



情報システム概論 (8 / 4) ~ VisionChipとその周辺

秋田純一



ここにいたる経緯(研究歴)

- 卒論:「三次元容量計算における誤差分析」
 - LSI内の配線間容量の数値計算に関するもの
- 修論:「確率遷移モデルに基づくCMOS論理回路の消費電力低減手法」
 - CMOS論理回路の動作モデルとLow Power化
- 博士論文: "A Study on High Speed and Low Power Image Sensors with Variable Resolution Scan using Tree Structure of Images"
 - 「イメージセンサ」(= 撮像素子)の研究



イメージセンサの道へ…

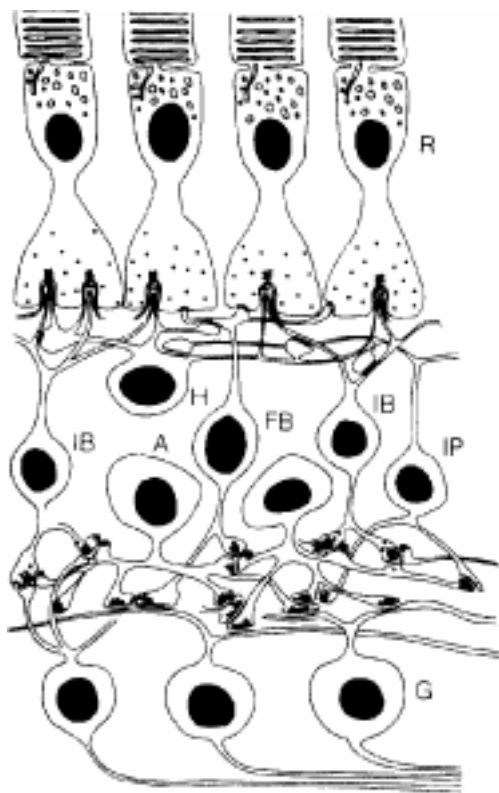
- [M1～M2] CMOS回路の(動的)消費電力を減らす研究
- [M2の終わりごろ]「文部省の重点領域の予算(数億円/年)がつきそうなんだけど、D1からこのテーマをやらないか？」
 - テーマ名「知的低電力デバイス・回路方式」
 - 「VisionChipというやつで、2次元を二分法に似た方式で走査したら低電力にならないか？」というアイディア



「画像」の世界との出会い

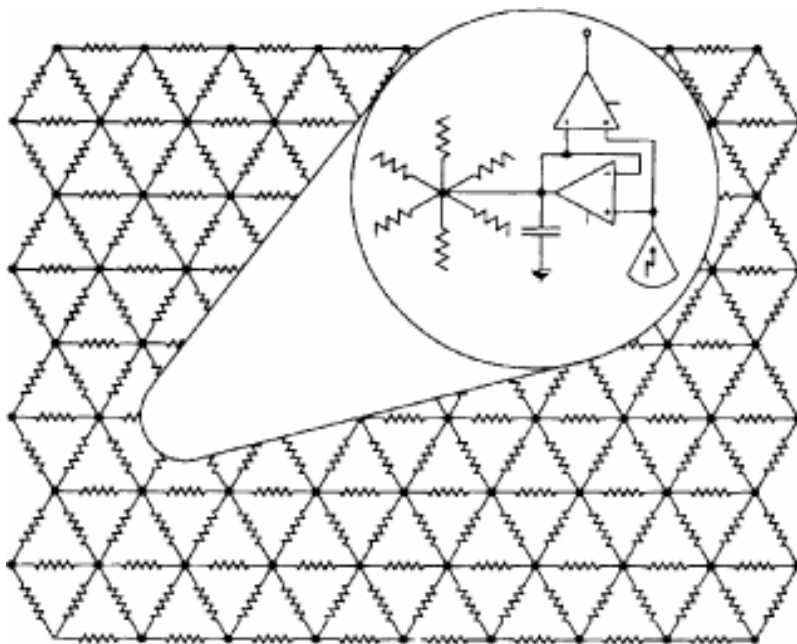
- 比較的なじみやすい概念
 - 並列処理
 - アルゴリズム・論理演算
- 知らない専門用語
 - Segmentation, digitize, neighbor, ...
- そもそも親しくない概念
 - 2次元の画素配列
 - 光電変換

VisionChipという概念



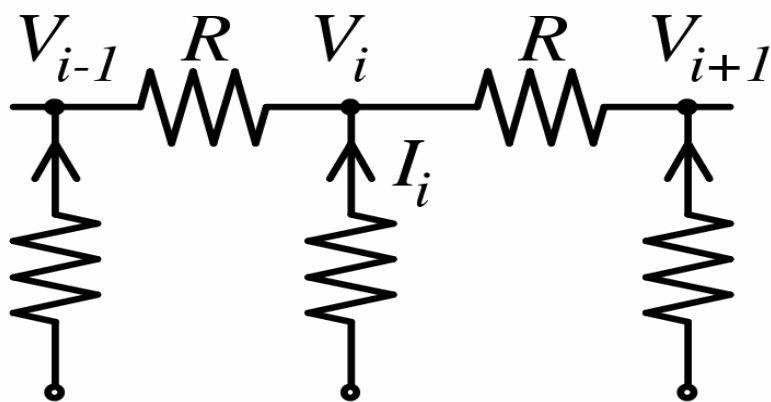
- [狭義] 網膜模倣イメージセンサの総称
 - Vision Chip
 - Silicon Retina
 - Smart Sensor
- [広義] 受光系と信号処理系を集積し、並列処理を特徴とするイメージセンサ

VisionChipの歴史



- C.Mead(1989),
“Analog VLSI and
Neural Systems”
 - 網膜の水平細胞を抵抗網で模擬
 - 受光素子からの光電流を注入
 - 抵抗網に発生する電位分布 = 平滑画像

C.MeadのVisionChipの原理

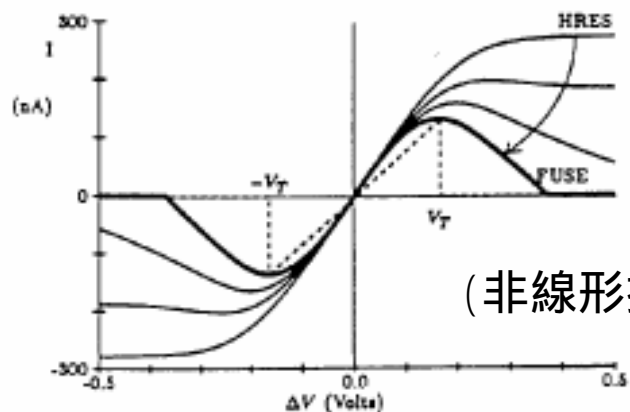
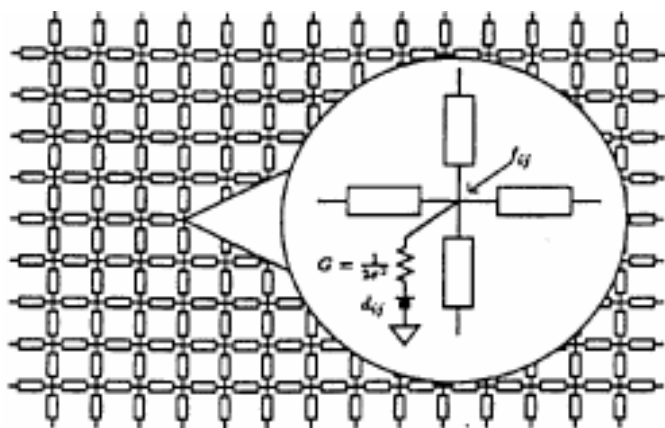


$$2V_i - V_{i-1} - V_{i+1} = RI_i$$

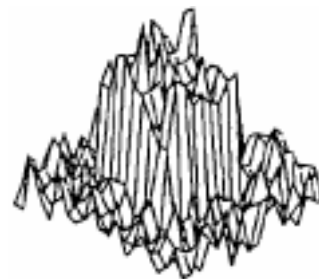
$$\rightarrow \nabla^2 V = RI$$

- 電位分布(V_i)
= 元画像(I_i)の空間2階微分
- 「平滑化をした画像」が得られる

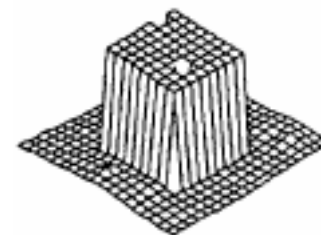
抵抗網の改良型VisionChip



(非線形抵抗)



(線形抵抗網)



(非線形抵抗網)
ノイズ除去
+
エッジ強調



C.MeadのVisionChipの問題点

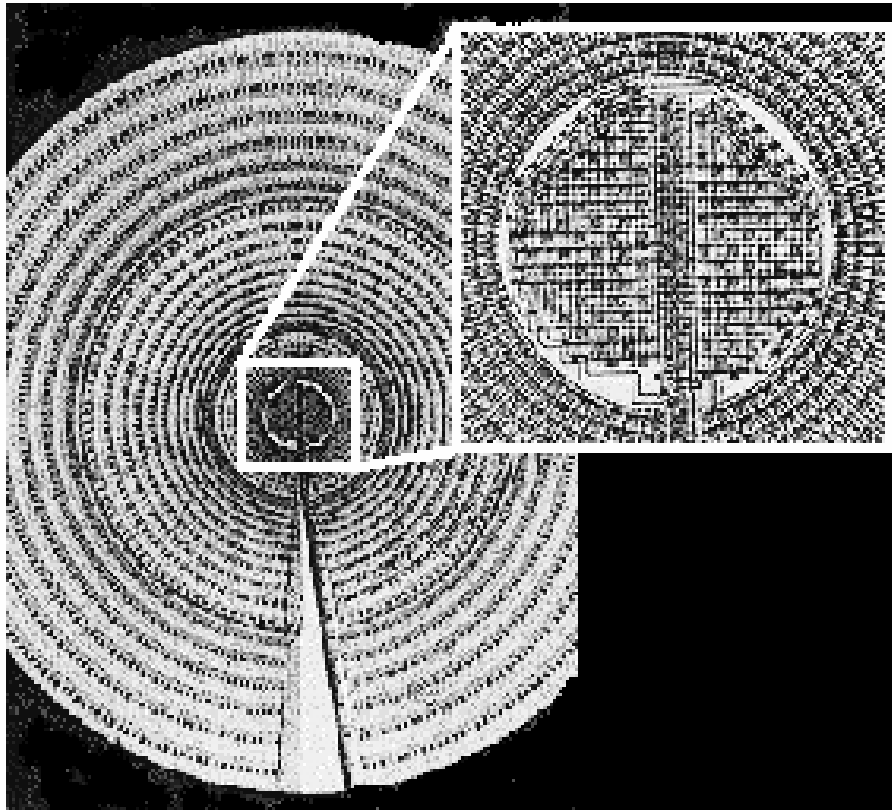
- 抵抗値のバラツキ = 画像のバラツキ
 - 実際に作ってみるとシャレにならない
 - 抵抗のバラツキを減らす回路構成 / バラツキの自己補正機構 [八木@阪大他]



動き検出VisionChip

- 2次正則化 + Euler-Lagrange方程式をアナログ回路網で求める方法
- 生体の網膜での信号伝播モデルをアナログ回路で模擬(主にニューラルネット畑)

中心窩VisionChip



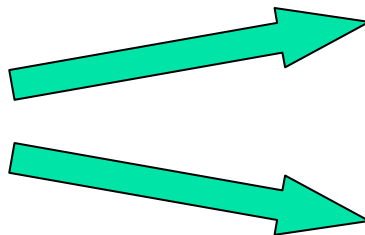
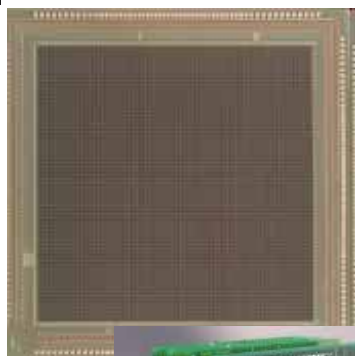
- 網膜の受光細胞の空間分布を模擬
 - 中心ほど解像度が高い
 - (信号を読み出すだけで演算機能はない)

製品化されたVisionChip

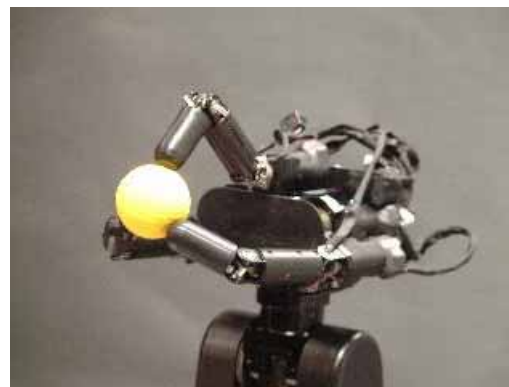


- 三菱電機M64270P他(通称「人工網膜チップ」)
- 100 × 100画素程度の撮像
- エッジ検出・平滑化などの処理後の画像も出力可

VisionChipの転換点:SPE



バッティング



ボールキャッチ

SPE (Sensory Processing Element)
(通称 "1msビジョン") (石川@東大)

- ビットシリアル演算
- SIMD型プロセッサ
- 超並列 (Massive Parallel) アーキテクチャ
- Self-Window法 (独自アルゴリズム) など



SPEのポイント

- 超並列

- PE (Processing Element)を本気で画素ごとに並べてしまった

- デジタル処理

- プログラムが可能 = 柔軟性

- 「1ms」という割り切り

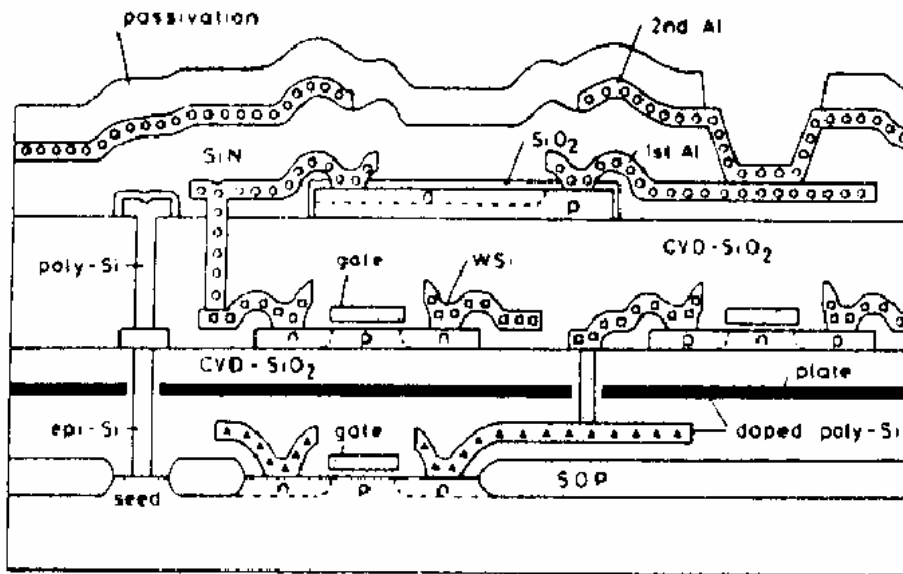
- Machine Visionであれば、1msより速くてもあまり意味がない(物体はそこまで速く動かない)
- 「時間のかかる」ビットシリアル演算回路の多用
= 回路規模の縮小 = 高解像度化・高開口率化



VisionChipの普及しない原因？

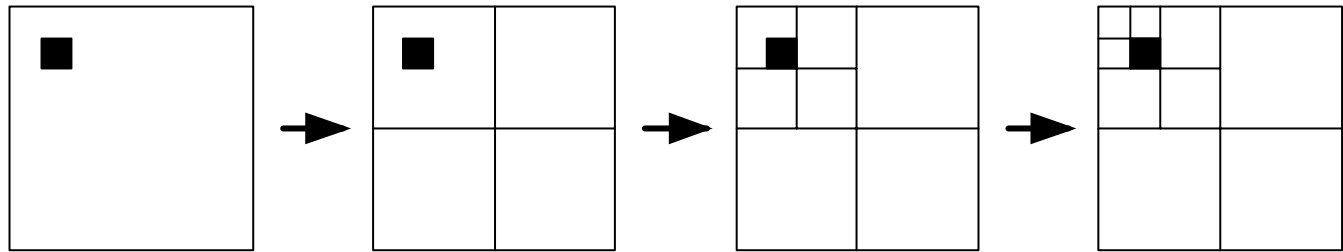
- ビデオカメラの普及・高性能化
 - 大半のことは、カメラ + PCでこなせるようになった
 - 高速カメラにも、PC側を一部並列化で対応
- 「プログラムできない」
 - 抵抗網型は機能が固定
 - SPEもかなり特殊なプログラム
- 原理的に解像度が低い

3次元LSIとVisionChip



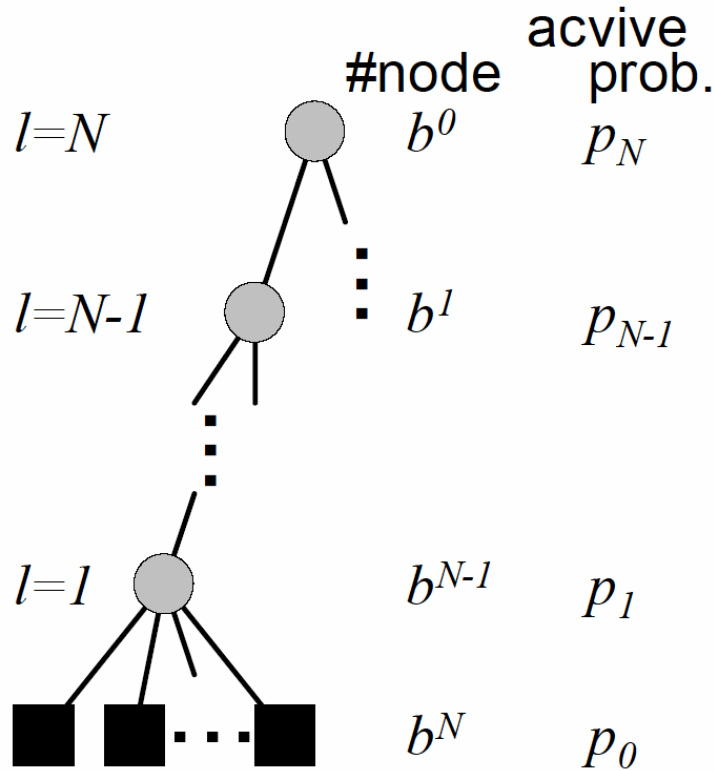
- 積層型の3次元LSI
 - 上層と下層を上下に結ぶ信号線
- VisionChipに最適
 - 最上層 = 受光素子
 - 受光素子 - 処理回路間が直結
 - (A/Dコンバータを作った例がある)

4進木走査とVisionChip



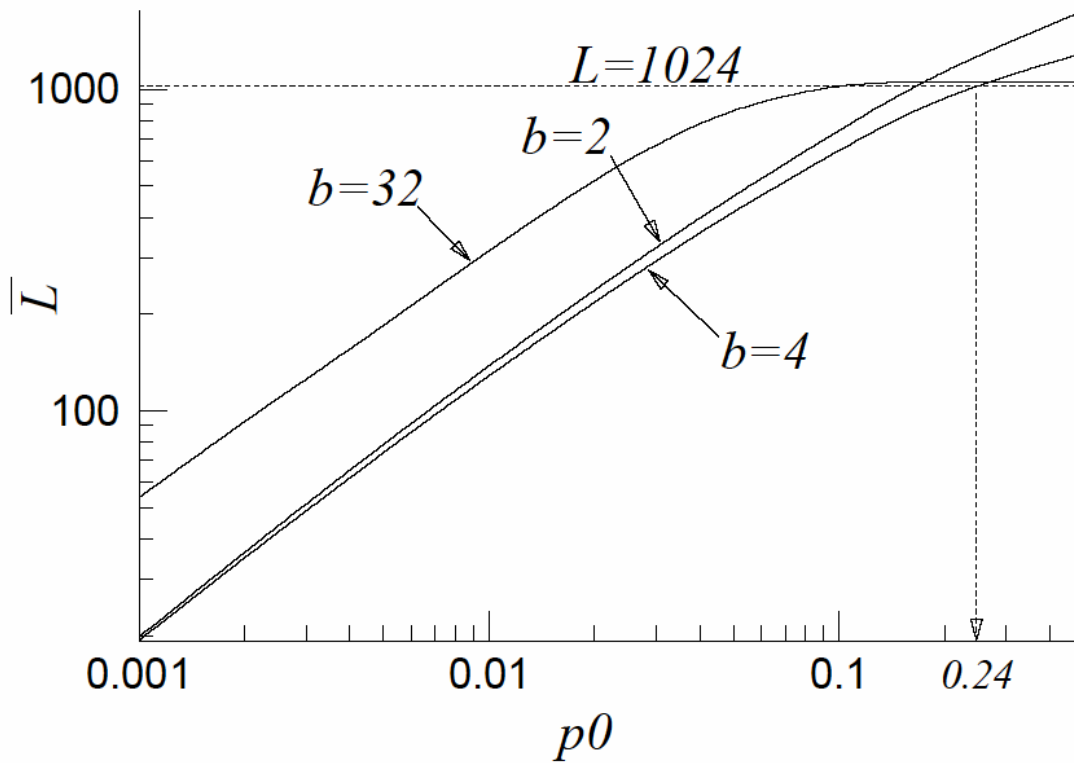
- 「二分探索」の二次元版
- 次第に領域を絞っていく(解像度を高めていく: Progressive JPEGのような感じ)

4進木走査とのモデル化



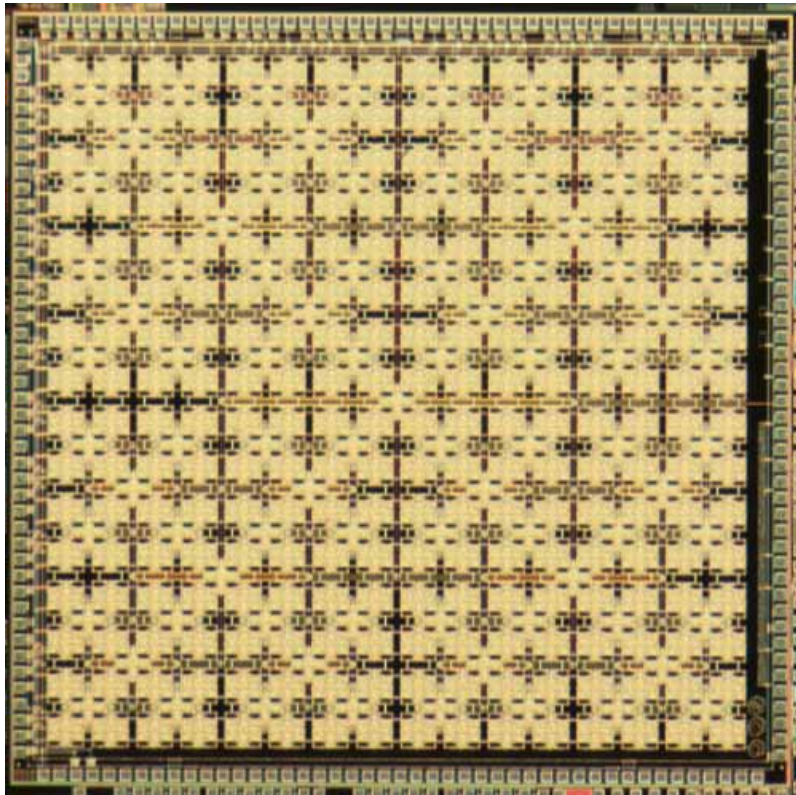
l : 木構造のレベル
 b : 分岐数
 N : 木構造の深さ
 p_l : レベル l のノードが1となる確率
 p_0 : 「画素」が1である確率
 \bar{L} : 読み出しステップ数の期待値
 $p_l = 1 - (1 - p_{l-1})^b$
 $p_l = 1 - (1 - p_0)^{b^l}$
 $\bar{L} = 1 + \sum_{l=1}^N b^{N-l+1} p_l$
 $= 1 + \sum_{l=1}^N b^{N-l+1} \{ 1 - (1 - p_0)^{b^l} \}$

4進木走査ステップ数の期待値



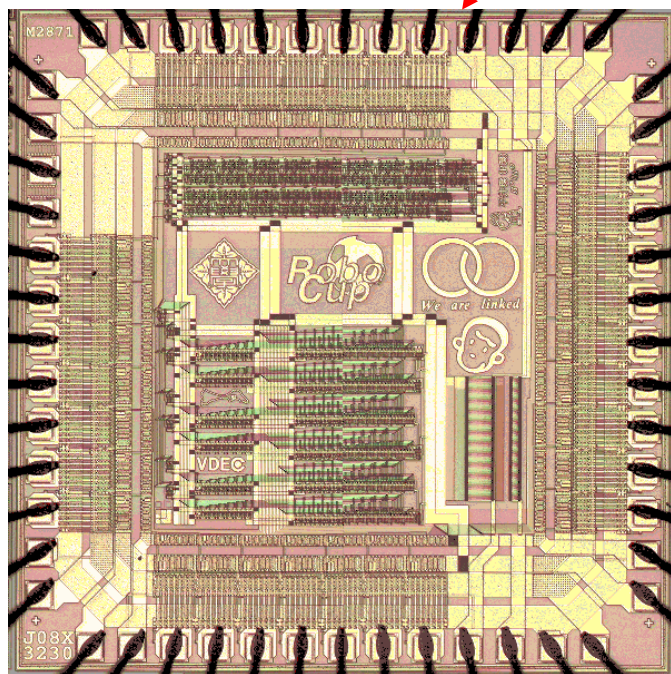
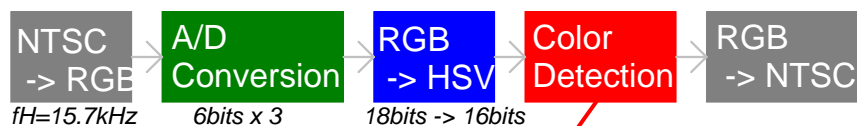
$$\bar{L} \propto O(\log N)$$

4進木走査イメージセンサ



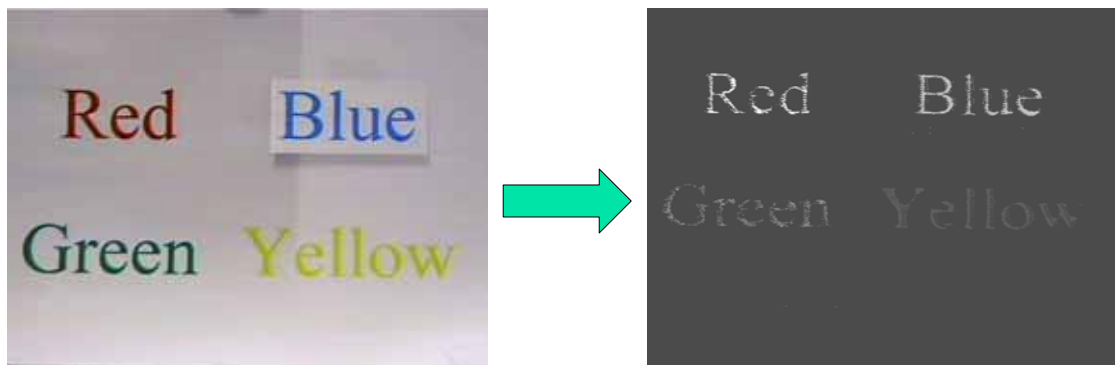
- 4進木構造をそのままレイアウト
 - 上位層ほど面積が大きい = バッファが大きい
 - 探索対象範囲のみにクロック供給 = 低電力

ちょっと寄り道: カラー画像処理



- 色抽出処理のみ
 - 対象色のみ抽出
- あえてビデオ信号を対象
 - カラー化のため...

色抽出処理の結果



- 400ns分の遅れのみ
- 時系列のまま画素信号の処理
 - フレームメモリが不要
- 「コーヒーフィルタ」のような情報処理



本当に必要な画像の情報とは？

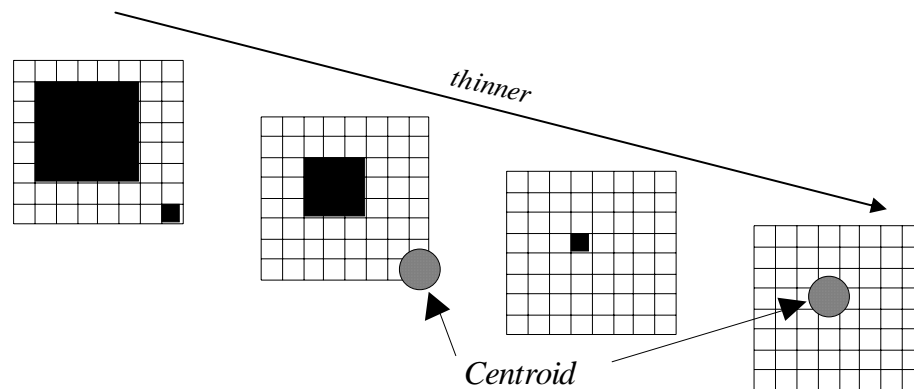
- 「画像」は、本来人間が見るもの
 - テレビ、ビデオ、・・・
 - 「画像」の意味は、人間が「理解」する
- 「理解」とは・・・？
 - 画像の意味(特徴量)を抽出する
 - Machine Visionでは、何らかの特徴量抽出が必要



画像の「特徴量」

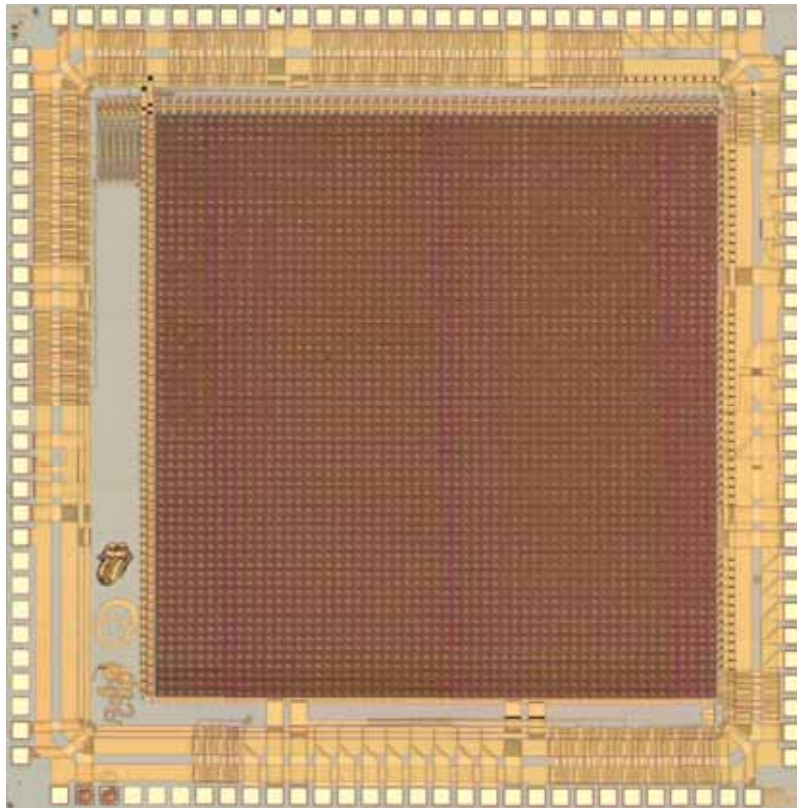
- 用途によっていろいろ…
- 「位置」の情報は…？
 - 「注視」の機能 = 対象領域を絞る前処理
 - 注視の後で、必要領域を詳細処理
= さまざまな処理の「前処理」
- 「ぱっと見」センサ…？
 - 実はあまり精度は必要ない(注視なので)
 - 「高精度化」とは逆行するパラダイム…？

オートマトンによる物体検出



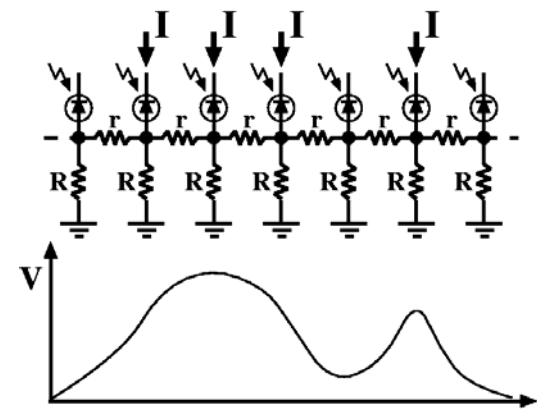
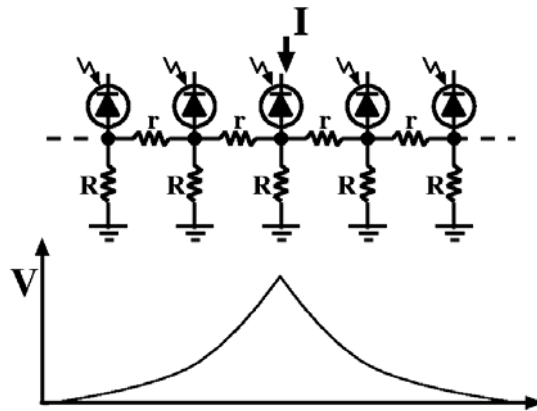
- 二値画像の収縮処理
- 最後に消えるところを「中心」とみなす
- セルオートマトンによる収縮処理の並列化

オートマトン遷移型物体検出 VisionChip (試作)



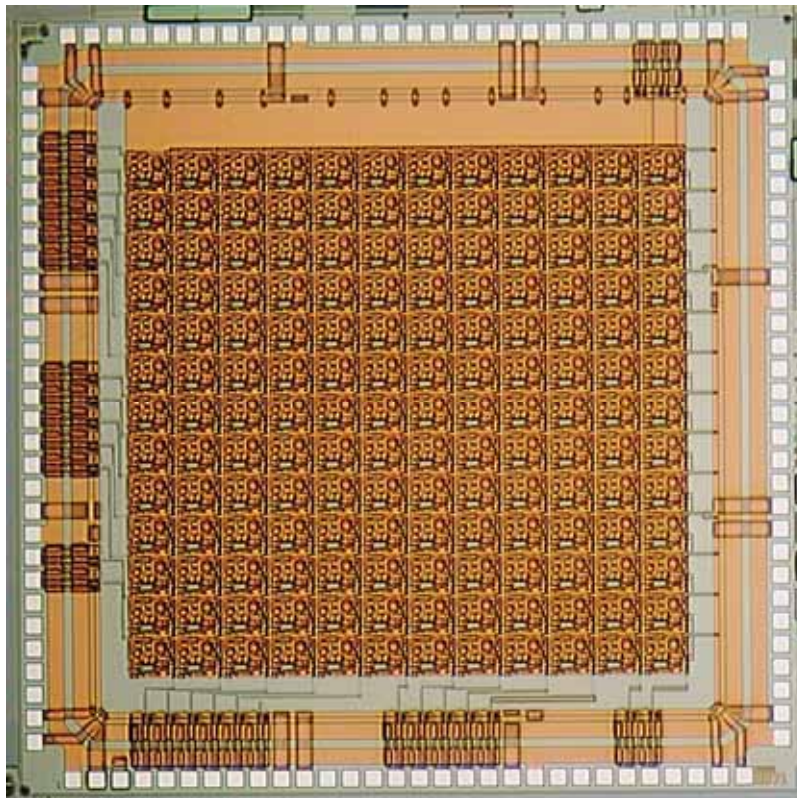
- CMOS 0.6 μ m 3A1
- 64 × 57画素
- 消滅点を4進木探索
 - 嗚呼、4進木探索...

抵抗網の電位分布と物体検出



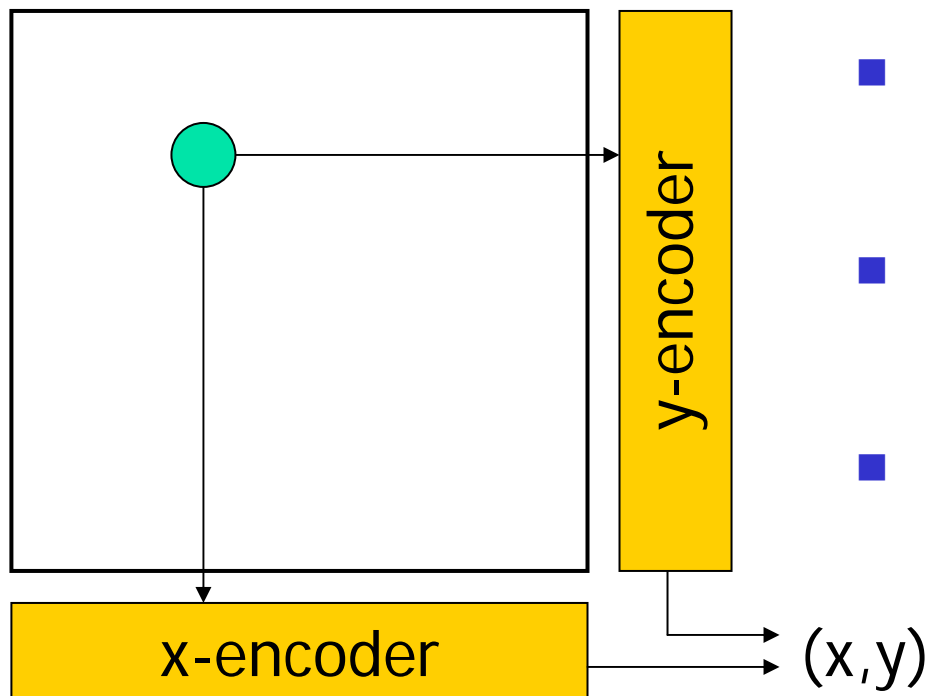
- 抵抗網の電位分布の「極大点」を「中心」とみなす
- 極大点を、画素ごとのコンパレータで並列比較

抵抗網電位分布型物体検出 VisionChip (試作)



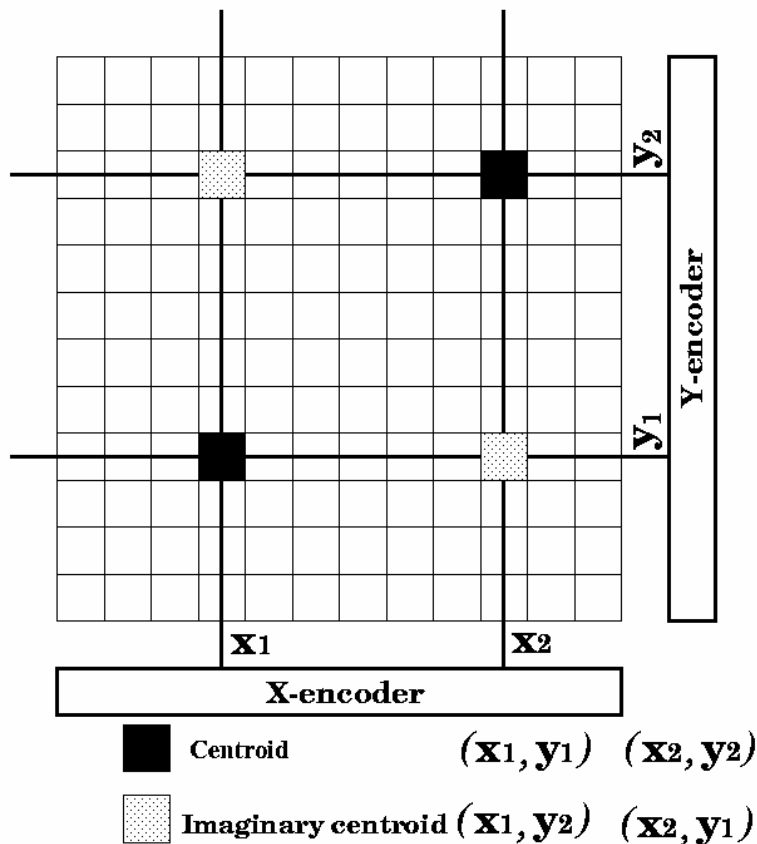
- CMOS 0.6 μm 3A1
- 23 \times 23画素

物体の「位置」の生成



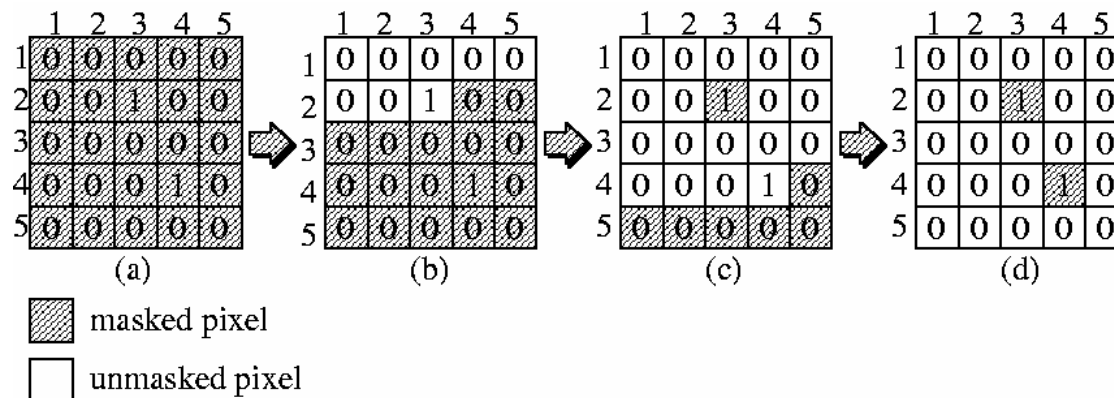
- 画像の特徴量としての「位置」=「座標」
- x 軸・ y 軸に射影してエンコーダで座標生成
- 射影 = x 方向・ y 方向の論理和

複数点に対する座標生成



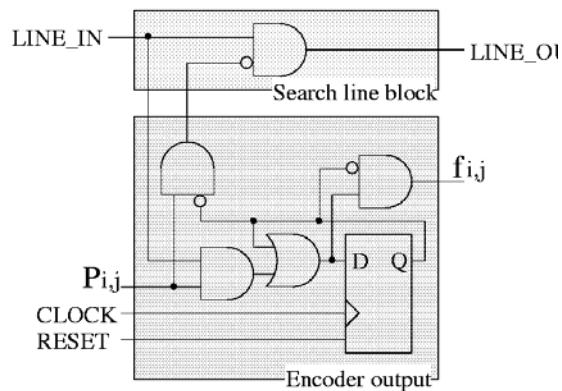
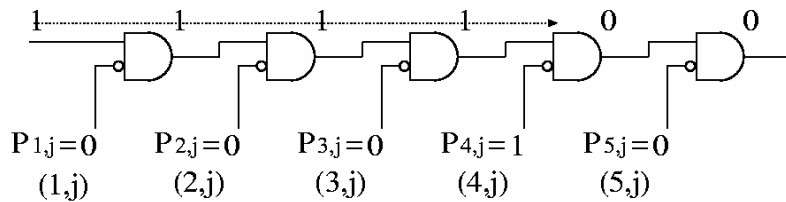
- 対象点が複数ある場合には、射影像からもとの位置を特定できない
- 「同時には1点しか見えないように」すればいい

順次マスク法



- 対象点を、1つずつ、順番に「見える」ようにする
- ただし対象点の探索をクロック同期で逐次探索しては無意味...

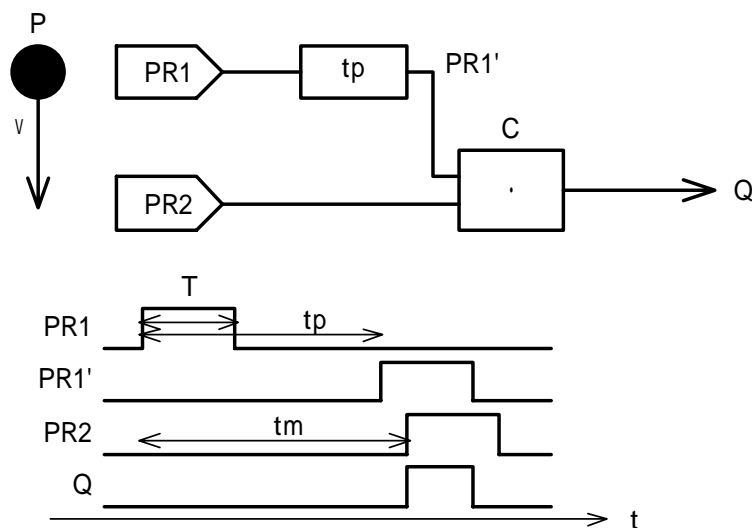
順次マスク法の探索



- 対象点の探索をクロックに非同期で行う
 - クロックに同期しない探索信号が「流れていく」(ゲート遅延時間のみ)
 - 対象点に「ぶつかったら」先には流れていかない
 - エンコードがすんだら、再び「流す」

(前置き)

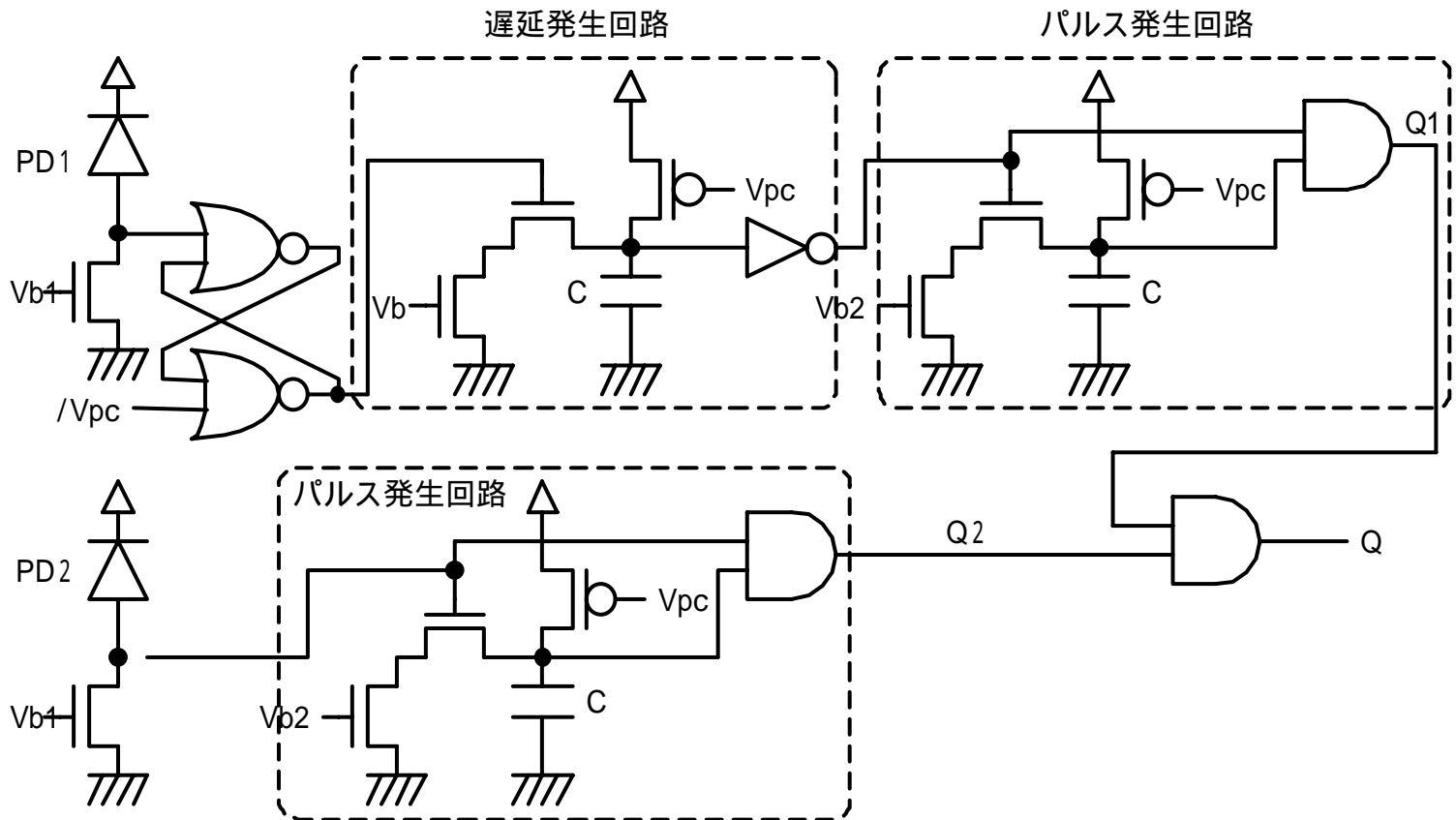
空間伝播時間型動き検出VisionChip



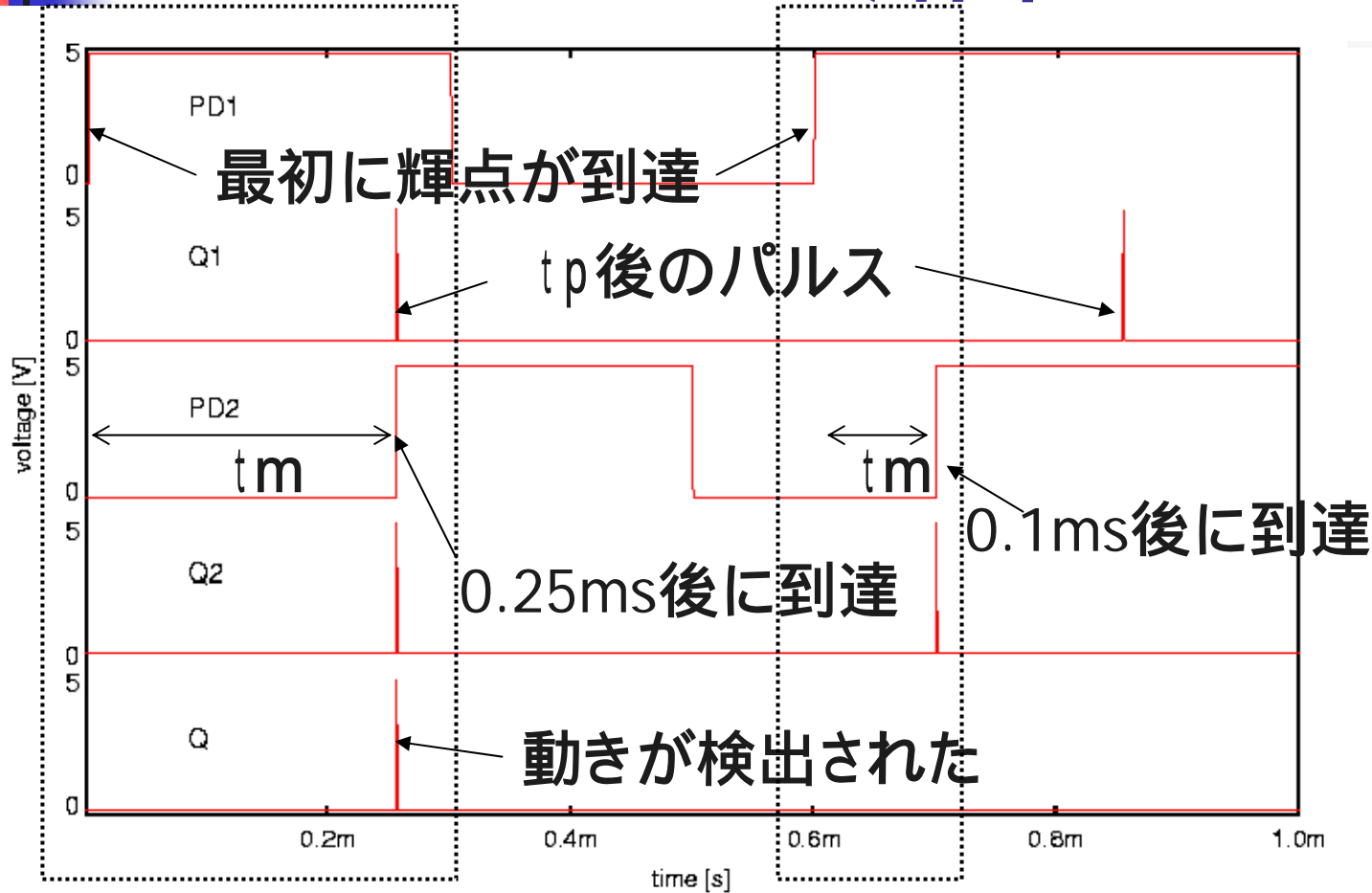
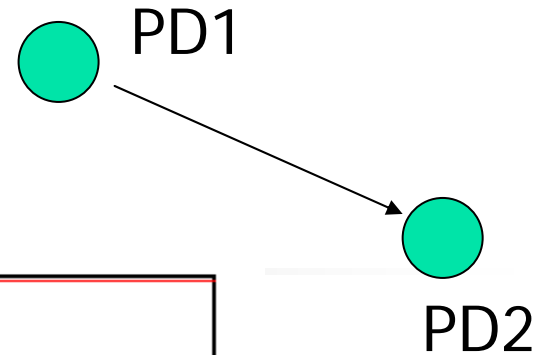
- 空間的に離れた2つの画素を、動く輝点が横切る
- 2つの時間差から動きの検出
 - 動きによる時間差
 - 遅延回路による時間差

空間伝播時間型

動き検出VisionChipの画素回路



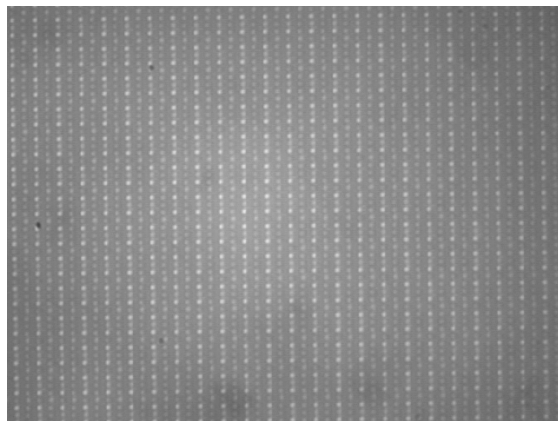
シミュレーション結果



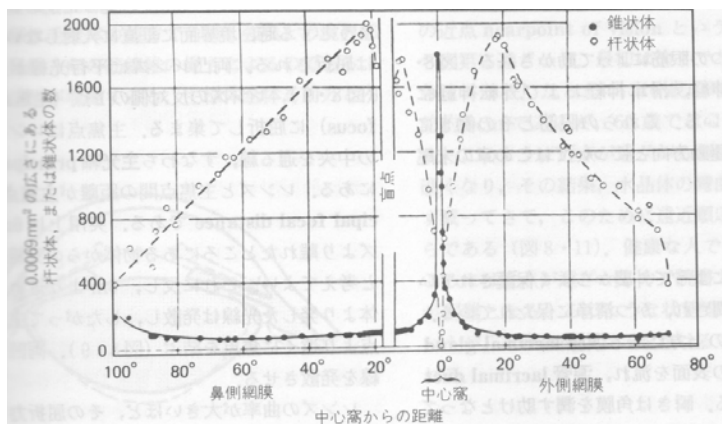
1回目の輝点
($t_m = 0.25\text{ms}$)

2回目の輝点
($t_m = 0.1\text{ms}$)

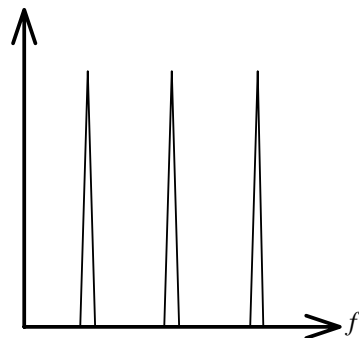
画素の配置・・・



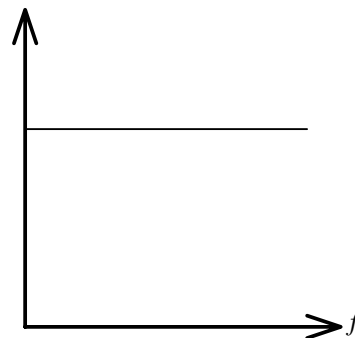
- 普通のイメージセンサの画素の配置は「格子状」
- 網膜上の受光細胞の配置は「不規則」



画素配置の空間スペクトル



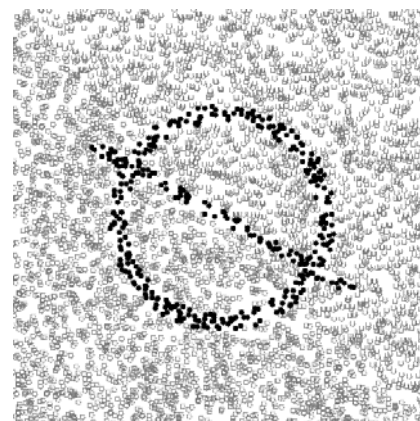
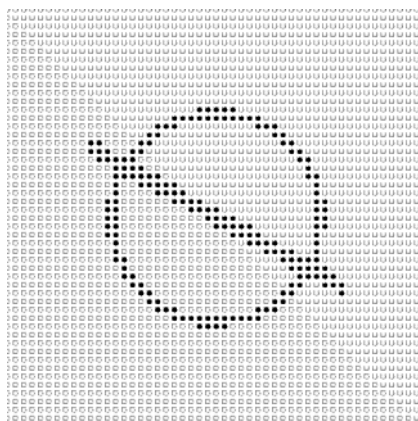
(a)規則配置の場合



(b)不規則配置の場合

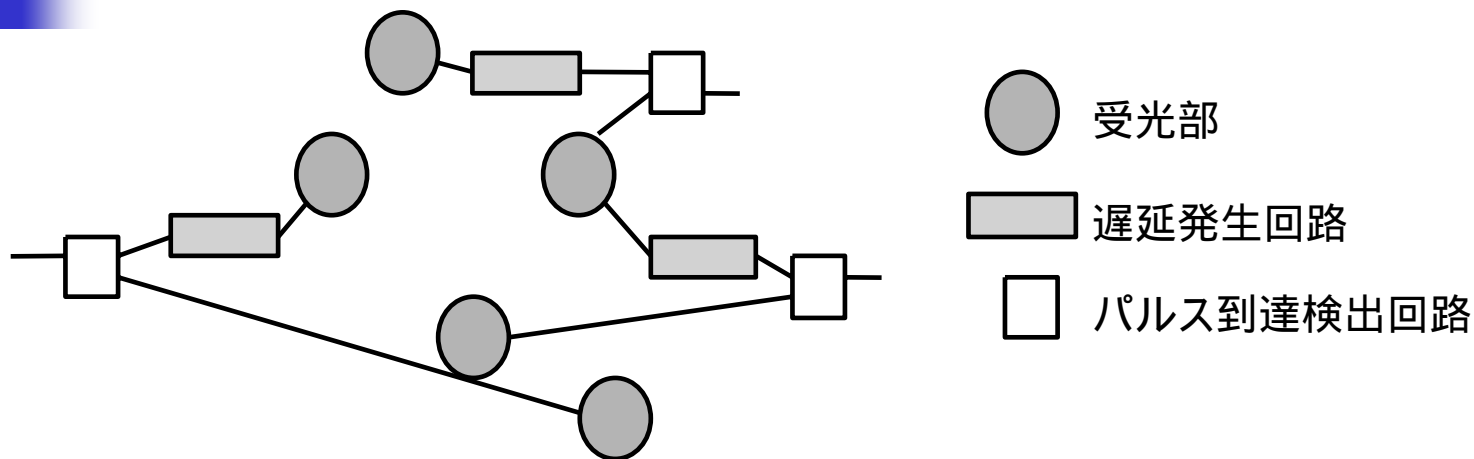
- 規則配置の空間スペクトルはパルス状
 - 表現可能な図形に周期・方向の特異性が強い
- 不規則配置の空間スペクトルはホワイトノイズ状
 - 表現可能な図形に周期・方向の自由度が高い

不規則画素配置VisionChip



- 表現図形は、画素配置での「近似」
- 例えば正方格子の画素配置では、斜め線や曲線がギザギザになる
- 不規則画素配置では、平均的な空間解像度が下がるが、自然な表現

任意方向動き検出VisionChipの 可能性



- 検出できる輝点の動きの方向と速度が、PD1・PD2の「空間的配置」と「遅延 t_p 」によってのみ決まる
- 画素配置の自由度が高い
- 検出結果の出力方法は今後の課題



生体様信号処理

- Neural Networkがベースのものが多い
 - Neuron (PE)のネットワークで、分散処理
 - 個々のNeuron (PE)は、ある程度バラツキがあり不正確
 - Network全体で相互補完してそれらしく動作



デジタル vs アナログ (一般論)

■ デジタル

- 精度が高い
- 素子のバラツキに強い(飽和領域)
- 高速化がしやすい(byスケーリング)
- 設計方法論が(ほぼ)確立している

■ アナログ

- 精度が低い
- 素子のバラツキに弱い
- 高速化がしにくい(スケーリングが効きにくい)
- 設計方法論が(あまり)確立していない



デジタル vs アナログ(別の見方)

■ デジタル

- 末端まで品質管理されたシステム
- 構成要素の「反乱」には極めて無力

■ アナログ

- 末端の特性のバラツキを想定したシステム
- 構成要素の「反乱」にも、システムとして「ある程度」「それなりに」対応できる



Rough Information Processing

- 「ぱっと見」の延長
- たしかに精度の高い処理は必要だが、どこまで必要？
 - $1.1 \times 1.1 = 1.21$ 1.2に四捨五入
 - 不必要な精度で演算する無駄・・・？
 - 例) ゲノムのホモロジー検索(類似度判定)
 - 類似度30%でも10%でも、「似ていない」のは同じ
 - 「似ていない」のは詳細演算はしないのもアリ



レポート

- 自分の修士までの研究と、博士課程での研究との関連について、そのいきさつや考えるところも含めてまとめてください。