ：

(i) 画像の入力を完全並列に行うので、デジタル画像処理プロセッサへの入力のようないわゆる I/O ボトルネックのうち、少なくとも“I”についてはボトルネックがない。

(ii) 画像処理が完全並列で行われる。

(iii) 画像処理プロセッサはアナログなので超高速である。処理時間は、要するにトランジスタ回路の寄生容量による過渡現象の“終了”時間であるから、典型的には1~2 μ s である。

(iv) 処理時間がピクセル数に依存しない。デジタル画像処理の速度が問題になる原因は I/O ボトルネック以外に、ピクセル数によるものであることに注意を促したい。

(v) 賢く回路を工夫すれば、同様の画像処理を行う

デジタルプロセッサに比べてはるかに少ない消費電力ですみまた面積も小さくすむ。

あるパラダイムに長所があるとすれば、対応する短所もあり、ビジョンチップも例外ではない：

(i) アナログ回路であることによる精度の問題。

(ii) プログラマビリティの問題：画像処理といっても、アナログ回路の過渡現象であるから、目的に応じて各々アナログ回路を工夫せねばならず汎用は難しい。

(iii) 記憶容量が小さい。

(iv) 目的に応じたよいアナログ回路を設計する時間が長い。

これらの長所、短所を考慮すると、ビジョンチップがカバーできる“守備範囲”は A/D 変換を行う前に遂行できる可能性のより高い、いわゆる初期視覚(early vision)アルゴリズムと呼ばれるものである場合が多い。すなわち、画像処理アルゴリズムの中でも、比較的初期レベルでの処理を行うことを目的とするものが多い。これは、例えば、コントラスト強調、平滑化、エッジ検出、動き検出、ステレオ等である。したがって、生体系の視覚情報処理アルゴリズム/アーキテクチャーを参考にしているものも多い。ビジョンチップの最初の提唱者 Mead²⁾の多くのチップもそのようなものであり、脊椎動物の網膜や蝸牛等にヒントを得ている。ただし、生体系とはまったく関係なく、純粋に工学的アルゴリズム/アーキテクチャーに基づくチップもいくらかでもあることに読者の注意を促したい。

Recent trends in vision chips (1996年1月8日受理)

*Takashi MATSUMOTO 早稲田大学理工学部電気電子情報工学科 (〒169 東京都新宿区大久保3-4-1), ** Haruo KOBAYASHI テラテック(株), ***Tetsuya YAGI 九州工業大学情報工学部制御システム工学科

1. 具体的チップの紹介

光センサーアレイでは、最も多く用いられているのが CMOS parasitic photo transistor と呼ばれるものであ

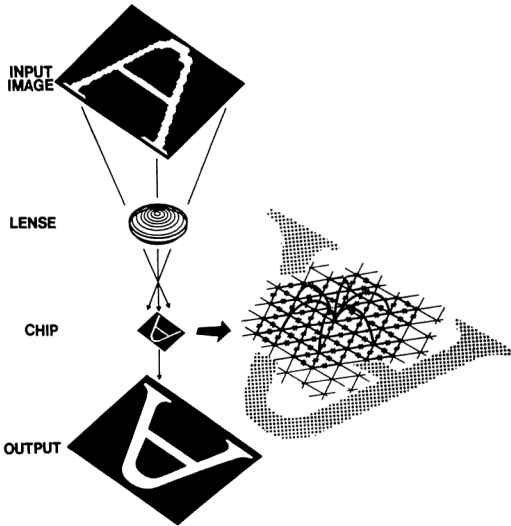


図1 ビジョンチップ概念図¹⁾ (© IEEE).

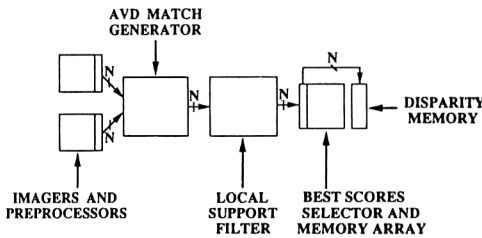


図2 視差検出チップ⁵⁾ (© IEEE).

る。よく用いられる理由は標準 CMOS プロセスで容易にインプレメントできるからである²⁾。しかるべきテクノロジーを備えている人々にとっては、これ以外の入力メカニズムはいくらでも考えられる。例えば、Takanagi ら³⁾の CMD (charge modulation device), Lange ら⁴⁾の GaAs を用いたもの等が考えられる。処理部分を CCD で行うなら、CCD センサーでももちろん可能である。

上に述べたセンサーに直接あるいは、比較的近接して画像処理回路が配置され、センサーから誘起された電流または電圧をもとに処理が行われる。これは多くのビジョンチップの心臓部であるといってよく、各処理目的に応じてよいアナログ回路を設計することがポイントである。

1.1 ステレオ

2つの入力画像からの視差 (disparity) 検出は重要な初期視覚問題のひとつであって多くのアルゴリズムがある。Hakkarainen らは Marr-Poggio-Drumheller アル

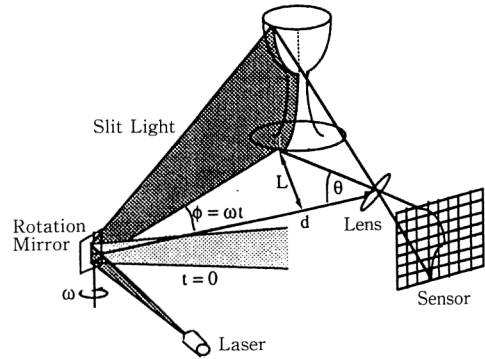


図3 レンジファインダー⁶⁾ (© IEEE).

ゴリズムと呼ばれるものを CCD と CMOS で実現した⁵⁾。

ステップ (i) 画像 (40 × 40 ピクセル) に適当なスケールで $\nabla^2 G$ 様フィルターをかける。

これは CCD 入力から直接

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

を繰り返すほどこした後

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

を一度ほどこして実現している。前者は Gaussian でなく binomial といわれるフィルターなので、正確には $\nabla^2 B$ フィルターと呼ぶべきかもしれない。

ステップ (ii) 各ピクセルについて、適当な大きさの近傍同士で上記 $\nabla^2 G$ 様フィルターの結果に対して対応をとる。対応をとるには何らかの距離が定義されている必要があり、この場合、両者の差の絶対値にしてあり、この回路設計が要点のひとつである。

ステップ (iii) 視差には“連続性”があると仮定し、適当なコンボリューション (低域フィルター) をかけて結果を出す。CCD を用いているし、各ピクセルごとに対応をとるプロセスが入るので、並列アナログではあるがクロックを必要とする。視差レンジ 10 ピクセルに対して 880 frame/s を実現している。図 2 にアーキテクチャーを示す。参考のため、ほぼ同じアルゴリズムを 65536 プロセッサコネクッションマシンで行うと 256 × 256 ピクセルに対して数秒かかる。

1.2 シリコンレンジファインダー⁶⁾

スリットを通したレーザー光反射を 2つのフォトセン

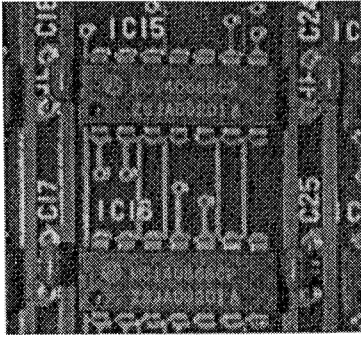


図4 プリント基板とキャパシター⁶⁾(© IEEE).

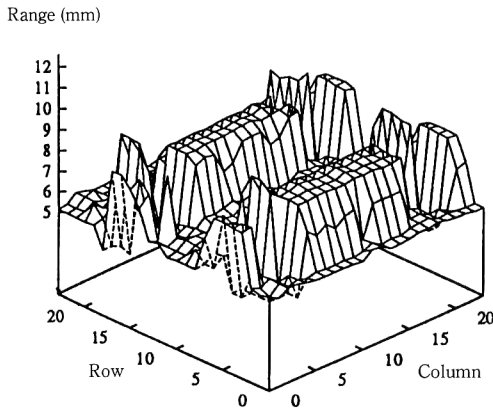


図5 Elevation map⁶⁾(© IEEE).

サーアレイで集め、3角測量の原理でレンジデータを得ている(図3)。24×24ピクセルをビデオフレームレートで実現した。このチップはいわゆるアクティブセンサーであり、ここで紹介する他のすべてのチップは、この意味でパッシブである。また、レンジデータなので、対象とする物体表面に、例えば文字や絵が書いてあったり汚れがあったりしても、それらと関係なくデータが得られ、目的によっては大きな利点となる。図4はプリント基板のICとキャパシター、図5はそのelevation mapである。

1.3 $\nabla^2 G$ フィルターのzero-crossings

$\nabla^2 G$ フィルターのzero-crossingはある仮定のもとでエッジに対応するといわれている。Trembrayらは、六角状フォトセンサーから、15近傍コンボリューション

$$\frac{\pi}{3} \left[\frac{3}{4} Fa(0)I_0 + \sum_{r=1}^{15} rFa(r)Ir \right]$$

を行うチップを実現した⁷⁾。ここに

$$F_{\sigma}(r) = \nabla^2 G_{\sigma}(r)$$

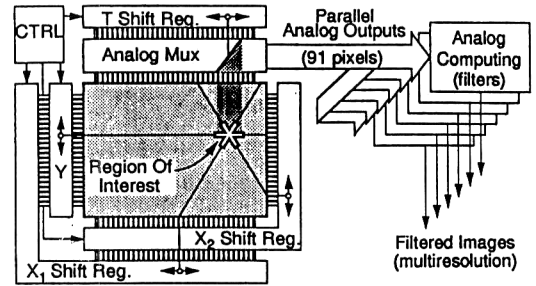


図6 $\nabla^2 G$ zero-crossings アーキテクチャー⁷⁾(© IEEE).

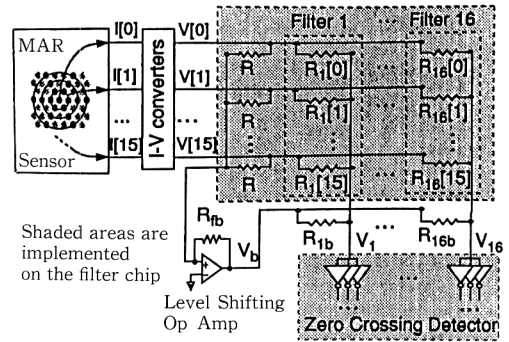


図7 $\nabla^2 G$ の部分⁷⁾(© IEEE).

である。ただし、フォトセンサー配列が六角形なので circular symmetry にすぐれており、15近傍の意味は注目点から15ピクセルの距離にあるという意味で、近傍数はそれよりはるかに多い。負性抵抗をさけるため2つのスケールの正抵抗回路網のコンボリューションの差をとっている。このチップも各ピクセルごとにフィルターを行うので、クロック(40MHz)を必要とする。これは、Shimmiら⁸⁾がIIR(infinite impulse response)構造をしているのに対して $\nabla^2 G$ 様フィルターのFIR(finite impulse response)版といえよう。フィルタースケールも(マニュアルで)変更可能である。図6は全体のアーキテクチャー、図7は $\nabla^2 G$ の部分である。

1.4 動画像圧縮イメージセンサー⁹⁾

条件付画素補充とは、時間方向の変化が大きい画素を選択し符号化する方式である。すなわち、入力画像は各画素ごとにメモリー中の一時点前の復号画像と比較され、その変化が閾値以上である場合に画素値が出力される。また、このときメモリー中の復号画像も更新される。この方式は、読み出すアドレスの制御を除けば各画素ごとに独立した構成とできるため、センサー面において並列性を生かすのに適する。このような回路を32×

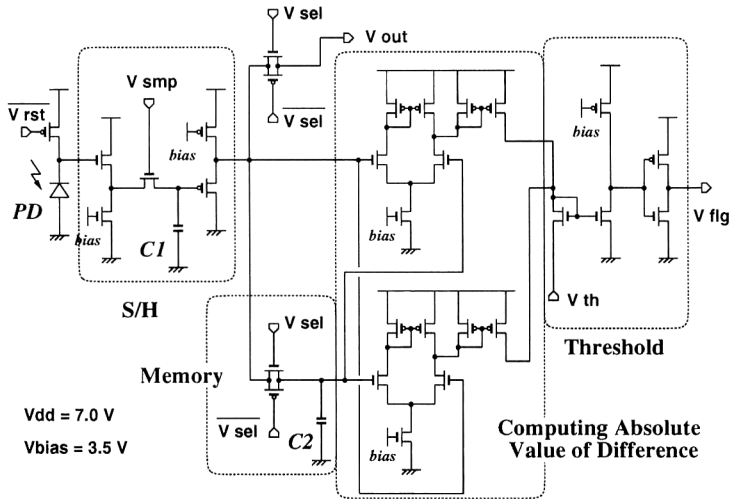


図8 動画像圧縮センサー⁹⁾(©計測自動制御学会).

32画素をポリ1層, メタル2層の2 μm CMOS プロセスにより試作した. 1画素の大きさは170 \times 170 μm^2 であり, そのうち開口部は20 \times 20 μm である. センサー全体ではおよそ7mm角である(図8).

1.5 その他

上述以外にはNabetら¹⁰⁾, 坂井ら¹¹⁾, Kobayashiら¹²⁾, Langeら⁴⁾等がある. 序文でビジョンチップ=光センサー+アナログ回路アレイと説明し, このカテゴリーに属するチップを紹介したが, このカテゴリーには属さないが, 光センサーアレイ埋蔵画像処理チップという意味では同様の目的をもつチップをひとつ紹介しておきたい. 石川¹³⁾らは60 \times 60フォトセンサーアレイに近接して, 同じ数のデジタル処理プロセッサを配置し, しかもチップサイズ8 \times 8mm²で実現した. 当然ながら, デジタル回路は思い切った単純化を行っており(図9), 各ピクセルは各々単純なA/Dコンバーターを組み込んである.

2. 生体の視覚から直接ヒントを得たチップ

前節では, 初期視覚情報処理の主として工学的アルゴリズムを様々な形で組み込んだチップを紹介した. ここでは, 生体の視覚メカニズムをより忠実に再現することを試みたチップについてもう少し掘り下げてみる.

生体の視覚メカニズムをできるだけ忠実に再現するチップを作ること自体が, 生体の視覚機能と構造を構成的な方法で解析するという意味からも興味深いことであるが¹⁴⁾, 工学的な視点からみてより重要なことは, 生体の視覚神経系のように柔軟で知的な視覚システムの実現を

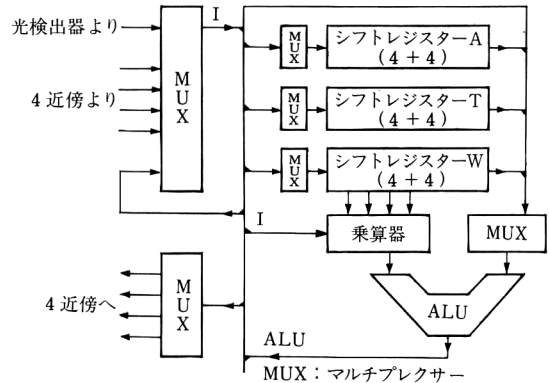


図9 デジタルビジョンチップ演算部¹³⁾(©計測自動制御学会).

目指すことにあるだろう.

2.1 順応機構をもった光センサー

網膜の視細胞は単に光の情報を電気信号に変換するのみでなく, 順応というきわめて合理的な機能を有している(図10). 図では横軸に光強度を対数値で, 縦軸には視細胞(錐体)の応答を規格化したものをプロットしてある. 視細胞は3~4 log程度の広いダイナミックレンジを有することがわかる. 当初Meadが設計した視細胞回路では, フォトダイオードにMOSトランジスタを直列につなぎ光強度の対数に応答する光センサー回路を実現していた²⁾. しかしこの広いダイナミックレンジのゆえに, 光強度の変化に対する応答ゲインはどうしても低くなり, MOSトランジスタの特性のばらつき等を考慮すると不具合が生じる. 一方, 視細胞ではこの応

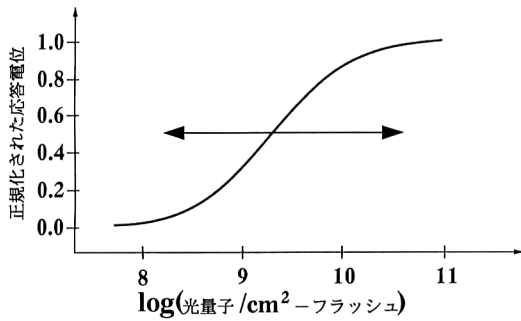


図10 網膜の視細胞（錐体）の応答曲線。視細胞を様々な光強度のフラッシュ光で刺激したときの応答振幅の値をプロットした。応答振幅は最大光強度で刺激したときの応答で正規化してある。一定の背景光をあたえたと、フラッシュ光に対する応答曲線は平衡移動する。文献24)のデータを修正した。

答ゲインを稼ぐために光順応が起こる。すなわち、ある一定の光刺激が定常的に視細胞に与えられると動作点がこれに準じて移動する。先の図でこの順応を表現すると、定常的に与えられた背景光の上に重ねて刺激光を与え、この刺激光に対する応答の大きさをプロットすると、背景光の強度をパラメータとし、グラフが水平方向に平行移動することになる。背景光の強度が強ければグラフは右方向に移動し、より強い光の領域にダイナミックレンジをもっていくことになる。

このような順応特性に習った光センサーがいくつか開発された¹⁵⁻¹⁷⁾。これらの光センサーにはバンドパスあるいはハイパスの時間特性が埋め込まれている。図11にMahowaldにより開発された順応型光センサーを示す¹⁵⁾。フォトトランジスタを流れる電流は、Q1によ

り供給される。 V_{out} はQ3, Q4およびC1を介してキャパシター C_F の電圧としてQ1のゲート電圧にフィードバックされている。このフィードバックの働きで、ある一定の明るさの背景光がある場合には、ゆっくりとした時定数で C_F がチャージされ、 V_{out} は順応した新しい定常レベルまで徐々にシフトする。この順応により4 log 以上の広い範囲で、光強度の変化を高感度でとらえることができる。

2.2 受容野の順応

網膜の順応機能は、初段の光センサーのレベルのみにとどまるものではない。光センサー細胞である視細胞から情報をうけとる信号処理回路の段階においても、巧妙な順応機構が存在する。この順応の例として、 $\nabla^2 G$ 様フィルターのスケールの調節機構についてふれよう。 $\nabla^2 G$ 様フィルターは、生体の視覚系では2次ニューロンである双極細胞において初めて発現する。最近の研究から、この双極細胞のフィルターのスケールは、網膜の順応状態に応じて変化するのではないかという実験的証拠が得られている¹⁸⁾。すなわち網膜が暗い環境におかれたときは、フィルターのスケールを大きくし、空間解像度を犠牲にして空間的な平滑化の度合を上げる。反対に明るい環境におかれたときはフィルターをよりシャープにし解像度を上げる。この機構は昼間のような明るい環境と、夕方のような薄暗い環境における画像の信号対雑音比に応じて、処理特性を変化させることに対応すると考えられる¹⁹⁾。

このような順応機構をもったチップのアーキテクチャーが提案されている¹²⁾ (図12)。実際の網膜では、 $\nabla^2 G$ のスケールの調節は脳からの遠心性の信号により行われ

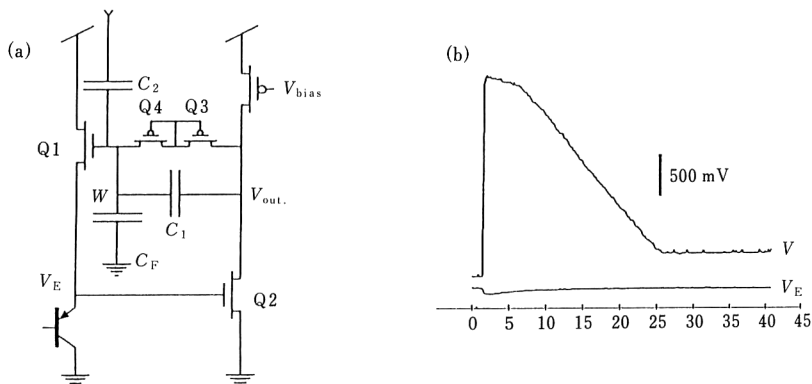


図11 順応機構をもった光センサー。回路図(a)と入出力特性(b)。(b)の上段は光センサーからの出力、下段はフォトトランジスタのエミッタ電圧(© Kluwer Academic Publishers)。

ていると考えられている。このアーキテクチャーでは、 $\nabla^2 G$ 様フィルターのスケールはチップ内の光センサーの応答により自動的に実行される。

受容野のスケール調節機能は、上記の順応の意味だけでなく視覚の様々な局面で必要とされると考えられる。生体の視覚には、異なる空間周波数のバンドにチューニングされたいくつかのチャンネルがあると考えられ²⁰⁾、このチャンネルは輪郭の知覚、奥行き知覚等に付随する不良設定性を回避するときうまく利用でき、チップにも応用されている¹⁵⁾。今後さらに研究する余地が残された課題と思われる。

2.3 動きの方向を検出するチップ

生体の視覚系には、刺激の光がある方向に動いたときに反応し、逆方向の動きには反応しない細胞がよくみつかれる。これを方向選択性ニューロンとよぶ。方向選択性ニューロンの神経回路は、特に昆虫など下等動物において解析が進んでいる。昆虫において見つかった神経回路を参考にしたチップがインプレメントされている²¹⁾。図13に方向選択性を発現させる基本的な回路の構成を示す。方向選択性には時間遅れ回路が用いられている(図13(b))。Pは光センサー細胞であり、この細胞は2次ニューロン(M)に対し直接入力し、隣接するMに方向性をもって時間遅れ信号を送る。図ではPが右隣のMに対し時間遅れを介して入力特性を与える。2次ニューロンMは非線形の入出力をもつ。この非線形特性は、乗算、整流等の回路が考えられる。この結果として、光刺激が左から右にPの間隔と遅れ時間に見合った速度で移動したときには、細胞の反応は大きくなる。これは結局空間的に分布したPの応答の相関をとっているのだから、correlation-motion-detectorと呼ばれる。このチップでは、方向検出回路を120度ずつ異なる3方向にとったものを1ピクセルとし、これを6角状に配列し速度検出を行っている(図13(b))。

2.4 チップ間の連絡

生体の視覚には、様々な“最適刺激”にチューニングされた細胞群が存在しているので、これらの機能を真似たチップもバラエティに富む。では、このような機能をどうやって統合し、生体のような柔軟な視覚システムを構築するのか？ 当然異なる機能をもった回路を単一のチップ上インプレメントすることは不可能である。そこで様々な機能を、異なるチップ上にインプレメントし、これを連絡する方法を考えなければならない。チップの出力ピンの数はきわめて限定されているので、異なるチップの対応するピクセル同士を結ぶわけにはいかない。

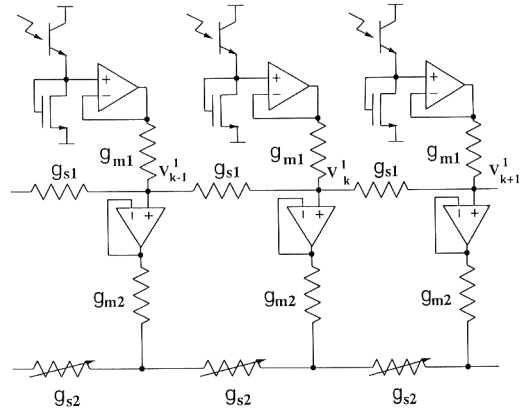


図12 受容野のスケールを順応的に変化させるチップのアーキテクチャー。光センサーからの出力 V_k を用いてピクセル間の電気的結合 G_{s2} を変化させる。

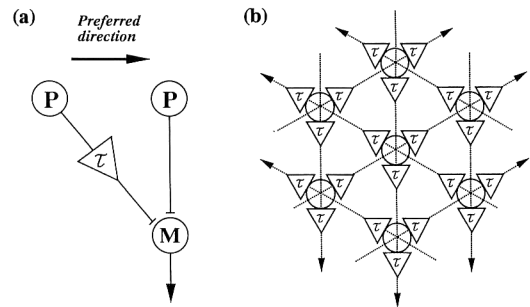


図13 方向選択性をもった回路。(a)は方向選択性をもたせるための基本回路。光センサー素子の応答は2次素子のMに送られる。(b)はチップのピクセル配列。円は光センサー素子と2次素子からなるピクセル。矢印の3方向に方向選択性をもっている。

かつビジョンチップは、与えられた画像から冗長な情報をフィルタリングし、後段につづくより高次の処理に必要な情報を優先して送るのがひとつの大きな役目である。このような背景を考慮すると、現在のところ最も有力視されているのが、address-event representationである¹⁵⁾。ここでは、異なるチップを少数のバスで結び、送り手のチップからはピクセルの住所付きで受け手のチップにインパルス列として情報を送る。バスは多くのピクセルの出力を共有するが、通常要求される視覚処理の範囲では、単一のピクセルから送るべきインパルスの頻度はきわめて低く、異なるピクセルの情報が混信する確率は低く抑えることができると考えられている。その他様々な方法でのチップ間の連絡方法が試されている。今後この領域での発展が、ビジョンチップの応用の可能性

を広げる上できわめて重要な要素であるのは間違いない。

筆者ら^{22,23)}による関連するサーベイもある。

文 献

- 1) T. Matsumoto, H. Kobayashi and Y. Togawa: "Spatial versus temporal stability issues in image processing neuro chips," *IEEE Trans. Neural Networks*, **3** (1992) 540-569.
- 2) C. Mead: *Analog VLSI and Neural Systems* (Addison-Wesley, Reading, MA, 1989).
- 3) I. Takayanagi, T. Isokawa and F. Nakamura: "A multiple output CMD imager for real-time image processing," 1993 IEDM Dig. Tech. Dig. 22.5.1 (1993) pp. 579-582.
- 4) E. Lange, E. Funatsu, J. Ohta and K. Kyuma: "Direct image processing using arrays of variable-sensitivity photodetectors," *ISSCC Dig. of Tech. Papers* (1995) pp. 228-229.
- 5) J. M. Hakkarainen and H. S. Lee: "A 40×40 CCD/CMOS absolute-value-of-difference processor for use in a stereo vision system," *IEEE J. Solid-State Circ.*, **28** (1993) 799-807.
- 6) K. Sato, A. Yokoyama and S. Inokuchi: "Silicon range finder—a real time range finding VLSI sensor—," *Proc. of CICC* (1994) pp. 339-342.
- 7) M. Trembray, M. d'Anjou and D. Poussart: "Hexagonal sensor with imbedded analog image processing for pattern recognition," *Proc. of ICCV* (1993) pp. 12.7.1-12.7.4.
- 8) T. Shimmi, H. Kobayashi, T. Yagi, T. Sawati, T. Matsumoto and A. Abidi: "A parallel analog CMOS signal processor for image contrast enhancement," *Proc. European Solid State Circuits Conference*, Sept. (1992).
- 9) 相澤清晴: "動画像圧縮イメージセンサ", 第15回光応用計測部会講演会 (1995) pp. 23-28.
- 10) B. Nabet, R. B. Daring and R. Pinter: "Implementation of front-end processor neural networks," *Neural Networks*, **5** (1992) 891-902.
- 11) 坂井丈泰, 澤地利明, 松本 隆: "Weak String フィルタとその MOS トランジスタによる実現", *テレビジョン学会* 技術報告, IPU 94-65, IDY 94-168 (1994) pp. 13-18.
- 12) H. Kobayashi, T. Matsumoto, T. Yagi and K. Tanaka: "Light-adaptive architectures for regularization vision chip," *Neural Networks*, **8** (1995) 87-101.
- 13) 石川正俊: "超並列・超高速ワンチップビジョンとその応用", *計測自動制御学会第15回光応用計測部会資料* (1995) pp. 33-37.
- 14) R. Douglas, M. Mahowald and C. Mead: "Neuromorphic analogue VLSI," *Annu. Rev. Neurosci.*, **18** (1995) 255-281.
- 15) M. Mahowald: *An Analog VLSI Systems for Stereoscopic Vision* (Kluwer Academic Publishers, London, 1994).
- 16) T. Delbruck and C. Mead: "Time-derivative adaptive silicon photoreceptor array," *Proc. SPIE*, **1541** (1991) 92-99.
- 17) J. Mann: "Implementing early visual processing in analog VLSI: light adaptation," *Proc. SPIE*, **1473** (1991) 128-136.
- 18) T. Shigematsu and M. Yamada: "Effects of dopamine on spatial properties of horizontal cells in the carp retina," *Neurosci. Res., Suppl.* **8** (1988) s69-s80.
- 19) 八木哲也, 大島成通, 舟橋康行: "初期視覚における網膜双極細胞の機能について", *電子情報通信学会論文誌 D-II*, **J78-D-II** (1995) 1123-1133.
- 20) H. R. Wilson and J. R. Bergen: "A four mechanism model for threshold spatial vision," *Vision Res.*, **19** (1979) 19-33.
- 21) T. Delbruck: "Silicon retina with correlation-based, velocity-tuned pixels," *IEEE Trans. Neural Networks*, **4** (1993) 529-541.
- 22) 松本 隆, 小林春夫, 八木哲也: "ビジョンチップ[1]—アナログ画像処理用ニューロチップ—", *電子情報通信学会誌*, **76** (1993) 783-791.
- 23) 松本 隆, 小林春夫, 八木哲也: "ビジョンチップ[2]—アナログ画像処理用ニューロチップ—", *電子情報通信学会誌*, **76** (1993) 851-858.
- 24) G. Fain and J. E. Dowling: "Intracellular recordings from single rods and cones in the mudpuppy retina," *Science*, **180** (1973) 1178-1181.